UNIVERSITY OF CAMBRIDGE

FUNCTIONAL INEQUALITIES IN QUANTUM INFORMATION THEORY

Cambyse Rouzé

Pembroke College, University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy.

November 2018

À mes parents.

Contents

Contents

Contents

Summary

Functional inequalities constitute a very powerful toolkit in studying various problems arising in classical information theory, statistics and many-body systems. Extensions of these tools to the noncommutative setting have been introduced in the beginning of the 90's in order to study the asymptotic properties of certain quantum Markovian evolutions. In this thesis, we study various extensions and problems arising from the specific noncommutative nature of such processes.

The first logarithmic Sobolev inequality to be proved, due to Gross, was for the Ornstein Uhlenbeck semigroup, that is the Brownian motion with friction on the real line. The generalization of this result to the quantum Ornstein Uhlenbeck semigroup was found very recently by Carlen and Maas, and de Palma and Huber by means of different techniques. The latter proof consists of a quantum generalization of the so-called entropy power inequality. Here, we consider another possible version of the entropy power inequality and use it to derive asymptotic properties of the frictionless quantum Brownian motion.

The proof of Carlen and Maas discussed in the previous paragraph relies on their new quantum extension of the classical notion of displacement convexity. This is classically known to imply most of the usual functional inequalities such as the modified logarithmic Sobolev inequality and Poincaré's inequality. Here, we further study the framework introduced by Carlen and Maas. In particular, we show how displacement convexity implies quantum functional and transportation cost inequalities. The latter are then used to derive certain concentration inequalities of quantum states in the spirit of Bobkov and Götze. These concentration inequalities are used in order to derive finite sample size bounds for the task of quantum parameter estimation.

The main advantage of classical logarithmic Sobolev inequalities over other methods resides in their tensorization property: the strong log-Sobolev constant of the product of independent Markovian evolutions is equal to the maximum over the set of strong log-Sobolev constants of the individual evolutions. However, this property is strongly believed to fail in the non-commutative case, due to the non-multiplicativity of noncommutative $\mathbb{L}_p \to \mathbb{L}_q$ norms. In this thesis, we show tensorization of the logarithmic Sobolev constants for the simplest quantum Markov semigroup, namely the generalized depolarizing semigroup. Moreover, we consider a new general method to overcome the issue of tensorization for general primitive quantum Markov semigroups by looking at their contractivity properties under the completely bounded $\mathbb{L}_p \to \mathbb{L}_q$ norms. This method was first investigated in the restricted case of unital semigroups by Beigi and King.

Noncommutative functional inequalities considered in the present literature only deal with primitive quantum Markovian semigroups which model memoryless irreversible dynamics converging to a specific faithful state. However, quantum Markov semigroups can in general display a much richer behavior referred to as decoherence: In particular, under some mild conditions, any such semigroup is known to converge to an algebra of observables which effectively evolve unitarily. Here, we introduce the concept of a decoherence-free logarithmic Sobolev inequality, and the related notion of hypercontractivity of the associated evolution, to study the decoherence rate of non-primitive quantum Markov semigroups. Moreover, we utilize the transference method recently introduced by Gao, Junge

and LaRacuente, in order to find decoherence times associated to a class of decoherent Markovian evolutions of great importance in the field of quantum error protection, namely collective decoherence semigroups.

Finally, we develop the notion of quantum reverse hypercontractivity, first introduced by Cubitt, Kastoryano, Montanaro and Temme in the unital case, and apply it in conjunction with the tensorization of the modified logarithmic Sobolev inequality for the generalized depolarizing semigroup in order to find strong converse rates in quantum hypothesis testing and for the classical capacity of classical-quantum channels. Moreover, the transference method also allows us to find strong converse bounds on the various capacities of quantum Markovian evolutions.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution except as declared in the Preface and specified in the text.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor, Nilanjana Datta, for guiding me throughout my PhD. I am particularly grateful to her for always making herself available, for our many blackboard discussions, and for her constant encouragement. I would also like to thank the Department of Pure Mathematics and Mathematical Statistics, as well as the Centre for Quantum Information and Foundations in the Department of Applied Mathematics and Theoretical Physics, and Pembroke College, for providing funding as well as travel grants that enabled me to attend various conferences and workshops.

The research that led to this thesis is the product of fruitful collaborations with Ivan Bardet, Salman Beigi, Nilanjana Datta, Eric Hanson, Marius Junge, Nicholas Laracuente, Felix Leditzky, Yan Pautrat and Daniel Stilck França. It was a pleasure to work and co-author papers with them, and I am looking forward to further extending my collaboration with each of them.

I also greatly enjoyed discussions with the following people, who, in one way or another, all contributed to my understanding of quantum information theory: Simon Andreys, Koji Azuma, Tristan Benoist, Mario Berta, Ángela Capel Cuevas, Hao-Chung Cheng, Min-Hsiu Hsieh, Aldéric Joulin, Richard Jozsa, Michael J. Kastoryano, Robert König, Will Matthews, Milán Mosonyi, Ion Nechita, Māris Ozols, Clément Pellegrini, Sergii Strelchuk, Michał Studzinski, Mark Wilde and Andreas Winter. I would also like to address special thanks to our crew of PhD students and Postdocs, including Ivan Bardet, Johannes Bausch, Felix Leditzky, Eric Hanson, Sathyawageeswar Subramanian and Mithuna Yoganathan, to whom I am greatly thankful for all the discussions on science related (and unrelated) topics, and for their support in times of need. They played an important role in creating a great environment which I had the privilege of being a part of.

Apart from the quantum information theory community, I could share my time in and out of Cambridge with many great people. My heartfelt thanks go to my dear friends Tiphaine Blay, Yacine Debbabi, Vera Di Stefano, Denis Kuzzay, Tristan Leclercq, Pablo Marcello Baquero, Lars Mennen, Vincent Midavaine, Simone Parisotto, Federico Pasqualotto, Andrea Rodríguez Martínez, Mark Rowland, Sam Thomas and Angela Wittman.

I would like to end these acknowledgments by thanking my dear family. To my parents Paul Rouzé and Mehrara Mehrgan Rouzé first, for providing me with the unlimited love, encouragement and moral support, I am forever indebted. Their constant search for new adventures and challenges have deeply inspired me throughout my life. To Kourosch and Kathia who always believed in me more than myself; to Anahita and Guillaume, whose home in our beloved hometown they always kept open; and to Darius and Maximilien, who I had the privilege to see growing up into the kindest young boys during these past four years: Je vous suis profondément reconnaissant.

Introduction

Ideal quantum systems undergo a unitary evolution. These closed evolutions are the ones traditionally considered in quantum computing. However, in realistic situations, the interactions between a system and its environment lead to a strong entanglement between them. This in turn results in the dynamical destruction of the initial entanglement between different subsystems of the system under consideration. This phenomenon, usually referred to as *environment induced decoherence* (EID), was introduced in the 70's by the German physicist Dieter Zeh as an attempt to solve the longstanding measurement problem. For example, [Joos and Zeh, 1985] showed that the "non-diagonal" elements (in the position basis) of the reduced density matrix of a particle subject to scattering effects of its environment vanish exponentially fast. In this case, *effective observables* (e.g. operators that are diagonal in the position basis) become fixed points of the evolution as time increases.

More precisely, let $\mathcal{H}_S$ be the Hilbert space corresponding to a quantum system S, and let $\mathcal{H}_E$ denote the Hilbert space of the environment E of S. Assume that at time $t = 0$, S and E are prepared in a pure uncorrelated state:

$$\psi_{SE} = \psi_S \otimes \psi_E \,.$$

Next, denote by H_{SE} the Hamiltonian modeling the interaction between E and S. In full generality, H_{SE} can be written as the sum of three terms:

$$H_{SE} = H_S \otimes \mathbb{1}_E + \mathbb{1}_S \otimes H_E + \lambda V_{SE} \,.$$

The first term H_S models the local interactions occurring within S alone, whereas H_E denotes the interactions within the environment. The last term H_{SE} results from the coupling of the system and its environment. The parameter $\lambda > 0$ is usually referred to as the *coupling strength*. Then, the state of the system at time $t > 0$ is given by

$$\rho_S(t) := \mathrm{Tr}_E \left(\mathrm{e}^{itH_{SE}} \left(\psi_S \otimes \psi_E \right) \mathrm{e}^{-itH_{SE}} \right) , \tag{-1.1}$$

where the partial trace Tr_E models the operation of discarding the state of the environment. Hence, apart from some trivial situations (e.g. $\lambda = 0$), the purity of the initial state ψ_S is decreases: $\rho_S(t)$ is a density operator on $\mathcal{H}_S$, that is a positive operator whose trace is equal to 1. Moreover, the energy of S is no longer preserved: $\mathrm{Tr}(\rho_S(t) H_S) \neq \langle \psi_S, H_S \psi_S \rangle$.

The analysis of (-1.1) as a function of time is in general difficult to carry out due to possible memory effects arising from the interaction term H_{SE}. The situation becomes much more tractable when considering that the coupling strength is small. In this case, the evolution (-1.1) can be approximated with the one of a memoryless system (Markovian approximation):

$$\rho_S(t) \sim \tilde{\rho}_S(t) \equiv \mathcal{P}_{t*}(\psi_S) \,.$$

When looking at the dual Heisenberg picture, that is the evolution of quantum observables, the maps

$(\mathcal{P}_t)_{t \geq 0}$ define a *quantum Markov semigroup* (QMS). This is a family of completely positive, unital maps $\mathcal{P}_t : \mathcal{B}(\mathcal{H}_S) \to \mathcal{B}(\mathcal{H}_S)$ on the algebra $\mathcal{B}(\mathcal{H}_S)$ of linear operators on $\mathcal{H}_S$ that satisfy the following semigroup property: for any $t, s \geq 0$,

$$\mathcal{P}_{t+s} = \mathcal{P}_t \circ \mathcal{P}_s .$$

In this thesis, we undertake a quantitative analysis of the asymptotic behavior of *quantum Markovian evolutions*. In order to achieve this goal, we introduce new non-commutative extensions of a set of powerful classical tools known as *functional inequalities*.

This contribution is by no means complete, nor the first attempt of the sort. In what follows, we provide a brief summary of already known results in this direction, and explain how the new results obtained as part of this thesis complement them.

Quantum functional inequalities for primitive semigroups

The investigation of non-commutative functional inequalities started almost at the same time as their classical analogues [Gross, 1975a, Lindsay, 1990, Carlen and Loss, 1993, Carlen and Lieb, 1993, Biane, 1997]: Let $(\mathcal{P}_t)_{t \geq 0}$ be a QMS on the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on a finite dimensional Hilbert space $\mathcal{H}$, and denote by $\mathcal{L} := \frac{d\mathcal{P}_t}{dt}\big|_{t=0}$ its generator. Assume moreover that there exists a unique full-rank state σ that is invariant under the evolution generated by $\mathcal{L}$. In this case, the semigroup is said to be *primitive*. Primitive semigroups model the simplest type of decoherence, where initial observables evolve towards their average computed in the invariant state σ. Equivalently, the algebra of effective observables is simply equal to the trivial algebra $\mathbb{C} \mathbb{1}_{\mathcal{B}(\mathcal{H})}$.

Perhaps the easiest example of a functional inequality is the *Poincaré inequality*: there exists a positive constant λ such that, for all $X \in \mathcal{B}(\mathcal{H})$,

$$\lambda \operatorname{Var}_\sigma(X) \leq \mathcal{E}_{2,\mathcal{L}}(X) := -\langle X, \mathcal{L}(X) \rangle_\sigma , \tag{PI}$$

where $\langle A, B \rangle_\sigma := \operatorname{Tr}(\sigma^{1/2} A^* \sigma^{1/2} B)$, and $\operatorname{Var}(X) := \langle X - \operatorname{Tr}(\sigma X), X - \operatorname{Tr}(\sigma X) \rangle_\sigma$. Denoting by $\mathbb{L}_2(\sigma)$ the Hilbert space corresponding to this inner product, PI simply implies the following exponential $\mathbb{L}_2(\sigma)$-convergence of any evolved observable $X_t := \mathcal{P}_t(X)$ towards the expected value $\operatorname{Tr}(\sigma X)$ of the initial observable X:

$$\operatorname{Var}(X_t) \leq \mathrm{e}^{-2\lambda t} \operatorname{Var}(X) .$$

In the case when the semigroup is symmetric with respect to $\langle ., . \rangle_\sigma$ (KMS-*symmetry*), the best constant $\lambda(\mathcal{L})$ achieving the bound in PI is simply the *spectral gap* of $\mathcal{L}$, that is the absolute value of the second (necessarily negative) highest eigenvalue of $\mathcal{L}$.

The Poincaré inequality provides a theoretically simple way of estimating the convergence of a quantum Markov semigroup towards an invariant state σ. For this, we introduce the family of non-commutative weighted $\mathbb{L}_p(\sigma)$ spaces, $p \geq 1$, with norm given by

$$\|X\|_{\mathbb{L}_p(\sigma)} := \left(\operatorname{Tr} |\sigma^{\frac{1}{2p}} X \sigma^{\frac{1}{2p}}|^p \right)^{\frac{1}{p}} .$$

Then, for a given initial state ρ, and denoting by $\rho_t := \mathcal{P}_{t*}(\rho)$ the state evolved up to time t, we get

$$\|\rho_t - \sigma\|_1 \equiv \|X_t - \mathbb{1}\|_{\mathbb{L}_1(\sigma)} \leq \|X_t - \mathbb{1}\|_{\mathbb{L}_2(\sigma)} \leq \mathrm{e}^{-\lambda(\hat{\mathcal{L}})t} \|X - \mathbb{1}\|_{\mathbb{L}_2(\sigma)} \leq \sqrt{\|\sigma^{-1}\|_\infty} \, \mathrm{e}^{-\lambda(\mathcal{L})t},$$

where $X_t := \sigma^{-\frac{1}{2}} \rho_t \sigma^{-\frac{1}{2}} \equiv \hat{\mathcal{P}}_t(X)$ can be interpreted as the density of ρ_t with respect to σ, $\hat{\mathcal{P}}_t$ denoting the dual of $\mathcal{P}_t$ with respect to $\langle .,. \rangle_\sigma$ so that $\lambda(\mathcal{L}) = \lambda(\hat{\mathcal{L}})$. However, estimating the spectral gap of $\mathcal{L}$ can turn out to be a difficult problem (see e.g. [Kastoryano and Brandão, 2016]). Moreover, the bounds found from this method are often loose. Fortunately, tighter bounds can be achieved from more elaborate techniques.

The notion of *hypercontractivity* and *logarithmic Sobolev inequalities* was fully extended to the non-commutative framework by [Olkiewicz and Zegarlinski, 1999]: A *quantum logarithmic Sobolev inequality* of order 2, defined in Section 7.2, can be introduced as follows: for any state ρ,

$$\alpha_2 \, D(\rho \| \sigma) \le \mathcal{E}_{2,\mathcal{L}}(X) \,, \tag{LSI$_2$}$$

where $X = \sigma^{-\frac{1}{4}} \rho^{\frac{1}{2}} \sigma^{-\frac{1}{4}}$ can be interpreted as the square root of the non-commutative density of ρ with respect to σ, and the logarithmic Sobolev constant $\alpha_2 \equiv \alpha_2(\mathcal{L})$ is the best constant achieving the bound in LSI$_2$. Alternatively, the quantum *modified logarithmic Sobolev inequality* is defined as follows [Kastoryano and Temme, 2013, Carbone and Martinelli, 2015]: for any initial state ρ,

$$4\alpha_1 \, D(\rho \| \sigma) \le \mathrm{EP}_\sigma(\rho) := -\mathrm{Tr}(\mathcal{L}_*(\rho)(\ln \rho - \ln \sigma)) \,, \tag{MLSI}$$

and the modified logarithmic Sobolev constant $\alpha_1(\mathcal{L})$ is the best constant achieving the bound in MLSI. Under the *quantum detailed balance condition*

$$\mathrm{Tr}(\sigma A \mathcal{L}(B)) = \mathrm{Tr}(\sigma \mathcal{L}(A) B) \tag{σ-DBC}$$

for any $A, B \in \mathcal{B}(\mathcal{H})$, the Stroock-Varopoulos inequality proved in Part III implies the so-called $\mathbb{L}_1$-regularity:

$$\mathrm{EP}_\sigma(\rho) \ge 2\,\mathcal{E}_{2,\mathcal{L}}(X) \quad \Rightarrow \quad \alpha_1(\mathcal{L}) \ge \frac{\alpha_2(\mathcal{L})}{2} \,. \tag{-1.2}$$

In practice, the constants $\alpha_1(\mathcal{L})$ and $\alpha_2(\mathcal{L})$ can be very different from each other: in the case of the *generalized depolarizing semigroup* $(\mathcal{P}_t^{\mathrm{depol}})_{t \ge 0}$, defined for any $X \in \mathcal{B}(\mathbb{C}^d)$ and all $t \ge 0$ as

$$\mathcal{P}_t^{\mathrm{depol}}(X) = \mathrm{e}^{-t} \, X + (1 - \mathrm{e}^{-t}) \, \mathrm{Tr}(\sigma X) \,, \tag{-1.3}$$

the constant $\alpha_1(\mathcal{L}^{\mathrm{depol}})$ was computed in [Müller-Hermes et al., 2016], whereas the constant $\alpha_2(\mathcal{L}^{\mathrm{depol}})$ is found in Theorem 7.2.4. In particular, $\alpha_1(\mathcal{L}^{\mathrm{depol}}) \ge \frac{1}{4}$ whereas $\alpha_2(\mathcal{L}^{\mathrm{depol}}) \sim \ln(d)^{-1}$ when $\sigma \equiv d^{-1} \mathbb{1}$. Mixing times can be found by noticing that

$$\mathrm{EP}_\sigma(\rho) = -\frac{dD(\rho_t \| \sigma)}{dt}\bigg|_{t=0} \quad \Rightarrow \quad D(\rho_t \| \sigma) \le \mathrm{e}^{-4\alpha_1(\mathcal{L})t} \, D(\rho \| \sigma) \,. \tag{-1.4}$$

This convergence in relative entropy typically provides much better bounds than the one provided by a simple Poincaré inequality. Indeed, by Pinsker's inequality:

$$\|\rho_t - \sigma\|_1 \le \sqrt{2D(\rho_t \| \sigma)} \le \mathrm{e}^{-2\alpha_1(\mathcal{L})t} \sqrt{2\ln \|\sigma^{-1}\|_\infty} \,.$$

The question of the usefulness of the constant $\alpha_2(\mathcal{L})$ can be posed. First, we recall that in the case of classical diffusions, $\alpha_2 = 2\alpha_1$. A good reason for the introduction of α_2 in the quantum (and classical discrete) setting is because it is usually easier to compute than α_1. Moreover, it was shown in [Temme et al., 2014] that $\alpha_2(\mathcal{L})$ is always strictly positive in finite dimensions, with a lower

bound depending on the smallest eigenvalue of σ and the spectral gap $\lambda(\mathcal{L})$. This and (-1.2) directly implies the positivity of $\alpha_1(\mathcal{L})$ for semigroups satisfying a detailed balance condition, and justifies the introduction of LSI_2.

Functional inequalities for non-primitive semigroups

The study of convergence to equilibrium of classical and quantum Markovian evolutions, and their related functional inequalities, usually assume the condition of *primitivity* of the semigroup: there exists a unique full-rank invariant state σ towards which the semigroup converges. One of the major contributions of this thesis is the study of the convergence of *non-primitive* semigroups [Blanchard and Olkiewicz, 2003, Deschamps et al., 2016, Carbone et al., 2013].

For sake of simplicity, we assume that the semigroup $(\mathcal{P}_t)_{t\geq 0}$ is *faithful*, which means that it admits at least one full-rank invariant state. This condition certifies the existence of a conditional expectation $E_{\mathcal{N}} : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ projecting any operator X on the algebra of effective observables $\mathcal{N}$ (see Section 0.1.4 for more details). For instance, assume that the system under consideration is constituted of n qubits which are subject to the exact same noise. This leads to an overall invariance of the evolution under permutation of the qubits that is usually referred to as *collective decoherence*. More simply, the depolarizing semigroup of Equation (-1.3) can be extended to the non-primitive case as follows: let $\mathcal{N}$ be a matrix subalgebra of $\mathcal{B}(\mathbb{C}^k)$, and $E_{\mathcal{N}} : \mathcal{B}(\mathbb{C}^k) \to \mathcal{N}$ be a conditional expectation onto $\mathcal{N}$. Then, the *simple* quantum Markov semigroup $(\mathcal{P}_t^{\mathcal{N}})_{t\geq 0}$ associated to $\mathcal{N}$ is defined for any $X \in \mathcal{B}(\mathbb{C}^k)$ and $t \geq 0$ by

$$\mathcal{P}_t^{\mathcal{N}}(X) := \mathrm{e}^{-t} X + (1 - \mathrm{e}^{-t}) E_{\mathcal{N}}[X].$$

This semigroup converges to $\mathcal{N}$ as $t \to \infty$. We recover the depolarizing semigroup by choosing $\mathcal{N} = \mathbb{C}\mathbb{1}_{\mathcal{B}(\mathbb{C}^k)}$ and $E_{\mathcal{N}}[.] = \mathrm{Tr}(\sigma\,.)\,\mathbb{1}_{\mathcal{B}(\mathbb{C}^k)}$.

The modified logarithmic Sobolev inequality MLSI was extended to this framework by [Bardet, 2017]: for any state ρ,

$$4\alpha_1 D(\rho\|E_{\mathcal{N}*}(\rho)) \leq \mathrm{EP}_{\sigma_{\mathrm{Tr}}}(\rho)\,, \tag{DF-MLSI}$$

where the invariant state σ_{Tr} appearing on the right hand side of DF-MLSI is defined as $\sigma_{\mathrm{Tr}} := k^{-1} E_{\mathcal{N}*}(\mathbb{1})$. If DF-MLSI is satisfied, an argument similar to the one of (-1.4) leads to the exponential convergence in relative entropy of the solution ρ_t towards $E_{\mathcal{N}*}(\rho_t)$. Once again, the question of the positivity of the optimal constant $\alpha_1(\mathcal{L})$ for any evolution occurring on a finite dimensional Hilbert space can be raised. Interestingly enough, the situation turns out to be very different from the primitive case.

In Chapter 8, we extend the logarithmic Sobolev inequality of order 2 to the non-primitive setting. A natural way of proceeding is by replacing the relative entropy on the left-hand side of LSI_2 by the relative entropy between ρ and its projection $E_{\mathcal{N}*}(\rho)$. However, we show in Chapter 8 that this inequality is never satisfied in a truly non-primitive case unless $\alpha_2 = 0$. A way to overcome this issue is through a weakening the inequality by adding an additional term on its right-hand side:

$$D(\rho\|E_{\mathcal{N}*}(\rho)) \leq c\,\mathcal{E}_{2,\mathcal{L}}(X) + \frac{d}{2}\,\|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}^2\,, \tag{DF-LSI$_2$}$$

for some $c > 0$, $d > 0$. Since $d > 0$, the Stroock-Varopoulos argument employed in the primitive case in order to get a lower bound on the rate of convergence in (-1.4) fails. Fortunately, one can employ a different technique in order to estimate the speed of decoherence in terms of the constants c and d. This

method uses the equivalence between the logarithmic Sobolev inequality and the hypercontractivity of the semigroup that still pertains in the non-primitive setting.

A primitive QMS $(\mathcal{P}_t)_{t\geq 0}$ with associated invariant state σ is said to be *hypercontractive* if for any $t \geq 0$, and $\ln(p(t)-1) := \frac{4t}{c}$,

$$\|\mathcal{P}_t : \mathbb{L}_2(\sigma) \to \mathbb{L}_{p(t)}(\sigma)\| \leq \exp\left\{2d\left(\frac{1}{2} - \frac{1}{p(t)}\right)\right\}. \qquad \text{(HC)}$$

The non-commutative Gross lemma [Olkiewicz and Zegarlinski, 1999] states that hypercontractivity is equivalent to the defective logarithmic Sobolev inequality DF-LSI$_2$ for primitive QMS satisfying σ-DBC. In the non-primitive case, the concept ot hypercontractivity needs to be defined with respect to a different family of norms, call it $\mathbb{L}_p(\mathcal{N})$, depending on the algebra $\mathcal{N}$. In particular, the following requirements should be fulfilled:

- $\mathbb{L}_p(\mathcal{N})$ should reduce to $\mathbb{L}_p(\sigma)$ for primitive QMS with unique full-rank invariant state σ;

- $\|X\|_{\mathbb{L}_p(\mathcal{N})} = \|X\|_{\mathbb{L}_2(\mathcal{N})}$ for all $X \in \mathcal{N}$, which implies that no further contraction can happen in $\mathcal{N}$;

- the QMS $(\mathcal{P}_t)_{t\geq 0}$ should contract from $\mathbb{L}_2(\mathcal{N})$ to $\mathbb{L}_p(\mathcal{N})$.

It turns out that the *amalgamated $\mathbb{L}_p$ norms* introduced in [Junge and Parcet, 2010] are well-suited. In Chapter 8, we extend Gross' lemma by showing a weak equivalence between hypercontractivity with respect to these amalgamated norms and DF-LSI$_2$. Assuming that the semigroup is hypercontractive with respect to these norms, we show the following convergence bound:

$$\|\mathcal{P}_{t*}(\rho - E_{\mathcal{N}*}(\rho))\|_1 \leq \max_{i\in I}\sqrt{d_{\mathcal{H}_i}}\,\ln(\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty)^{\frac{\lambda c}{4}}\,\mathrm{e}^{1+d-\lambda(\mathcal{L})t},$$

as long as one assumes that $\dim(\mathcal{H}) \geq 3$. The coefficients $d_{\mathcal{H}_i}$ refer to the dimensions of the blocks appearing in the block decomposition of the algebra $\mathcal{N}$ (see Equation (0.9)). We also show universal lower bounds on the constants c and d in the spirit of [Temme et al., 2014]. Notice however that the problem of finding a universal lower bound on the constant $\alpha_1(\mathcal{L})$ in DF-MLSI is still open.

Computing the exact constants c and d in DF-LSI$_2$ is in general difficult. In Chapter 9, we show how one can get estimates on decoherence times by showing that, given a faithful QMS that is reversible with respect to the completely mixed state, there exists a classical Markov semigroup whose constants control the ones of the original QMS. This so-called *transference method* was introduced in [Gao et al., 2018b] who were exclusively concerned with the transfer of classical diffusions. Here, we broaden their scope to incorporate evolutions on finite groups. One important class of QMS for which the method applies is the one of collective decoherence already mentioned above. For those evolutions, we show that the *decoherence time*, that is the time it takes for the evolution to approximately reach $\mathcal{N}$:

$$\tau_{\mathrm{deco}}(\varepsilon) = \inf\left\{t \geq 0\,;\,\|\mathcal{P}_{t*}(\rho) - E_{\mathcal{N}} \circ \mathcal{P}_{t*}(\rho)\|_1 \leq \varepsilon \quad \forall \rho\right\},$$

is independent of the size of the system: for instance, consider the Lindblad generator of the *weak collective decoherence semigroup* on $\mathcal{B}((\mathbb{C}^2)^{\otimes n})$:

$$\mathcal{L}_n^{\mathrm{wcd}}(X) := \Sigma_z^n\,X\,\Sigma_z^n - X, \qquad \text{where} \qquad \Sigma_z^n := \sum_{i=1}^{n}\mathbb{1}_{\mathbb{C}^2}^{\otimes i-1} \otimes \sigma_z \otimes \mathbb{1}_{\mathbb{C}^2}^{n-i}.$$

In this example, the algebra $\mathcal{N}$ of effective observables coincides with the *fixed point algebra*

$$\mathcal{F}(\mathcal{P}^{\mathrm{wcd},n}) = \{X\,:\,\forall t \geq 0,\,\mathcal{P}_t^{\mathrm{wcd},n}(X) = X\}.$$

We show in Chapter 8 that the following estimate holds:

$$\|\mathcal{P}_t^{\mathrm{wcd},n} - E_{\mathcal{F}} : \mathbb{L}_1(d_{\mathcal{H}}^{-1}\mathbb{1}_{\mathcal{H}}) \to \mathbb{L}_1(d_{\mathcal{H}}^{-1}\mathbb{1}_{\mathcal{H}})\| \le \|P_t^{\mathrm{heat}} - \mathbb{E}_\lambda : \mathbb{L}_1(\mathbb{T}^1) \to \mathbb{L}_1(\mathbb{T}^1)\|,$$

where P_t^{heat} denotes the classical heat semigroup on the circle $\mathbb{T}^1$ which converges to the Lebesgue measure λ. Moreover, the latter was in shown e.g. in [Saloff-Coste, 1994] to satisfy the following bound

$$\|P_t^{\mathrm{heat}} - \mathbb{E}_\lambda : \mathbb{L}_1(\mathbb{T}^1) \to \mathbb{L}_1(\mathbb{T}^1)\| \le \sqrt{2 + \sqrt{\pi/t}}\; \mathrm{e}^{-\frac{t}{2}} \; .$$

This implies that the weak collective decoherence converges exponentially fast to its fixed point algebra independently of the size n of the system.

Tensorization

The great advantage of classical logarithmic Sobolev inequalities over kinds of functional inequalities resides in their *tensorization property*: the logarithmic Sobolev constants α_1, α_2 of n copies of a Markovian evolution are equal to the ones of a single copy. One easy way to understand this is by noticing that, in the commutative case, operator norms are multiplicative. This simple fact no longer holds true in the non-commutative setting. In Chapter 10, we find a uniform bound on $\alpha_1(\mathcal{L}^{\mathrm{depol}})$ for the generalized depolarizing semigroups and their tensor powers. The proof of this result is a generalization of the proof of a similar result in the classical case [Mossel et al., 2013]. We also show that the constant $\alpha_2(\mathcal{L}^{\mathrm{depol}})$ tensorizes in the qubit case.

In [Beigi and King, 2016], the authors proposed to define the hypercontractivity property in terms of the *completely bounded norms*, which are known to be multiplicative even in the non-commutative framework, and proved that it is equivalent to the so-called notion of a complete logarithmic Sobolev inequality for a primitive QMS $(\mathcal{P}_t)_{t\ge0}$. This provides a way to recover the tensorization property in the non-commutative framework. In practice, this amounts to embedding the QMS $(\mathcal{P}_t)_{t\ge0}$ into the non-primitive faithful QMS $(\mathrm{id}_k \otimes \mathcal{P}_t)_{t\ge0}$ on $\mathcal{B}(\mathbb{C}^k \otimes \mathcal{H})$ for each $k \in \mathbb{N}$, and study the latter's hypercontractivity properties. Let σ be the unique invariant state of $(\mathcal{P}_t)_{t\ge0}$. Then $\mathcal{N}_k := \mathcal{B}(\mathbb{C}^k) \otimes \mathbb{1}_{\mathcal{H}}$ and $\sigma_{\mathrm{Tr}} = \frac{\mathbb{1}_k}{k} \otimes \sigma$. Since we proved the impossibility of a positive logarithmic Sobolev constant of order 2 for these evolutions in Chapter 8, we are entitled to introduce a weak constant d similarly to DF-LSI$_2$ in our definition of the complete logarithmic Sobolev inequality. This is problematic since, contrary to the constant c, the constant d is additive under tensorization.

Concentration of quantum states

Concentration of measure is the phenomenon according to which almost all the points of a metric probability measure space (Ω, d, μ) are close to a subset of positive measure. One way of deriving such inequalities is from *transportation-cost inequalities*: such an inequality is said to hold if there exists a constant $c > 0$ such that for any probability measure $\nu << \mu$,

$$W_1(\mu, \nu) \le \sqrt{2c\,\mathrm{Ent}\left(\frac{d\nu}{d\mu}\right)},$$

where $W_1(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \int d(x,y)\, d\pi(x,y)$, and the minimization is over all the probability measures on $\Omega \times \Omega$ having μ and ν as marginals. The transportation-cost inequality was shown by [Marton, 1986] to encode the concentration properties of the measure μ. In particular, it implies the existence of a constant C such that, given any Borel set $A \subset \Omega$ of measure $\mu(A) \ge \frac{1}{2}$, and any

r-enlargement $A_r \coloneqq \{x \in \Omega,\, d(x, A) \le r\}$,

$$\mu(A_r) \le 1 - \mathrm{e}^{-Cr^2}\,.$$

The original proof of Marton is of a probabilistic nature. Later, a functional analytical proof was proposed by [Bobkov and Goetze, 1999] who used the following dual formulation of the Wasserstein distance:

$$W_1(\mu, \nu) = \sup_{\|f\|_{\mathrm{lip}} \le 1} \int f\, d\mu - \int f\, d\nu\,,$$

where the supremum is taken over the set of functions $f : \Omega \to \mathbb{R}$ of Lipschitz constant $\|f\|_{\mathrm{lip}} \coloneqq \sup_{x \ne y} |f(x) - f(y)|/d(x, y)$ less than 1. Following the approach of [Bobkov and Goetze, 1999], we prove a concentration phenomenon for quantum states in Section 12.7: first, we define the following quantum Wasserstein distance: given two states ρ, σ,

$$W_1(\rho, \sigma) \coloneqq \sup_{\|X\|_{\mathrm{Lip}} \le 1} |\operatorname{Tr}(X(\rho - \sigma))|\,,$$

where the supremum is taken over all operators $X \in \mathcal{B}(\mathcal{H})$ of quantum Lipschitz constant less than one. Here, the Lipschitz constant is defined as follows:

$$\|X\|_{\mathrm{Lip}} \coloneqq \|[L, X]\|_\infty\,,$$

for some fixed operator $L \in \mathcal{B}(\mathcal{H})$. The above commutator replaces the finite difference in the definition of $\|f\|_{\mathrm{lip}}$. Given a state σ on $\mathcal{H}$, we then define the *quantum transportation-cost inequality* as follows: there exists a constant c such that, for any state ρ of support included in the one of σ,

$$W_1(\rho, \sigma) \le \sqrt{2\, c\, D(\rho \| \sigma)}\,. \tag{TC$_1$}$$

Similarly to their classical counterparts, we show that such inequalities can be for instance derived whenever a primitive quantum Markov semigroup of invariant state σ satisfies a logarithmic Sobolev inequality. The link between the semigroup chosen and the above inequality is through the definition of the Lipschitz constant, where the operators L are chosen as the Lindblad operators associated to the evolution. Moreover, this implies the following Gaussian concentration of the state σ: for any observable X,

$$\operatorname{Tr}\left(\sigma\, \mathbb{1}_{[r, \infty)}(X - \operatorname{Tr}(\sigma X))\right) \le \exp\left(\frac{-r^2}{2\, c\, \|(\sigma^{-\frac{1}{2}} X \sigma^{\frac{1}{2}})\|_{\mathrm{Lip}}^2}\right)\,, \tag{Gauss}$$

which means that the operator X is concentrated around its expected value in the state σ. Similarly, we show that a weaker concentration can be obtained from a Poincaré inequality.

Non-commutative curvature and displacement convexity

The convergence to equilibrium of a diffusion process on a smooth Riemannian manifold M can be described by the properties of the underlying manifold. This was first observed by [Bakry and Émery, 1985] who showed e.g. that the logarithmic Sobolev constant is lower bounded by the curvature of M. Extending this geometric interpretation to the case of Markovian evolutions on spaces of less regularity has been a subject of focus in the classical community over the past two decades. A milestone was

reached with the appearance of articles establishing a connection between Bakry-Émery's criterion and the well established theory of optimal transport [McCann, 1997, Jordan et al., 1998, Otto and Villani, 2000a]. In these papers, the focus was shifted from the geometry of the underlying sample space, say a smooth Riemannian manifold M, to the set of probability measures $\mathcal{P}(M)$ on M. The latter is given the structure of a metric space by means of the *Wasserstein distance* of order 2: for any $\mu, \nu \in \mathcal{P}(M)$,

$$W_2(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \left(\int_{\mathcal{M}} d(x, y)^2 d\pi(x, y) \right)^{1/2},$$

where $\Pi(\mu, \nu)$ denotes the set of probability measures on $M \times M$ of marginals μ and ν. [Benamou and Brenier, 2000] showed (originally in the case of $M = \mathbb{R}^d$) that the metric space $(\mathcal{P}_2(M), W_2)$ can be provided with a *weak Riemannian structure*:

$$W_2(\mu, \nu) = \inf \left(\int_0^1 \int_M |v_t|^2 \, d\mu_t \, dt \right)^{\frac{1}{2}}, \tag{-1.5}$$

where $\mathcal{P}_2(M)$ is the set of probability measures on which W_2 takes finite values, and where the infimum is taken over distributional solutions (μ_t, v_t) of the following *continuity equation*:

$$\frac{d}{dt}\mu_t + \nabla.(v_t \, \mu_t) = 0. \tag{-1.6}$$

Using [Benamou and Brenier, 2000]'s formulation of the Wasserstein distance, a diffusion semigroup can be canonically defined as the *gradient flow* of the relative entropy functional in the space $(\mathcal{P}_2(M), W_2)$. Bakry and Émery's conditions were then reinterpreted in terms of the convexity properties of the latter along geodesic paths in $\mathcal{P}_2(M)$. The mesoscopic notion of transportation of *masses* in $\mathcal{P}(M)$ being more robust than the local notion of transportation of *points* in M, this new approach lead to a particularly prolific generalization of Bakry-Émery's criterion to non-smooth metric-measure length spaces. Since then, the original framework of Bakry and Émery was further extended to the case of Markov chains over discrete sets (e.g. lattice spin systems) by [Maas, 2011, Erbar and Maas, 2012]. This framework is briefly described in Chapter 4.

Recently, [Carlen and Maas, 2014, Carlen and Maas, 2017] introduced an extension of the latter in the context of primitive quantum Markov semigroups. In Chapter 12, we extend their notion of *quantum geodesic convexity* to the case of primitive QMS acting on the algebra $\mathcal{B}(\mathcal{H})$ of linear operators acting on a finite dimensional Hilbert space $\mathcal{H}$ satisfying the detailed balance condition: for any constant speed geodesic $(\gamma(s))_{s \in [0,1]}$ in $\mathcal{D}(\mathcal{H})$,

$$D(\gamma(s) \| \sigma) \leq (1-s)D(\gamma(0) \| \sigma) + sD(\gamma(1) \| \sigma) - \frac{\kappa}{2}s(1-s) W_{2,\mathcal{L}}(\gamma(0), \gamma(1))^2, \qquad \text{(Displ. Conv.)}$$

where the quantum Wasserstein distance $W_{2,\mathcal{L}}$ is defined in [Carlen and Maas, 2017] similarly to Equation (-1.5). Moreover, we show that Displ. Conv. implies a quantum version of the celebrated HWI *inequality*, according to which, for any full-rank state ρ on $\mathcal{H}$:

$$D(\rho \| \sigma) \leq W_{2,\mathcal{L}}(\rho, \sigma)\sqrt{\mathrm{EP}_\sigma(\rho)} - \frac{\kappa}{2}W_{2,\mathcal{L}}(\rho, \sigma)^2. \tag{HWI}$$

We show that, in the case of $\kappa > 0$, HWI implies that the modified logarithmic Sobolev constant $\alpha_1(\mathcal{L})$ defined through MLSI is lower-bounded by $\frac{\kappa}{2}$.

In the case when Displ. Conv. holds with constant $\kappa = 0$, we show that a Poincaré inequality can be retrieved, with constant depending on the diameter of the set of states under the Wasserstein distance $W_{2,\mathcal{L}}$. A similar conclusion can be reached concerning the modified logarithmic Sobolev

inequality in the case of a doubly stochastic QMS. We end the chapter in Section 12.8 with a preliminary discussion on extensions of the framework to the infinite dimensional setting of continuous variable quantum systems.

Entropy power inequality and isoperimetry

Perhaps the first logarithmic Sobolev inequality to be discovered was the one of the classical Ornstein Uhlenbeck semigroup modeling the behavior of the velocity of a massive Brownian particle under the influence of friction:

$$\text{Ent}(f^2) \le 2 \int_{\mathbb{R}} |\nabla f|^2 d\mu_G \,, \tag{Gross}$$

where $\text{Ent}(f) := \int f \ln f d\mu - \int f d\mu \ln \int f d\mu$, and where μ is the standard Gaussian measure on $\mathbb{R}^n$. The original proof of (Gross) by [Gross, 1975b] used the equivalence with the hypercontractivity proved before by [Nelson, 1973a]. A few decades earlier, [a.J. Stam, 1959] had proved the superadditivity of the entropy power function, also known as the *entropy power inequality*: given two independent random variables X and Y on $\mathbb{R}^n$

$$N(X + Y) \ge N(X) + N(Y) \,, \tag{EPI}$$

where $N(X) = \exp(2S(X)/2)$, $S(X) := -\int f_X(x) \ln f_X(x) d^n x$ being the differential Shannon entropy of the signal X of corresponding probability distribution function f_X. The entropy power inequality was in fact originally proposed by [Shannon, 1948] as a way to establish the capacity region of the Gaussian broadcast channel. It can be proved to be equivalent to the following *entropy convex combination inequality*: for any two independent random variables X and Y on $\mathbb{R}^n$, and any $\lambda \in [0,1]$:

$$S(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \ge \lambda S(X) + (1-\lambda)S(Y) \,.$$

In words, the entropy of a mixture of two independent signals X and Y is greater than the average entropy of X and Y. More than thirty years after its original proof by Stam, [Carlen, 1991] showed that EPI implies Gross. This alternative proof makes use of the so-called *entropic isoperimetric inequality*: given a random variable X on $\mathbb{R}^n$ of associated density f_X:

$$\text{I}(f_X)N(X) \ge 2\pi \,\text{e}\, n \,, \tag{Isop.}$$

which can be derived from (EPI). Here, $\text{I}(f_X) := \int \frac{|\nabla f_X|^2}{f_X} = -\int f_X \Delta \ln f_X$ is the Fisher information of the random variable X.

Recently, the entropy power inequality has been a subject of focus in the quantum community [Koenig and Smith, 2014, Audenaert et al., 2016, Carlen et al., 2016]. In analogy with their classical counterparts, the *quantum entropy power* and *divergence-based quantum Fisher information* of a "well-enough behaved" state ρ on $\mathbb{L}_2(\mathbb{R}^n)$ are defined as follows:

$$N(\rho) := \text{e}^{S(\rho)/n} \,, \qquad J(\rho) := \sum_{j=1}^{n} \text{Tr}(\rho\,[P_j, [P_j, \ln(\rho)]]) + \text{Tr}(\rho\,[Q_j, [Q_j, \ln(\rho)]]) \,,$$

where $S(\rho) := -\text{Tr}(\rho \ln \rho)$ denotes the entropy of ρ, and the operators $P_j : \psi \mapsto -i\partial_{x_j}\psi$, $Q_j : \psi \mapsto (x \mapsto x_j\psi(x))$, are the usual momentum and position operators. In Chapter 11, we prove the following

quantum isoperimetric inequality:

$$J(\rho)\,N(\rho) \geq 2\,\mathrm{e}\,n\,. \tag{qIsop.}$$

We also show how this inequality provides convergence times for the quantum heat semigroup whose generator is defined in Equation (5.44) as:

$$\mathcal{L}^{\mathrm{qheat}}\left(\cdot\right) := -\frac{1}{4}\sum_{j=1}^{2n}\left[Q_j,[Q_j,\cdot]\right] + \left[P_j,[P_j,\cdot]\right]. \tag{-1.7}$$

Here, the double commutators should be interpreted as non-commutative second order differentiations in the directions P_j and Q_j. We show that any state ρ_t of finite first and second moments converges to its *Gaussification*, that is a Gaussian state ρ_t^G of same first and second moments as ρ_t, polynomially fast: for any $\varepsilon \in (0,1)$ there exists $t_\varepsilon > 0$ as well as $\alpha_\varepsilon > 0$ such that for any $t \geq t_\varepsilon$:

$$D(\rho_t\|\rho_t^G) \leq \alpha_\varepsilon\ t^{\varepsilon-1}\,. \tag{-1.8}$$

Similarly to [Carlen, 1991], one motivation behind the derivation of qIsop. was to find the logarithmic Sobolev constant of the *quantum Ornstein Uhlenbeck semigroup* whose generator takes the following form:

$$\mathcal{L}^{\mathrm{qOU}}(X) = -\sum_{i=1}^{n}\left([Q_i,[Q_i,X]] + [P_i,[P_i,X]]\right) - \left(X + i\,[Q_iXP_i - P_iXQ_i]\right)\,. \tag{-1.9}$$

This generator, which models the evolution of a two-energy levels atom which traverses a photonic cavity, reduces to the one of the classical Ornstein Uhlenbeck semigroup when restricting its action to the "diagonal" of functions of the position operators. However, qIsop. turns out to be the wrong generalization of the isoperimetric inequality to serve that purpose. More recently, [De Palma and Huber, 2018] observed that the original *quantum-quantum entropy power inequality* proved by [Koenig and Smith, 2014] does the job. We recall their proof for sake of completeness in Section 11.3.

Applications to quantum information theory

Part V is devolved to applications of the results obtained in the previous chapters to quantum information theory. These results are organized in three categories:

Applications to quantum statistics Initiated in the mid 60's, the field of quantum statistics deals with the problem of extracting information from a particular quantum system (its quantum state, its evolution, etc.) from the analysis of the data obtained after performing certain quantum measurements. The importance of this basic task can be observed by noticing that it is at the core of many more complex quantum information processing protocols. Perhaps one of the first important discoveries was the generalization of *Stein's lemma* to the quantum setting: assume that n identical and independent copies of a quantum system have been prepared in either a state ρ or a state σ. Bob's task is to infer which state the system has been prepared in by means of a quantum measurement, also referred to as a *test*. Two kinds of error can be made: either Bob infers the system is in the state σ when it is in state ρ, or vice versa. The corresponding probabilities of error are respectively called the *type I error* and *type II error*, and are denoted as follows:

$$\alpha(T_n) = \mathrm{Tr}(\rho^{\otimes n}(\mathbb{1}_{\mathcal{H}^{\otimes n}} - T_n))\,, \qquad \beta(T_n) = \mathrm{Tr}(T_n\sigma^{\otimes n})\,,$$

where $0 \leq T_n \leq \mathbb{1}$ models any allowed test performed on the n received copies. Then, for any $\varepsilon > 0$,

$$- \lim_{n \to \infty} \frac{1}{n} \ln \min_{0 \leq T_n \leq \mathbb{1}_{\mathcal{H}^{\otimes n}}} \{ \beta(T_n) : \alpha(T_n) \leq \varepsilon \} = D(\rho \| \sigma) .$$

This is the content of Stein's lemma, which provides the quantum relative entropy with an operational interpretation. This means that the smallest error rate that one can hope to achieve for the type II error while imposing the type I error to stay lower than a threshold ε is given by the quantum relative entropy. While the above limit provides a good understanding of the asymptotic problem, finding finite n bounds on the type II error may be more relevant to tackle practical situations where only finite resources are available. Fortunately, a recent classical method [Liu et al., 2017] based on the *reverse hypercontractivity* for the random walk on the complete graph can be extended to the quantum regime and yields the following *finite sample size* strong converse bound: for any $n \in \mathbb{N}$ and any test $0 \leq T_n \leq \mathbb{1}_{\mathcal{H}^{\otimes n}}$ such that the type II error $\beta(T_n)$ decreases at a rate $r > D(\rho \| \sigma)$, the type I error grows exponentially fast with a rate given by

$$\alpha(T_n) \geq 1 - \mathrm{e}^{-nf}, \tag{-1.10}$$

where $f = \left(\sqrt{\gamma + (r - D(\rho \| \sigma))} - \sqrt{\gamma} \right)^2$, and hence tends to zero in the limit of $r \to D(\rho \| \sigma)$ (see Section 13.1.2). The same method can also be used to derive similar strong converse bounds on the classical capacity of classical-quantum channels (see Section 15.3).

Another important task in quantum statistics is the one of the estimation of a quantum state indexed by a real parameter θ. Once again, we assume that n copies of the same state ρ_θ are being produced. Then, an estimator is described by a sequence of positive operator valued measurements (POVMs in short) $\mathbf{M} := \{ M^{(n)} \}_{n \in \mathbb{N}}$, where, for each n, $M^{(n)} : \mathcal{B}(\mathbb{R}) \mapsto \mathcal{P}(\mathcal{H}^{\otimes n})$ is a POVM on the Hilbert space $\mathcal{H}^{\otimes n}$ of the n systems, $\mathcal{B}(\mathbb{R})$ standing for the Borel algebra associated to $\mathbb{R}$. The merit of such a sequence can be quantified in terms of the following error exponent (see [Hayashi, 2002, Nagaoka, 2005, Masahito, 2005]):

$$\beta(\mathbf{M}, \theta, \varepsilon, n) := -\frac{1}{n \varepsilon^2} \log \mathbb{P}_{M^{(n)}} \left(\hat{\theta}_n \in [\theta - \varepsilon, \theta + \varepsilon]^c \right) ,$$

where $\mathbb{P}_{M^{(n)}} \left(\hat{\theta}_n \in [\theta - \varepsilon, \theta + \varepsilon]^c \right) := \mathrm{Tr} \left(M^{(n)} ([\theta - \varepsilon, \theta + \varepsilon]^c) \rho_\theta^{\otimes n} \right)$ is the probability that the estimated value $\hat{\theta}_n$ is at least ε away from the true parameter θ. In the asymptotic setting $n \to \infty$, it was shown in Lemma 14 of [Hayashi, 2002] that, under some technical assumptions, any POVM $\mathbf{M}$ satisfies

$$\limsup_{\varepsilon \to 0} \limsup_{n \to \infty} \beta(\mathbf{M}, \theta, \varepsilon, n) \leq \frac{\mathrm{I}_{\mathrm{SLD}}(\theta)}{2} , \tag{-1.11}$$

where $\mathrm{I}_{\mathrm{SLD}}(\theta) := \mathrm{Tr} \left(\rho_\theta (L_\theta^{\mathrm{SLD}})^2 \right)$ is the quantum symmetric logarithmic derivative (SLD for short) Fisher information (see Section 1.3). This bound was also shown to be saturated for a sequence of projective-valued measurements $\mathbf{M}_\theta$ associated to the self-adjoint operator

$$X_\theta^{(n)} := \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}^{\otimes (k-1)} \otimes \left(\frac{L_\theta^{\mathrm{SLD}}}{\mathrm{I}_{\mathrm{SLD}}(\theta)} + \theta \, \mathbb{1} \right) \otimes \mathbb{1}^{\otimes (n-k)} , \tag{-1.12}$$

where the estimated value $\hat{\theta}_n$ is determined to be the outcome of the measurement $M_\theta^{(n)}$. Using the concentration of quantum states arising from the modified logarithmic Sobolev inequality for the tensor product of generalized depolarizing semigroups converging to $\rho_\theta^{\otimes n}$, we show in Section 13.2 that any estimator constructed from such a family of observables of the form of $X^{(n)}$, with identical local

terms X, is such that the following lower bound holds for any integer n:

$$\beta(\mathbf{M}, \theta, \varepsilon, n) \geq \frac{1}{2 \, \|\rho_\theta^{-1/2} \, X \, \rho_\theta^{1/2}\|_{\text{Lip}}^2} \, . \tag{-1.13}$$

Applications to entanglement theory In various information processing tasks (compression, communication, etc.), the advantage of quantum protocols over their classical analogues lies in the possibility of encoding information in *entangled states*. We recall that a *separable state* on a bipartite system $\mathcal{H}_A \otimes \mathcal{H}_B$ is one that can be written as the mixture of product states:

$$\rho_{AB} = \sum_{i=1}^r p_i \, \rho_A^{(i)} \otimes \rho_B^{(i)} \, . \tag{-1.14}$$

Any state that cannot be decomposed as in Equation (-1.14) is called *entangled*. One of the main goals of quantum information processing is to figure out ways to create and manipulate quantum states without losing their entanglement properties. Therefore, *entanglement breaking channels*, i.e. quantum channels that only output separable states when acting on one half of a bipartite quantum state, represent a kind of noise that a quantum system needs to be protected from in such non-classical protocols. In Chapter 14, we provide estimates on the time it takes for a continuous or discrete time Markovian evolution applied to one half of a bipartite quantum system to ensure that it becomes entanglement breaking.

Our upper bounds are based on the observation that full rank product quantum states lie in the relative interior of the set of separable quantum states, as already proved in [Lami and Giovannetti, 2016]. We then use techniques similar to those of [Gurvits and Barnum, 2002] to obtain estimates on the radius of the separable ball around such states in different metrics. Combining these with the tools to estimate decoherence time of Markovian evolutions studied in Chapter 8, we obtain estimates on how long it takes for all outputs to be in the separable ball.

To derive lower bounds, we exploit the fact that quantum channels that only output separable states remain positive maps when composed with partial transposition. Thus, if we can show that the output of a state under the channel has a negative partial transposition, the channel still preserves some entanglement. Applying this reasoning with the maximally entangled state as an input, we are then able to obtain criteria based on the spectrum of the quantum channel to certify that it still preserves some entanglement. Unlike our upper bounds, we do not make any assumptions on the structure of the quantum channels to prove these lower bounds, although we derive specialized versions for quantum channels of particular interest, such as quantum Markov semigroups in continuous time.

Application to estimates on capacities of quantum Markovian evolutions In Chapter 15, we are interested in the estimation of the optimal amount of information that can be sent for different information processing tasks involving quantum inputs and noise. We consider the following tasks. The *capacity* associated to each of these tasks, i.e. the optimal achievable asymptotic rates at which the task can be performed, can be expressed in terms of an appropriate entropic quantity. The main difficulty of quantum channel coding in comparison to its classical analogue lies in the fact that the entropic expressions characterizing most of the capacities are intractable for general quantum channels. For instance, in the case of classical communication over identical uses of a quantum channel Φ, the classical capacity $\mathcal{C}(\Phi)$ of Φ is characterized by the *regularized Holevo information*:

$$\mathcal{C}(\Phi) = \chi_{\text{reg}}(\Phi) \equiv \lim_{n \to \infty} \frac{1}{n} \chi(\Phi^{\otimes n}) \, ,$$

where the *Holevo information* $\chi(\Phi)$ of a quantum channel Φ is defined in Equation (15.2). In general, the regularized Holevo information does not reduce to its *single-letter* expression: $\chi_{\mathrm{reg}}(\Phi) \neq \chi(\Phi)$, in sharp contrast with the classical setting. This is due to the so-called *superadditivity of the Holevo information*: there exist channels Φ_1 and Φ_2 such that

$$\chi(\Phi_1 \otimes \Phi_2) > \chi(\Phi_1) + \chi(\Phi_2).$$

Fortunately, recent progress has been made in finding good *strong converse bounds*[1] on various capacities. In the case when the channel is assumed to arise from a quantum convolution semigroup (see Section 5.5.3), we use the transference methods of Chapter 9 in order to estimate the behavior of the capacity as a function of time.

Main contributions and outline of the thesis

This thesis is based on the following publications and preprints:

[DPR17] Nilanjana Datta, Yan Pautrat and Cambyse Rouzé. "Contractivity properties of a quantum diffusion semigroup". *Journal of Mathematical Physics* 58, 012205 (2017). Accepted at the 20th Annual Conference on Quantum Information Processing.DOI:10.1063/1.4974223.

[DR17a] Nilanjana Datta and Cambyse Rouzé. "Concentration of quantum states from quantum functional and transportation cost inequalities" (2017). Accepted for publication in *Journal of Mathematical Physics.* arXiv:1704.02400.

[DR17b] Nilanjana Datta and Cambyse Rouzé. "Relating relative entropy, optimal transport and Fisher information: a quantum HWI inequality" (2017). arXiv:1709.07437.

[BDR18] Salman Beigi, Nilanjana Datta and Cambyse Rouzé. "Quantum reverse hypercontractivity: its tensorization and application to strong converses" (2017). arXiv:1804.10100. Submitted to *Communications in Mathematical Physics.*

[BR18] Ivan Bardet, Cambyse Rouzé. "Hypercontractivity and logarithmic Sobolev Inequality for non-primitive quantum Markov semigroups and estimation of decoherence rates" (2018). arXiv:1803.05379. Accepted at the 21st Annual Conference on Quantum Information Processing.

[BJLRS18] Ivan Bardet, Marius Junge, Nicholas LaRacuente, Cambyse Rouzé and Daniel Stilck França. "Functional inequalities via group transference techniques and application to estimation of decoherence times and capacities" (2018). Accepted at the 22nd Annual Conference on Quantum Information Processing.

[HRS18] Eric Hanson, Cambyse Rouzé, and Daniel Stilck França. "Estimating when a quantum channel becomes entanglement breaking" (2018).

It is organized in five parts: In Part I, Chapter 0, we briefly recall the main notions of quantum theory, with an emphasis on the Heisenberg picture and the algebra of observables. This rather unconventional choice, as opposed to the more traditional approach focusing on states, can be justified from the major role played by certain algebras of operators throughout this thesis. In Chapter 1, we introduce various measures quantifying the distance between quantum states, observables or channels, namely norms, entropies and Fisher informations. Here the term distance is used in a broad sense.

---

[1]If the probability of error made by trying to achieve a rate that lies above capacity converges to 1 exponentially fast in the limit of infinitely many uses of the channel, the task is said to satisfy a *strong converse property*.

Part II regroups a corpus of techniques that are well-known to the community of classical probabilists, statisticians and mathematical analysts. After a brief review of some basic definitions from the theory of classical Markov processes in Chapter 2, we provide a summary of functional inequalities in the commutative setting in Chapter 3. Connection to optimal transport, and in particular the equivalence between the Ricci curvature lower bound and the displacement convexity of Boltzmann's H entropy is explained in Chapter 4.

In Part III, we introduce quantum Markov semigroups. In Chapter 5, we provide a summary of the main definitions and concepts that mirrors the exposition of Chapter 2. In Chapter 6, we introduce the concept of decoherence for non-primitive Markovian evolutions.

The new contributions provided in this thesis are organized into a mathematical part (Part IV) and an information-theoretic part (Part V). In particular, we prove the following original results:

1 In Chapter 5, we provide in Theorem 5.2.6 a characterization of weak* continuous quantum Markov semigroups on the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators acting on a separable Hilbert space $\mathcal{H}$ that satisfy the detailed balance condition with respect to some faithful normal state σ. In Theorem 5.4.2, we prove a quantum Stroock-Varopoulos inequality for these semigroups. This result appeared in the restricted case of a finite dimensional Hilbert space in [BDR18]. Finally, we derive a new expression for the quantum Ornstein Uhlenbeck semigroup in **??**.

2 In Chapter 7, we find the the exact expression for the logarithmic Sobolev constant of order 2 of the generalized depolarizing semigroup (Theorem 7.2.4). Moreover, we develop the framework of quantum reverse hypercontractivity for primitive finite dimensional quantum Markov semigroups. In particular, we show the equivalence between reverse hypercontractivity and the modified logarithmic Sobolev inequality under the detailed balance condition in Corollary 7.4.10. In Section 13.1, Theorem 13.1.2 and Chapter 15, Theorem 15.3.1, we derive finite sample size strong converse results for the tasks of asymmetric binary hypothesis testing and classical communication over a quantum channel based on the quantum reverse hypercontractivity of the generalized depolarizing semigroup. These results were obtained in collaboration with Salman Beigi and Nilanjana Datta, and appeared in [BDR18].

3 In Theorems 7.5.2 and 7.5.4, we derive a non-commutative Nash inequality and ultracontractivity for the quantum heat semigroup on CCR algebra over a finite dimensional Hilbert space under some positivity conditions on the initial state. There results were obtained in collaboration with Nilanjana Datta and Yan Pautrat, and appeared in [DPR17].

4 In Chapter 8, we introduce the framework of logarithmic Sobolev inequalities and hypercontractivity for non-primitive quantum Markov semigroups acting on the algebra of $n \times n$ complex-valued matrices. In particular, we show the quasi-equivalence between these two concepts in Theorem 8.4.1, and find universal constants by standard interpolation techniques in Corollary 8.4.11. We prove the impossibility of a tight logarithmic Sobolev inequality for a truly non-primitive semigroup in Theorem 8.5.1, which we use to prove the non-uniform convexity of the amalgamated $\mathbb{L}_p$ in Corollary 8.5.2. Finally, decoherence times are given in terms of the structure of the decoherence free subalgebra of the semigroup and its hypercontractivity constants in Proposition 8.6.1. These results were obtained in collaboration with Ivan Bardet, and appeared in [BR18].

5 In Chapter 9, we propose a technique to derive decoherence times for a doubly stochastic, finite dimensional quantum Markov semigroup acting on a projective representation of a group G in

terms of the mixing times of an associated classical Markov semigroup acting on G. These follow directly from the norm estimates of Theorem 9.2.1. In Theorem 15.2.1, we also derive entropic estimates that allow us to bound various capacities for continuous-time quantum channels in Section 15.2. These results were obtained in collaboration with Ivan Bardet, Marius Junge, Nicholas LaRacuente and Daniel Stilck França and will appear in [BJLRS18].

6 Chapter 10 is devoted to the study of tensorization in the non-commutative setting. In Theorem 10.1.1, we derive a constant lower bound on the modified logarithmic Sobolev constant α_1 for the generalized depolarizing semigroup. On the other hand, the tensorization of α_2 for qubits is obtained in Theorem 10.1.4. These results were obtained in collaboration with Salman Beigi and Nilanjana Datta, and appeared in [BDR18].

7 In Chapter 11, we prove a non-commutative entropic isoperimetric inequality (Theorem 11.1.7) for the CCR algebra. The proof relies on a quantum Blachman-Stam inequality which we derive in Theorem 11.1.3. We use the former to prove the convergence of a quantum state ρ_t evolving according the quantum heat semigroup towards a Gaussian state that shares the same mean and variance as ρ_t (Proposition 11.2.5). These results were obtained in collaboration with Nilanjana Datta and Yan Pautrat, and the first two appeared in [DPR17].

8 Chapter 12 is devoted to the analysis of the non-commutative curvature introduced by [Carlen and Maas, 2017] in the context of finite dimensional, primitive quantum Markov semigroups. Different equivalent formulations for displacement convexity are derived in Theorem 12.3.3. These are showed to imply a non-commutative HWI inequality (Theorem 12.5.2), which itself leads to the modified logarithmic Sobolev inequality in positive curvature (Corollary 12.5.3). We also introduce a non-commutative transportation-cost inequality of order 2 and relate it to other quantum functional inequalities in Theorems 12.5.6 and 12.5.10. Theorems 12.6.3 and 12.6.7 are dedicated to the derivation of modified logarithmic Sobolev and Poincaré constants from diameter estimates in the context of non-negative curvature. Finally, we derive quantum concentration inequalities from non-commutative transportation cost and Poincaré inequalities in Theorems 12.7.5 and 12.7.7. These are used to obtain finite blocklength upper bounds on the error probability for the task of parameter estimation of a quantum state in Proposition 13.2.1. These results were obtained in collaboration with Nilanjana Datta and appeared in [DR17a] and [DR17b]. Finally, a rapid excursion to the infinite dimensional setting of the CCR algebra is proposed in Section 12.8.

9 In Chapter 14, we find upper bounds on the time it takes for a quantum Markovian evolution to become entanglement breaking, based on the functional inequalities for non-primitive quantum Markov semigroups obtained in Chapter 8. We also derive matching lower bounds. These results were obtained in collaboration with Eric Hanson and Daniel Stilck França, and will appear in [HRS18].

Part I.

The axioms of quantum theory

Chapter 0.

Quantum open systems

In this first chapter, we introduce the axioms of quantum theory following an algebraic approach. The latter, usually referred to as the *algebraic framework of (quantum) physics*, was pioneered by [Segal, 1947] and further developed by [Haag and Kastler, 1964]. Three reasons can be invoked to justify this choice: firstly, it encapsulates the duality between a state and an observable by providing a very general common mathematical ground on which any physical theory (classical or quantum) can be described. Secondly, the algebraic framework provides an efficient way of extending the notions of differential calculus, geometry and probability to non-commutative settings, and to compare them to their commutative cousins. Thirdly, some quantum systems with infinite degrees of freedom can only be rigorously defined algebraically. Here we provide an outline of the concepts that will enter the rest of this thesis, which will be properly redefined in due time. We do not claim any originality in this section, and our exposition is mainly inspired by [Strocchi, 2008] and [Holevo, 2011].

Layout of the chapter: In Section 0.1, we derive the axioms of quantum theory from an algebraic point of view. In Section 0.2, we discuss an important example of a quantum system, namely a bosonic system which we will come back to in the subsequent parts of the manuscript. We end the chapter in Section 0.3 by introducing basics of operator algebras required in the rest of this thesis.

0.1. Operational derivation of quantum axioms

Classical Hamiltonian systems are described by a set of canonical variables (q, p) called *configuration* (e.g. the position and momentum of a particle) modeled as points of a manifold Ω commonly called a *phase space*. Physical quantities describing the system (e.g. temperature, magnetic field), also referred to as *observables*, are functions $(q, p) \mapsto f(q, p)$ of the configuration. These in particular include the configurations q and p themselves. Observables are usually assumed to belong to the commutative algebra $C(\Omega)$ of real continuous functions on Ω provided with the sup norm $\|.\|_\infty$. Conversely, assuming that Ω is compact, the Stone-Weierstrass theorem implies that any configuration (q, p) can be uniquely determined by the values taken by all the observables at that particular point. This fact is usually known as the *duality relation between states and observables*.

Thinking of the state of a physical system as a configuration in Ω might seem too restrictive: such a framework only allows for the description of theories that can be tested with arbitrarily accurate observations. This working assumption, called the *deterministic description*, makes sense for a large class of phenomena in which random fluctuations can be both practically and theoretically ignored (e.g. planetary motion). However, regardless of the number of successes the deterministic point of view led to in the classical physics of the XVIII-XIXth centuries, it was definitely refuted at the turn of the

XXth with the first experiments occurring at the atomic scale.

One way to allow the implementation of physical theories that are stable against both experimental and theoretical random fluctuations is to replace the intuitive concept of a configuration by the one of a probability measure on the Borel algebra of events on the phase space Ω[1]. Moreover, the Riesz representation theorem states that the information encoded in any such probability measure μ is fully contained in the positive normalized linear functional that it defines,

$$\mathbb{E}_\mu[f] := \int_{C(\Omega)} f \, d\mu, \qquad \forall \text{ observable } f,$$

and which we refer to as an *expectation*. The former notion of a configuration $(q, p) \in \Omega$ is then retrieved when considering the Dirac distribution $\delta_{(q,p)}$. Dirac distributions are special in the sense that they cannot be written as linear combinations of other measures.

The above stochastic description of a physical system is at the core of Boltzmann's statistical mechanics of complex systems, typically constituted of 10^{23} particles. It also suggests the possibility of defining a classical system through an abstract Abelian C*-algebra $\mathcal{A}$ of observables, together with a set of positive normalized linear functionals on $\mathcal{A}$, with no more explicit mention of the underlying phase space. This is due to the Gelfand-Naimark representation theorem which states that any Abelian (unital) C*-algebra $\mathcal{A}$ is isometrically isomorphic to the algebra of complex continuous functions $C(\Omega)$ on a compact Hausdorff topological space Ω. Philosophically, we moved from the traditional geometric description of a classical system in terms of its configuration to an equivalent purely algebraic view point. A relaxation of the condition of commutativity of the algebra $\mathcal{A}$ leads us to the quantum realm.

0.1.1. An operational axiomatization of physics

Here, we proceed to a more systematic mathematical description of a physical system. From an operational point of view, a physical system can be described by a set $\mathcal{S}$ of *states* in which it can be prepared, together with a set $\mathcal{O}$ of measurable quantities called *observables*.

Operational requirements It is clear that for any observable $A \in \mathcal{O}$ and any $\lambda \in \mathbb{R}$, one can define the observable λA measured by rescaling the apparatus by λ (think of a pointer scale). Similarly, A^2 may be interpreted as the observable associated with squaring the apparatus scale, and iteratively one defines the observables $A^{m+n} = A^m A^n$ as well as any polynomial of A. An element A is said to be positive if all the results of the measurements of A are positive numbers. This is equivalent to assuming that $A = B^2$ for some other observable $B \in \mathcal{O}$. We also define the observable A^0 whose results of measurements always give the value 1.

A state $\omega \in \mathcal{S}$ acts on any observable $A \in \mathcal{O}$ in order to provide a corresponding *expectation* $\omega(A)$. This can be operationally understood as the task of taking the average over the results of repeated measurements of A on multiple identical copies of the physical system prepared in the state ω. The latter then defines a real functional on $\mathcal{O}$. We also assume the following natural axioms:

- *Homogeneity*: For any $A, B \in \mathcal{O}$, $\omega \in \mathcal{S}$ and $\lambda \in \mathbb{R}$,

$$\omega(\lambda A) = \lambda \omega(A), \quad \forall \lambda \in \mathbb{R}.$$

---

[1]Here, we implicitly assume the following statistical postulate [Holevo, 2011]: "even if individual results in a sequence of identical, independent realizations of an experiment may vary, the occurrence of one or another result in a long enough sequence of realizations can be characterized by a definite stable frequency".

- *Linearity*: For all $A \in \mathcal{O}$ and m, n integers:

$$\omega\left(A^n + A^m\right) = \omega\left(A^n\right) + \omega\left(A^m\right).$$

- *Observables separate states*: Given two states $\omega_1, \omega_2 \in \mathcal{S}$:

$$\forall A \in \mathcal{O}, \quad \omega_1\left(A\right) = \omega_2\left(A\right) \quad \Rightarrow \quad \omega_1 = \omega_2.$$

This follows from the fact that a state should be operationally characterized by the values of the expectations $\omega\left(A\right)$ for A spanning the set of observables.

- *Completeness of states with respect to observables*: Two observables $A, B \in \mathcal{O}$ sharing the same expectation on all states in $\mathcal{S}$ should not be distinguishable (e.g. two different apparatus may effectively relate to the same observable). This defines an equivalence relation $\sim$ between elements of $\mathcal{O}$:

$$A \sim B \quad \overset{\text{def}}{\Leftrightarrow} \quad \forall \omega \in \mathcal{S}, \quad \omega\left(A\right) = \omega\left(B\right).$$

From now on, we replace the set $\mathcal{O}$ by its associated set of equivalence classes $\mathcal{O}/_\sim$, that is two indistinguishable observables are assumed to be equal.

- *Normalization*: For all $A \in \mathcal{O}$, $\omega\left(A^0\right) = 1$. This together with the previous axiom implies that the 0-th powers of observables all fall in the same equivalence class, which we call the *identity*, denoted by $\mathbb{1}$. Hence, for any state ω,

$$\omega\left(\mathbb{1}\right) = 1.$$

Moreover, the property $A^m A^n = A^{m+n}$ implies that $A\,\mathbb{1} = \mathbb{1}\,A = A$.

- *Positivity*: For any observable $A \in \mathcal{O}$, the positivity of A is equivalent to the requirement that for all state $\omega \in \mathcal{S}$:

$$\omega\left(A\right) \geq 0.$$

- *Mixtures*: Given two states ω_1 and ω_2 and $\lambda \in [0, 1]$, the convex linear combination

$$\omega = \lambda\,\omega_1 + \left(1 - \lambda\right)\omega_2,$$

called *mixture* or *mixed state*, is also a valid state. A state is called *pure* if any decomposition of the above form leads to either $\lambda = 0$ or $\lambda = 1$.

The C*-algebraic framework   Any apparatus is limited by the range of values that it can take. Typically, a pointer state can only reach a bounded region of real numbers. This implies that the results of measurements of an observable $A \in \mathcal{O}$ in any state should yield a bounded number. Hence, the following positive number associated to an observable can be given the operational meaning of being the maximum numerical value that can be displayed on any concrete apparatus measuring A:

$$\|A\| := \sup_{\omega \in \mathcal{S}} |\omega\left(A\right)| < \infty.$$

By homogeneity of the states,

$$\|\lambda A\| = |\lambda|\,\|A\|, \quad \forall \lambda \in \mathbb{R}.$$

By the separability criterion,

$$\|A\| = 0 \quad \Rightarrow \quad A = 0.$$

Moreover, the following identity holds:

$$\|A^2\| = \|A\|^2.$$

Proof. For any state $\omega \in \mathcal{S}$, $\omega(\|A\|\mathbb{1} \pm A) \geq 0$ follows from the definition of $\|A\|$. $\|A\|\mathbb{1} \pm A \geq 0$ by positivity. Hence, $(\|A\|\mathbb{1} - A)(\|A\|\mathbb{1} + A)$ is a positive polynomial of A and

$$\forall \omega \in \mathcal{S}, \quad \|A\|^2 - \omega(A^2) = \omega\left((\|A\|\mathbb{1} - A)(\|A\|\mathbb{1} + A)\right) \geq 0 \quad \Rightarrow \quad \|A\|^2 \geq \|A^2\|.$$

On the other hand, $(\|A\|\mathbb{1} \pm A)^2 = \|A\|^2\mathbb{1} + A^2 \pm 2\|A\|A$ is positive, which implies that for any state ω,

$$2\|A\|\,|\omega(A)| \leq \|A\|^2 + \omega(A^2) \leq \|A\|^2 + \|A^2\| \quad \Rightarrow \quad \|A\|^2 \leq \|A^2\|.$$

$\square$

Next, we motivate the notion of a sum of two observables $A, B \in \mathcal{O}$: for a wide class of pairs of observables $A, B \in \mathcal{O}$ (e.g. kinetic and potential energy) there exists an observable (say the total energy) denoted by $A + B \in \mathcal{O}$, such that

$$\forall \omega \in \mathcal{S}, \quad \omega(A + B) = \omega(A) + \omega(B).$$

In what follows, we assume that to any two observables we can still associate their sum so that the above equation is satisfied. This axiom is physically less motivated than the former ones and constitutes a first non-trivial extrapolation of physics. However, it is a very useful mathematical assumption since it extends the set $\mathcal{O}$ of observables to a vector space, which we still denote by $\mathcal{O}$.

From powers of the sum $A + B$ one can define the *Jordan* (or *symmetric*) product

$$A \circ B := \frac{1}{2}\left((A + B)^2 - A^2 - B^2\right) = B \circ A,$$

so that $A \circ A = A^2$. This product is commutative, but in general not associative. The additional condition that the Jordan product satisfies $A \circ (A^2 \circ B) = A^2 \circ (A \circ B)$ (weak associativity) endows $\mathcal{O}$ with the structure of a *Jordan algebra*. Segal showed that the above structure (together with a few more technical requirements) allows one to recover most of the features of quantum theory. However, the discussion becomes easier if we assume that the elements of the Jordan algebra $\mathcal{O}$ generate a complex associative algebra $\mathcal{A}$ such that:

- The symmetric product arises from the associative product in $\mathcal{A}$:

$$\forall A, B \in \mathcal{O}, \quad A \circ B \equiv \frac{1}{2}(AB + BA)$$

- An involution is defined on $\mathcal{A}$ such that for all $A, B \in \mathcal{O}$,

$$\forall \lambda \in \mathbb{C}, \quad (\lambda A + B)^* = \bar{\lambda} A + B, \quad (AB)^* = BA.$$

- For any $A \in \mathcal{A}$, A^*A is positive and the states can be extended from $\mathcal{O}$ to $\mathcal{A}$ by linearity to linear functionals on $\mathcal{A}$, with the natural extension of the positivity property:

$$\forall A \in \mathcal{A}, \quad \omega(A^*A) \geq 0.$$

For simplicity, we also denote the set of these extended states by $\mathcal{S}$.

- Similarly, the quantity $\|.\|$ originally defined on $\mathcal{O}$ can be extended to a norm on $\mathcal{A}$, also denoted by $\|.\|$:

$$\|A\| := \sup_{\omega \in \mathcal{S}} |\omega(A)|.$$

We assume the following two properties for the norm

$$\forall A, B \in \mathcal{A}, \quad \|AB\| \le \|A\| \|B\|, \quad \|A^* A\| = \|A^*\| \|A\|.$$

By positivity of $\omega((\lambda A + \mathbb{1})^*(\lambda A + \mathbb{1}))$, one can then prove the following two statements:

$$\forall A \in \mathcal{A}, \quad \omega(A^*) = \overline{\omega(A)}, \quad \|A^*\| = \|A\|.$$

The algebra $\mathcal{A}$ generated by our original set $\mathcal{O}$ of observables now satisfies the axioms of a C*-algebra with identity $\mathbb{1}$ and self-adjoint operators corresponding to the set $\mathcal{O}$.

In summary, we adopted the following mathematical axiomatization:

- A physical system is represented by its unital C*-algebra $\mathcal{A}$ of observables.

- A state of the given physical system is a normalized positive linear functional on $\mathcal{A}$.

- The set $\mathcal{S}(\mathcal{A})$ of states separates the elements of $\mathcal{A}$, and conversely observables separate states.

0.1.2. The Hilbert space formulation of quantum theory

In Section 0.1.1, we described the observables of a physical system in terms of a (possibly non-Abelian) C*-algebra $\mathcal{A}$, and the set $\mathcal{S}(\mathcal{A})$ of states of this system as positive linear functionals on it. We now seek for a way to derive concrete realizations of this abstract framework arising from physical requirements. We recall that in the *classical* Abelian case, the Gelfand-Naimark theorem allows such a concrete realization by representing observables by continuous functions on a compact Hausdorff topological space. The situation is rather different in the non-commutative setting.

Let us first recall that a *-*homomorphism* between two unital *-algebras $\mathcal{A}$ and $\mathcal{B}$ is a linear, multiplicative mapping $\pi : \mathcal{A} \to \mathcal{B}$ which preserves all the algebraic relations including the involution *. If π is bijective, it is called a *-*isomorphism*, a *-*automorphism* in the case when $\mathcal{A} = \mathcal{B}$. In the particular case when $\mathcal{B} = \mathcal{B}(\mathcal{H})$ is the C*-algebra of bounded linear operators on a Hilbert space $\mathcal{H}$, then π is called a *representation* of $\mathcal{A}$ in $\mathcal{H}$. Such a representation is *faithful* if $\ker(\pi) = \{0\}$, and it is called *irreducible* if $\{0\}$ and $\mathcal{H}$ are the only closed subspaces invariant under the image $\pi(\mathcal{A})$. In this last case, each vector $\Psi \in \mathcal{H}$ is *cyclic*, i.e. $\pi(\mathcal{A})\psi$ is dense in $\mathcal{H}$. More generally, we denote by $(\mathcal{H}, \pi, \Psi)$ a representation π in a Hilbert space $\mathcal{H}$ with a cyclic vector Ψ. The next theorem connects the framework that we developed in Section 0.1.1 to the more "traditional" Hilbert space formulation used in quantum information theory.

Theorem 0.1.1 (Gelfand-Naimark-Segal construction)**.** *Given a unital C*-algebra $\mathcal{A}$ and a state ω on $\mathcal{A}$, there exists a Hilbert space $\mathcal{H}_\omega$, with associated inner product denoted by $\langle .,. \rangle$, and a representation $\pi_\omega : \mathcal{A} \to \mathcal{B}(\mathcal{H}_\omega)$ such that*

- $\mathcal{H}_\omega$ *contains a cyclic vector Ψ_ω;*

- $\forall A \in \mathcal{A}, \quad \omega(A) = \langle \Psi_\omega, \pi_\omega(A) \Psi_\omega \rangle$;

- *Every other representation π in a Hilbert space $(\mathcal{H}_\pi, \langle .,. \rangle')$ with cyclic vector Ψ_π such that $\omega = \langle \Psi_\pi, \pi(A) \Psi_\pi \rangle'$ is unitarily equivalent to π_ω.*

The representation $(\mathcal{H}_\omega, \pi_\omega, \Psi_\omega)$ is called the GNS representation of $\mathcal{A}$ associated to the state ω. This representation is irreducible if and only if ω is pure, and faithful whenever ω is, that is whenever $\omega(A^*A) > 0$ for all $\mathcal{A} \ni A \neq 0$.

The GNS theorem does not provide a faithful representation in general. This is the content of the Gelfand-Naimark theorem:

Theorem 0.1.2 (Gelfand-Naimark). *A C^*-algebra $\mathcal{A}$ is isomorphic to a normed closed *-algebra of bounded operators acting on a Hilbert space.*

The normalized vectors of $\mathcal{H}$ hence describe a *full* set of states, that is a set capable of separating the observables.

Example 0.1.3 (Finite dimensional quantum systems). When $\mathcal{A} \equiv \mathcal{B}(\mathcal{H})$ is the set of bounded operators acting on a finite dimensional Hilbert space $\mathcal{H}$, states are represented by positive, trace-class operators ρ called *density operators*:

$$\omega(A) := \mathrm{Tr}(\rho A) \qquad \mathrm{Tr}(\rho) = 1 \,.$$

The set of density operators on $\mathcal{H}$ is denoted by $\mathcal{D}(\mathcal{H})$.

0.1.3. Measurements

So far, we have only described a generic mathematical framework for the description of expected values of certain observables of a physical system. However, one would ideally want to be able to talk about the probability that an event occurs. In the commutative (*classical*) setting, this justifies the conceptual jump between the concept of a bounded continuous function on a topological space to the one of a bounded measurable function defined on an abstract probability space $(\Omega, \mathcal{F}, \mu)$. In particular, this means we want to think of indicator functions $\mathbb{1}_A$, $A \in \mathcal{F}$, as being part of the algebra of observables, so that

$$\mathbb{E}_\mu[\mathbb{1}_A] = \mu(A)$$

is the probability that event A occurs. Now, let $X : \Omega \to \mathbb{R}$ be an $\mathcal{F}$-measurable real random variable of law μ. This means that for any Borel set $B \in \mathcal{B}(\mathbb{R})$

$$\mathbb{P}(X \in B) = \mu\left(X^{-1}(B)\right) \equiv \mu_X(B) = \mathbb{E}_{\mu_X}[\mathbb{1}_B] \qquad \Rightarrow \qquad \mathbb{E}[X] = \int_\mathbb{R} \lambda \, d\mu_X(\lambda) \,. \tag{0.1}$$

The operator of left multiplication:

$$\mathbb{L}_\infty(\Omega, \mathcal{F}, \mu) \to \mathcal{B}(\mathbb{L}_2(\Omega, \mathcal{F}, \mu)),$$
$$X \mapsto (L_X : \varphi \mapsto X \varphi),$$

allows us to think of X as a self-adjoint operator on $\mathbb{L}_2(\Omega, \mathcal{F}, \mu)$, whose spectral decomposition is written as

$$L_X = \int_\mathbb{R} \lambda \, dL_{\mathbb{1}_X}(\lambda), \tag{0.2}$$

where for every set $A \in \mathcal{B}(\mathbb{R})$, $\mathbb{1}_X(A) \coloneqq \mathbb{1}_{X^{-1}(A)}$ is bounded measurable in $\mathbb{L}_\infty(\Omega, \mathcal{F}, \mu)$, so that the spectral measure $L_{\mathbb{1}_X}$ on $\mathbb{L}_2(\Omega, \mathcal{F}, \mu)$ given by $L_{\mathbb{1}_X}(A)\varphi \coloneqq \mathbb{1}_{(X^{-1}(A))}\varphi$ defines an orthogonal projector of $\mathcal{B}(\mathbb{L}_2(\Omega, \mathcal{F}, \mu))$. In particular, $\mathbb{E}[\mathbb{1}_X(A)] = \mu(X^{-1}(A)) = \mu_X(A)$, as expected.

Back to the general C*-algebraic scenario, the Gelfand-Naimark theorem (0.1.2) implies that any observable $X \in \mathcal{O}$ can be thought of as a self-adjoint bounded operator acting on a Hilbert space $\mathcal{H}$. By spectral theory, X has the following decomposition

$$X = \int_{\mathbb{R}} \lambda \, dP_X(\lambda),$$

where $P_X : \mathcal{B}(\mathbb{R}) \to \mathcal{P}(\mathcal{H})$ is the so-called spectral measure of X taking values in the set of projections on $\mathcal{H}$. This means that, given a state ω of the system, there exists a probability measure $\mu_{X,\omega}$ on $\mathcal{B}(\mathbb{R})$ such that the probability that a measurement outputs a number $x \in B$ is given by

$$\mathbb{P}(X \in B) \equiv \mu_{X,\omega}(B) = \omega\left(P_X(B)\right) \quad \Rightarrow \quad \omega(X) = \int_{\mathbb{R}} \lambda \, \omega(dP_X(\lambda)). \tag{0.3}$$

By direct analogy with Equations (0.1) and (0.2), one sees that indicator functions should be replaced by projections. However, just as indicator functions are not continuous, the spectral projection of a representant X of an observable might not be part of the C*-algebra represented in $\mathcal{B}(\mathcal{H})$. In order to incorporate them in it, on needs to introduce the concept of a von Neumann algebra.

Let $\mathcal{H}$ be a Hilbert space. A *von Neumann algebra* $\mathcal{M}$ is a strongly closed *-subalgebra of $\mathcal{B}(\mathcal{H})$ (see Section 0.3 for a short review of von Neumann algebras). Since the strong topology is weaker than the norm topology, any von Neumann algebra is a C*-algebra. As expected, the algebra of complex bounded measurable functions $\mathbb{L}_\infty(\Omega, \mu)$ can be interpreted as a von Neumann algebra when acting by multiplication on the Hilbert space $\mathbb{L}_2(\Omega, \mu)$ of square integrable functions. In this sense, indicator functions are represented by projectors in the image of this representation, which completes the analogy between the commutative and the non-commutative cases.

So far, we only mentioned measurable events arising from the spectral measure P_X of an observable $X \in \mathcal{M}$. If the von Neumann algebra $\mathcal{M}$ is non-commutative, there exists a slightly more general notion of quantum measurement. Given a measurable space $(\Omega, \mathcal{F})$, a *positive-operator valued measure* (POVM), also known as *resolution of the identity*, is a collection $\mathbf{M} \coloneqq \{M(B) : B \in \mathcal{F}\}$ of self-adjoint operators in $\mathcal{M}$ such that

- $M(\varnothing) = 0$, and $M(\Omega) = \mathbb{1}_{\mathcal{H}}$;

- For all $B \in \mathcal{F}$, $M(B) \geq 0$;

- For any countable partition B_j of $B \in \mathcal{F}$, $\sum_j M(B_j)$ weakly converges to $M(B)$.

In the commutative setting, the only existing resolutions of the identity are defined through indicator functions $M(B) \coloneqq \mathbb{1}_B$, $B \in \Omega$. In the non-commutative case, any spectral measure associated to a self-adjoint operator is an orthogonal POVM, that is one for which $M(A)M(B) = 0$ if $A \cap B = \varnothing$. In fact Naimark's theorem states that any POVM on $\mathcal{H}$ can be dilated into the spectral measure of a self-adjoint operator in a larger Hilbert space $\tilde{\mathcal{H}}$.

0.1.4. Composite systems and conditional expectations

In this section, we are interested in the description of systems $\mathcal{A}_{12}$ composed of two subsystems $\mathcal{A}_1$ and $\mathcal{A}_2$. Mathematically, such systems are modeled by the *tensor product* $\mathcal{A}_1 \otimes \mathcal{A}_2$ of $\mathcal{A}_1$ and $\mathcal{A}_2$. More generally, given two normed vector spaces $\mathcal{V}_1$ and $\mathcal{V}_2$, and a pair $(x, y) \in \mathcal{V}_1 \times \mathcal{V}_2$, there exists a

unique element $x \otimes y$ in the dual of the space $\mathcal{B}(\mathcal{V}_1 \times \mathcal{V}_2, \mathbb{R})$ of bounded bilinear maps $B : \mathcal{V}_1 \times \mathcal{V}_2 \to \mathbb{R}$ defined by

$$\forall B \in \mathcal{B}(\mathcal{V}_1 \times \mathcal{V}_2, \mathbb{R}), \qquad x \otimes y\,(B) = B(x, y),$$

and such that $\|x \otimes y\| = \|x\| \|y\|$. Next, we denote by $\mathcal{V}_1 \otimes \mathcal{V}_2$ the *algebraic tensor product* of $\mathcal{V}_1$ and $\mathcal{V}_2$. It consists of the finite linear combinations of elementary tensors $x \otimes y$, for $x \in \mathcal{V}_1$ and $y \in \mathcal{V}_2$. Assuming that $\mathcal{A}_1 \coloneqq \mathcal{V}_1$ and $\mathcal{A}_2 \coloneqq \mathcal{V}_2$ are algebras, there is a unique multiplication on $\mathcal{A}_1 \otimes \mathcal{A}_2$ such that for all $A, A' \in \mathcal{A}_1$ and $B, B' \in \mathcal{A}_2$:

$$(A \otimes B)\,(A' \otimes B') \;=\; AA' \otimes BB'.$$

If $\mathcal{A}_1$ and $\mathcal{A}_2$ are $*$-algebras, the involution can also be extended to $\mathcal{A}_2 \otimes \mathcal{A}_2$ so that $(A \otimes B)^* = A^* \otimes B^*$ for all $(A, B) \in \mathcal{A}_1 \times \mathcal{A}_2$. Finally, if $\mathcal{A}_1$ and $\mathcal{A}_2$ are C*-algebras, there exists at least one C*-norm on $\mathcal{A}_1 \otimes \mathcal{A}_2$, thus providing the latter with a C*-algebraic structure. A C*-algebra $\mathcal{A}$ is called *nuclear* if for any other C*-algebra $\mathcal{B}$, there is only one C*-norm on $\mathcal{A} \otimes \mathcal{B}$. This is the case, for example, of finite C*-algebras.

Closely related to the concept of composite systems is the one of a conditional expectation. Let X belong to the space $\mathbb{L}_1(\Omega, \mathcal{F}, \mu)$ of integrable functions on the probability space $(\Omega, \mathcal{F}, \mu)$. We recall that the conditional expectation of X given an event $B \in \mathcal{F}$ is defined as

$$\mathbb{E}[X|B] = \frac{\mathbb{E}[X\,\mathbb{1}_B]}{\mu(B)}.$$

Moreover, given a discrete sub σ-algebra $\mathcal{G}$ of $\mathcal{F}$ generated by a countable family of disjoint events $(B_i)_{i \in I}$, the conditional expectation of X given $\mathcal{G}$ is defined as

$$\mathbb{E}[X|\mathcal{G}] \coloneqq \sum_{i \in I} \mathbb{E}[X|B_i]\,\mathbb{1}_{B_i}.$$

This random variable then outputs $\mathbb{E}[X|\mathcal{G}](\omega) = \mathbb{E}[X|B_i]$, the conditional expectation of X knowing B_i whenever the input ω belongs to B_i. This should be understood as a coarse-graining of the information available on the random variable X: if one is only able to separate events up to the sets B_i constituting $\mathcal{G}$, the best information on X knowing that an event belonging to B_i has occurred is $\mathbb{E}[X|B_i]$. One can easily check that the random variable $\mathbb{E}[X|\mathcal{G}]$ satisfies the three following properties

- $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable,

- $\mathbb{E}[X|\mathcal{G}]$ is integrable, since $\mathbb{E}\big[|\mathbb{E}[X|\mathcal{G}]|\big] \le \mathbb{E}\big[\mathbb{E}[|X|\,|\mathcal{G}]\big] = \mathbb{E}[|X|]$ and

- for any other $\mathcal{G}$-measurable random variable Y,

$$\mathbb{E}[X\,Y] = \mathbb{E}\big[\mathbb{E}[X|\mathcal{G}]\,Y\big]. \tag{0.4}$$

In the case of a general sub σ-algebra $\mathcal{G}$, these properties can in fact be turned into a definition: given any integrable random variable X, there exists a unique (up to null events) random variable, denoted by $\mathbb{E}[X|\mathcal{G}]$, such that the three above conditions are satisfied.

The notion of conditional expectation in the context of non-commutative tracial von Neumann algebras was introduced by [Umegaki, 1954, Umegaki, 1956, Umegaki, 1959, Umegaki, 1962] (see also [Nakamura and Turumaru, 1954]), and follows from the non-commutative generalization of the theory of integration [Dye, 1952, Dixmier, 1953, Segal, 1953]). The generalization of the theory to the

non-tracial setting was made by [Takesaki, 1972] (see also [Takesaki, 1979, Størmer, 1997]). Let ω be a faithful, semi-finite, normal weight on a von Neumann algebra $\mathcal{M}$, and let $\mathcal{N}$ be a von Neumann subalgebra of $\mathcal{M}$ such that the restriction $\omega|_{\mathcal{N}}$ of ω to $\mathcal{N}$ is semi-finite. A linear map $E_{\mathcal{N}} : \mathcal{M} \to \mathcal{N}$ is called the *conditional expectation* of $\mathcal{M}$ into $\mathcal{N}$ with respect to ω if the following conditions are satisfied:

- For all $X \in \mathcal{M}$, $\|E_{\mathcal{N}}[X]\| \leq \|X\|$,

- For all $X \in \mathcal{N}$, $E_{\mathcal{N}}[X] = X$,

- $\omega \circ E_{\mathcal{N}} = \omega$.

By Theorem III.3.4, as well as the proof of Theorem IX.4.2 of [Takesaki, 1979], a conditional expectation $E_{\mathcal{N}}$ satisfies the following properties:

- *Positivity*: for all $X \in \mathcal{M}$, $E_{\mathcal{N}}[X^*X] \geq 0$. In fact, a conditional expectation is *completely positive* (see Section 0.1.5),

- for any $X \in \mathcal{M}$ and $A, B \in \mathcal{N}$,

$$E_{\mathcal{N}}[AXB] = A\, E_{\mathcal{N}}[X]B. \tag{0.5}$$

Therefore, similarly to Equation (0.4),

$$\omega(E_{\mathcal{N}}[AXB]) = \omega(AE_{\mathcal{N}}[X]B).$$

As we mentioned above, the existence and almost sure uniqueness of the conditional expectation is always true in a classical setting. For non-commutative algebras, this is no longer true. The following theorem from [Takesaki, 1972] (see also Theorem IX.4.2 of [Takesaki, 2003]) provides a necessary and sufficient condition for it to be true (we refer to Section 0.1.5 for the definition of the modular automorphism group associated with a state).

Theorem 0.1.4. *Let $\mathcal{M}, \mathcal{N}$ be two von Neumann algebras with $\mathcal{N} \subseteq \mathcal{M}$. The existence of a conditional expectation $E_{\mathcal{N}} : \mathcal{M} \to \mathcal{N}$ with respect to a faithful normal state on $\mathcal{M}$ is equivalent to the global invariance of $\mathcal{N}$ under the associated modular automorphism group:*

$$\forall t \in \mathbb{R}, \qquad \alpha_t^{\omega}(\mathcal{N}) = \mathcal{N}.$$

In this case, the conditional expectation $E_{\mathcal{N}}$ is normal and uniquely determined by ω.

We end this section with two canonical examples where the notion of a conditional expectation will keep occurring:

Example 0.1.5 (Classical probability theory)**.** As always, the classical case is retrieved as follows: let $(\Omega, \mathcal{F}, \mu)$ be a probability space and $\mathcal{M} := \mathbb{L}_{\infty}(\Omega, \mathcal{F})$ with associated faithful normal state given by the expected value:

$$\mathbb{E}_{\mu}(.) := \int_{\Omega} (.)\, d\mu. \tag{0.6}$$

Now let $\mathcal{G}$ be a sub σ-algebra of $\mathcal{F}$ on Ω. Then, $\mathcal{N} := \mathbb{L}^{\infty}(\Omega, \mathcal{G}) \subset \mathcal{M}$, and one easily verifies that $E_{\mathcal{N}} = \mathbb{E}[.|\mathcal{G}]$. In particular, let $(\Omega, \mathcal{F}) := (\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ be a product measurable space, and let $\mathcal{G} := \mathcal{F}_1$ be understood as a sub σ-algebra of $\mathcal{F}$ (more precisely, take $\mathcal{G} := \{A \times \Omega_2, \ A \in \mathcal{F}_1\}$). Then, the predual $E_{\mathcal{N}*}$ of the conditional expectation $E_{\mathcal{N}} \equiv \mathbb{E}[.|\mathcal{G}]$ maps μ onto its marginal μ_{Ω_1}.

Example 0.1.6 (Finite dimensional quantum systems)**.** Let $\mathcal{H}_1, \mathcal{H}_2$ be two finite dimensional Hilbert spaces, and $\mathcal{B}(\mathcal{H}_1), \mathcal{B}(\mathcal{H}_2)$ the sets of linear operators on $\mathcal{H}_1$, respectively $\mathcal{H}_2$. One trivially embeds $\mathcal{B}(\mathcal{H}_1)$ into $\mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ the following way:

$$\mathcal{B}(\mathcal{H}_1) \ni A \mapsto A \otimes \mathbb{1}_{\mathcal{H}_2}. \tag{0.7}$$

It is then easy to check that the partial trace $\mathrm{Tr}_{\mathcal{H}_2}$, uniquely defined by the following identity:

$$\forall A \in \mathcal{B}(\mathcal{H}_A), \ X \in \mathcal{B}(\mathcal{H}), \quad \mathrm{Tr}\left((A \otimes \mathbb{1}_{\mathcal{H}_2})X\right) = \mathrm{Tr}\left(A\, \mathrm{Tr}_{\mathcal{H}_2}(X)\right),$$

is the conditional expectation on $\mathcal{M} := \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ onto $\mathcal{N} := \mathcal{B}(\mathcal{H}_1)$ seen as a von Neumann subalgebra of $\mathcal{M}$ under the above embedding (0.7). Similarly to the classical case, its predual, also given by $\mathrm{Tr}_{\mathcal{H}_2}$, is the non-commutative version of the action of taking the marginal of a probability measure: it maps any state $\rho \in \mathcal{D}(\mathcal{H})$ to its so-called *reduced state*

$$\rho_1 := \mathrm{Tr}_{\mathcal{H}_2}(\rho).$$

Quantum systems composed of two parts $\mathcal{H}_1 \otimes \mathcal{H}_2$ are commonly referred to as *bipartite*.

More generally, let $\mathcal{N}$ be a C*-subalgebra of $\mathcal{B}(\mathcal{H})$, where $\mathcal{H}$ is a finite dimensional Hilbert space. In this case, by the structure theorem for C*-algebras acting on finite dimensional Hilbert spaces, $\mathcal{N}$ can always be decomposed into the direct sum of a commutative part and subparts where it restricts to a factor [Kadison and Ringrose, 2015]. More precisely, up to a unitary transformation, the Hilbert space $\mathcal{H}$ admits the following decomposition

$$\mathcal{H} = \bigoplus_{i \in I} \mathcal{H}_i \otimes \mathcal{K}_i, \tag{0.8}$$

such that $\mathcal{N}$ is unitarily isomorphic to the algebra

$$\mathcal{N} = \bigoplus_{i \in I} \mathcal{B}(\mathcal{H}_i) \otimes \mathbb{1}_{\mathcal{K}_i}. \tag{0.9}$$

In this case, given any state of the form

$$\sigma_{\mathrm{Tr}} := \frac{1}{d_{\mathcal{H}}} \sum_{i \in I} d_{\mathcal{K}_i} \mathbb{1}_{\mathcal{H}_i} \otimes \tau_i, \tag{0.10}$$

where for each $i \in I$, τ_i is a full-rank density operator on $\mathcal{K}_i$, $\mathcal{N}$ is invariant under the action of the modular operator $\Delta_{\sigma_{\mathrm{Tr}}}(.) := \sigma_{\mathrm{Tr}}(.)\sigma_{\mathrm{Tr}}^{-1}$. In fact, the algebra is pointwise invariant: for any $X := \sum_{i \in I} X_i \otimes \mathbb{1}_{\mathcal{K}_i} \in \mathcal{N}$,

$$\Delta_{\sigma_{\mathrm{Tr}}}(X) = \sigma_{\mathrm{Tr}} X \, \sigma_{\mathrm{Tr}}^{-1} = \sum_{i \in I} X_i \otimes \mathbb{1}_{\mathcal{K}_i} = X.$$

By Theorem 0.1.4, this implies the existence of a unique conditional expectation $E_{\mathcal{N}}$ associated to σ_{Tr}[2]. One can simply verify that $E_{\mathcal{N}}$ is defined as follows: for any $X \in \mathcal{B}(\mathcal{H})$,

$$E_{\mathcal{N}}[X] = \sum_{i \in I} \mathrm{Tr}_{\mathcal{K}_i}\left((\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i)P_i X P_i\right) \otimes \mathbb{1}_{\mathcal{K}_i}. \tag{0.11}$$

In particular, $\sigma_{\mathrm{Tr}} = E_{\mathcal{N}*}\left(\frac{\mathbb{1}_{\mathcal{H}}}{d_{\mathcal{H}}}\right)$, where for each i, P_i denotes the projector onto $\bigoplus_{j=1}^{\dim(\mathcal{K}_i)} \mathcal{H}_i$, and $\mathrm{Tr}_{\mathcal{K}_i}$

---

[2]We recall that the modular automorphism group is given by $(\Delta_{\sigma_{\mathrm{Tr}}}^{it})_{t \in \mathbb{R}}$, see Section 0.1.5.

is the partial trace with respect to $\mathcal{K}_i$. The choice of σ_{Tr} will appear to be particularly relevant in the study of non-commutative functional inequalities and decoherence in Part IV. This comes from the fact that σ_{Tr} is *tracial* on $\mathcal{N}$, that is, for all $X \in \mathcal{N}$ and all $Y \in \mathcal{B}(\mathcal{H})$,

$$\text{Tr}(\sigma_{\text{Tr}} XY) = \text{Tr}(\sigma_{\text{Tr}} YX).$$

We refer to Chapter 6 for more details.

0.1.5. Dynamics

A large part of this thesis concerns the study of the dynamics of *open quantum systems*, that is quantum systems interacting with their environment. In the two paragraphs below, we briefly introduce the concept of time evolution of a quantum system, which we will explore in more depth in Parts III and IV.

Closed dynamics   We first describe the case of a *closed quantum system* modeled by a C^*-algebra $\mathcal{A}$. Since the system does not interact with its environment, the evolution is assumed to be time-reversible. Therefore, the time dynamics is in full generality given by a $*$-automorphism α on $\mathcal{A}$. Classically, that is when the algebra $\mathcal{A}$ consists of the set of bounded measurable operators on a measurable space $(E, \mathcal{F})$, the dynamics is defined as follows: for any f bounded measurable, and any $x \in E$:

$$\alpha(f)(x) = f \circ \varphi(x),$$

where φ is an invertible map on E (see e.g. [Pillet, 2006]). Moreover, given any reference state ω, the GNS theorem implies the following action of α on the GNS Hilbert space $\mathcal{H}_\omega$:

Theorem 0.1.7. *Let $\mathcal{A}$ be a C^*-algebra, ω a state on $\mathcal{A}$ and α a $*$-automorphism of $\mathcal{A}$ that leaves ω invariant, i.e., such that $\omega \circ \alpha = \omega$. Then, there exists a unique unitary $U \in \mathcal{B}(\mathcal{H}_\omega)$ such that for all $X \in \mathcal{A}$:*

$$\pi_\omega\left(\alpha(A)\right) = U \pi_\omega(A) U^*.$$

The reference state ω will for example be taken as the ground state ω of a Hamiltonian system. The above result can be interpreted as a justification of the more traditional notion of "unitary evolution" in the Hilbert space formulation of quantum theory. Physical evolutions of the above sort are usually associated with a time parameter $t > 0$, and denoted by α_t. Reversibility then means that α_t is an invertible map, and we denotes its inverse by α_{-t}. We further assume that the dynamics can be restarted at any point α_s along its trajectory to get the same result α_{t+s} as flowing forward for time $t + s$ from α_0. This natural property is known as the *group property*, and reads as follows:

$$\forall s, t \in \mathbb{R}, \qquad \alpha_{s+t} = \alpha_s \circ \alpha_t.$$

Assume that the group $(\alpha_t)_{t \geq 0}$ is *strongly continuous*, which means that for any $A \in \mathcal{A}$, $\|\alpha_t(A) - A\| \to 0$ as $t \to 0$ (in the case of a von Neumann algebra, one simply assumes that the group is weak* continuous). Then, there exists a *generator* δ of the dynamics of domain

$$\text{dom}(\delta) := \left\{ A \in \mathcal{A} : \lim_{t \to 0} t^{-1}(\alpha_t(A) - A) \text{ exists} \right\},$$

that is a (possibly unbounded) operator $(\delta, \mathrm{dom}(\delta))$ defined by

$$\forall A \in \mathrm{dom}(\delta), \qquad \delta(A) = \lim_{t \to 0} \frac{\alpha_t(A) - A}{t},$$

where the limit is taken with respect to the norm topology on $\mathcal{A}$. Consider now, for each $t \in \mathbb{R}$, the unitary U_t associated to α_t through the GNS representation $\mathcal{H}_\omega$ of a reference state. This in particular implies that for all $t \geq 0$, and any $A \in \mathcal{A}$

$$\pi_\omega(\alpha_t(A)) = U_t \, \pi_\omega(A) U_t^*.$$

It is easy to verify that the family $(U_t)_{t \geq 0}$ itself forms a strongly continuous group of unitary operators. The following important theorem, due to Stone, will allow us to reconnect up to the better known notion of Schrödinger evolution in the traditional Hilbert space formulation of quantum mechanics:

Theorem 0.1.8 (Stone, see e.g. [Reed and Simon, 1972])**.** *For all $t \in \mathbb{R}$, let U_t be a bounded linear operator on a Hilbert space $\mathcal{H}$. Then $(U_t)_{t \geq 0}$ is a strongly continuous one-parameter group of unitary operators if and only if there exists a densely defined self-adjoint operator H with domain $\mathrm{dom}(H)$ such that*

$$U_t = \mathrm{e}^{-itH}.$$

Moreover, $\mathrm{dom}(H) = \{ \psi \in \mathcal{H} : \exists \phi \in \mathcal{H} \text{ such that } \lim_{t \to 0} \| t^{-1}(\mathrm{e}^{-itH} \psi - \psi) - \phi \| = 0 \}$.

In the framework of dynamical systems, the operator H is usually referred to as the *Hamiltonian* of the system. In the case when $\mathcal{A} = \mathcal{B}(\mathcal{H})$ for a finite dimensional Hilbert space $\mathcal{H}$, we recover the usual Schrödinger equation: for any $\rho \in \mathcal{D}(\mathcal{H})$ and $A \in \mathcal{B}(\mathcal{H})$

$$A_t := \alpha_t(A) = \mathrm{e}^{-itH} \, A \, \mathrm{e}^{itH} \qquad \Leftrightarrow \qquad \frac{d}{dt} A_t = -i [H, A_t], \qquad \text{(Heisenberg picture)}$$

$$\rho_t := \hat{\alpha}_t(\rho) = \mathrm{e}^{itH} \, \rho \, \mathrm{e}^{-itH} \qquad \Leftrightarrow \qquad \frac{d}{dt} \rho_t = i [H, \rho_t], \qquad \text{(Schrödinger picture)}$$

where the map $\hat{\alpha}_t$ is the dual of α_t with respect to the Hilbert Schmidt inner product on $\mathcal{B}(\mathcal{H})$: $\langle A, B \rangle_{\mathrm{HS}} := \mathrm{Tr}[A^* B]$.

KMS states: Given a closed evolution α as above, it is simple to verify that states of the form

$$\sigma := \frac{\mathrm{e}^{-\beta H}}{\mathrm{Tr}(\mathrm{e}^{-\beta H})} \tag{0.12}$$

are invariant with respect to the unitary evolution: $\hat{\alpha}_t(\sigma) = \sigma$. Such a state can be shown to be the only state satisfying the following KMS condition: for any $A, B \in \mathcal{B}(\mathcal{H})$,

$$\mathrm{Tr}(\sigma B A) = \mathrm{Tr}(\sigma A \, \alpha_{i\beta}(B)).$$

States of the above type still exist in infinite dimensions. However, they would be well defined only if the partition function $Z(\beta) := \mathrm{Tr}(\mathrm{e}^{-\beta H})$ is finite, which in particular implies that H has discrete spectrum. One way to avoid this restriction is to rather take the KMS-condition as a starting point. First, given a weak* continuous group $(\alpha_t)_{t \in \mathbb{R}}$ of *-automorphisms on a von Neumann algebra $\mathcal{M}$, an element $X \in \mathcal{M}$ is called *entire analytic* for α_t if there exists a function $f : \mathbb{C} \to \mathcal{M}$ such that

  (i) $f(t) = \alpha_t(X)$ for all $t \in \mathbb{R}$;

  (ii) $z \mapsto \eta(f(z))$ is entire analytic for all $\eta \in \mathcal{M}_*$.

The set $\mathcal{M}_\alpha$ of entire analytic elements is a weak* dense, $(\alpha_t)_{t\geq 0}$-invariant *-subalgebra of $\mathcal{M}$. Then[3]

Definition 0.1.9. Let $(\alpha_t)_{t\geq\mathbb{R}}$ be a weak* continuous group of *-automorphisms on a von Neumann algebra $\mathcal{M}$, and $\beta \in \mathbb{R}$. A state ω is said to be a (α, β)-KMS *state* if it is invariant under the action of $(\alpha_t)_{t\in\mathbb{R}}$ and the following holds for any A, B in a weak* dense $(\alpha_t)_{t\in\mathbb{R}}$-invariant *-subalgebra of $\mathcal{M}_\alpha$:

$$\omega(A\alpha_{i\beta}(B)) = \omega(BA). \qquad \text{(KMS-condition)}$$

The KMS-condition is known to be equivalent to the following identity holding true for all A, B in a weak* dense, $(\alpha_t)_{t\in\mathbb{R}}$-invariant *-subalgebra of $\mathcal{M}_\alpha$ (see Proposition 5.3.3 of [Bratteli and Robinson, 1997]):

$$\omega(\alpha_{-i\beta/2}(A)\alpha_{i\beta/2}(B)) = \omega(BA). \qquad (0.13)$$

In practice, we always consider the following weak* continuous group of *-automorphisms (we refer to Section 1.2 for an brief introduction to modular theory): given a faithful normal state ω on a von Neumanna algebra $\mathcal{M}$, with corresponding cyclic vector Ω_ω in the standard form $(\pi_\omega(\mathcal{M}), \Omega_\omega)$, the *modular automorphism group* α^ω associated to ω is defined as follows: for any $X \in \mathcal{M}$ and all $t \in \mathbb{R}$:

$$\alpha_t^\omega(X) := \pi_\omega^{-1}(\Delta_\omega^{it} \pi_\omega(X) \Delta_\omega^{-it}), \qquad (0.14)$$

which belongs to $\mathcal{M}$ by Tomita-Takesaki theorem (cf. Theorem 2.5.14 of [Bratteli and Robinson, 1979]), where Δ_ω is the modular operator associated with the pair $(\pi_\omega(\mathcal{M}), \Omega_\omega)$. It can be verified that ω is a $(\alpha^\omega, -1)$-KMS state. In the case when $\mathcal{M} = \mathcal{B}(\mathcal{H})$, with $\mathcal{H}$ finite dimensional, and ω is with associated full-rank density operator σ and cyclic vector $\Omega_\omega := \sigma^{\frac{1}{2}}$, the KMS-condition with respect to the modular automorphism group α^ω reduces to:

$$\omega(A\alpha_{i\beta}^\omega(B)) = \text{Tr}\left(\sigma^{\frac{1}{2}} A \Delta_\sigma B \Delta_\sigma^{-1}(\sigma^{\frac{1}{2}})\right) = \text{Tr}\left(\sigma^{\frac{1}{2}} A\sigma B\sigma^{-1}\sigma^{\frac{1}{2}}\right) = \text{Tr}(A\sigma B) = \omega(BA).$$

Open dynamics   In this paragraph, we start touching what constitutes the main subject of this thesis, which is the theory of *open quantum systems*. As opposed to their closed counterparts, these systems are assumed to interact with an *environment* (heat bath, reservoir, magnetic field, etc.). The latter should be considered as being constituted of a much larger number of particles than the system itself. As a result of such interactions, a system Σ gets coupled with its environment E. This typically leads to a local loss of the quantum features of the reduced state ω_Σ that are believed to be crucial for certain key tasks of quantum information and quantum computing. A rigorous mathematical investigation of open quantum systems is postponed to Parts III and IV. Here instead, we introduce some of their main ingredients on a simple example following the treatment of [Spohn and Lebowitz, 1978] (see also [Davies, 1974, Davies, 1976, Lidar et al., 2001]).

Assume a system Σ of interest, of associated finite dimensional Hilbert space $\mathcal{H}_\Sigma$, interacts with an environment E (typically infinite dimensional), also called a reservoir. The dynamics of Σ when isolated is governed by the closed evolution generated by a Hamiltonian H_Σ. We also assume for sake of simplicity that the reservoir is only composed of one heat bath with associated inverse temperature β[4] and evolution generated by the Hamiltonian H_E. Since the bath is assumed to be at equilibrium,

---

[3]For more details about the physical meaning of KMS states and KMS symmetry, we refer to section IV of [Spohn and Lebowitz, 1978] and references therein.

[4]Here, we assume the Boltzmann constant is equal to 1.

it is in the Gibbs state

$$\rho_E^\beta := \frac{\mathrm{e}^{-\beta H_E}}{\mathrm{Tr}\ \mathrm{e}^{-\beta H_E}},$$

which can be immediately shown to be invariant under the unitary evolution generated by H_E. Next, the interaction between Σ an E is modeled by a bounded interaction Hamiltonian $\lambda H_{\Sigma E}$, where $\lambda \geq 0$ is called the *coupling strength*. We will further assume that $\mathrm{Tr}_E\,(\mathbb{1}_\Sigma \otimes \rho_E^\beta)\,H_{\Sigma E} = 0$ (the state of the environment is invariant under the interaction with the system). The Hamiltonian of the coupled system $\mathcal{H}_\Sigma \otimes \mathcal{H}_E$ is then given by a sum of three terms:

$$H_\lambda := H_\Sigma \otimes \mathbb{1}_E + \mathbb{1}_\Sigma \otimes H_E + \lambda\,H_{\Sigma E},$$

and its time evolution is then given by the unitary $U_t^\lambda := \mathrm{e}^{itH_\lambda}$. Then, for an initial state of the joint system given by $\rho_0 := \rho_\Sigma \otimes \rho_E^\beta \in \mathcal{D}(\mathcal{H}_\Sigma \otimes \mathcal{H}_E)$, the state evolved at time t takes the following form:

$$\rho_t := U_t^\lambda\,\rho_0\,(U_t^\lambda)^* \equiv \alpha_t^\lambda(\rho_0).$$

Since we are only interested in the evolution of the reduced state on σ, partial tracing over the environment leads to the definition of a family of maps $\mathcal{P}_t^\lambda : \mathcal{D}(\mathcal{H}_\Sigma) \to \mathcal{D}(\mathcal{H}_\Sigma)$, $t \geq 0$, so that

$$\mathcal{P}_t^\lambda(\rho_\Sigma) := \mathrm{Tr}_E\left[U_t^\lambda\,(\rho_\Sigma \otimes \rho_E^\beta)\,(U_t^\lambda)^*\right].$$

In general, the reduced evolution $(\mathcal{P}_t)_{t\geq 0}$ contains memory effects due to the interaction term $\lambda\,H_{\Sigma E}$[5]. An expression that shows explicitly these effects can be derived: define the projector

$$P_\beta : \begin{cases} \mathcal{T}_1(\mathcal{H}_\Sigma \otimes \mathcal{H}_E) \to \mathcal{T}_1(\mathcal{H}_\Sigma) \otimes \rho_E^\beta, \\ \qquad T \qquad \mapsto \mathrm{Tr}_E(T) \otimes \rho_E^\beta, \end{cases}$$

where given a Hilbert space $\mathcal{H}$, $\mathcal{T}_1(\mathcal{H})$ represents the Banach space of trace-class operators in $\mathcal{H}$ (see Chapter 1), and define $(\gamma_t^\lambda)_{t\geq 0}$ to be the time evolution on $\mathcal{T}_1(\mathcal{H}_\Sigma \otimes \mathcal{H}_E)$ generated by

$$T \mapsto P_\beta \circ [H_\lambda, .] \circ P_\beta(T) + (\mathrm{id} - P_\beta) \circ [H_\lambda, .] \circ (\mathrm{id} - P_\beta)(T).$$

The following formula relating the evolutions $(\alpha_t^\lambda)_{t\geq 0}$ and $(\gamma_t^\lambda)_{t\geq 0}$ can then be verified by direct differentiation:

$$\forall t \geq 0, \quad \alpha_t^\lambda = \gamma_t^\lambda - i\lambda \int_0^t \gamma_{t-s}^\lambda \circ (P_\beta \circ [H_{\Sigma E}, .] \circ (\mathrm{id} - P_\beta) + (\mathrm{id} - P_\beta) \circ [H_{\Sigma E}, .] \circ P_\beta) \circ \alpha_s^\lambda\,ds. \quad (0.15)$$

Next, the two following equations are derived from Equation (0.15):

$$P_\beta \circ \alpha_t^\lambda \circ P_\beta = \gamma_t^\lambda \circ P_\beta - i\lambda \int_0^t \gamma_{t-s}^\lambda \circ P_\beta \circ [H_{\Sigma E}, .] \circ (\mathrm{id} - P_\beta) \circ \alpha_s^\gamma \circ P_\beta\,ds, \qquad (0.16)$$

$$(\mathrm{id} - P_\beta) \circ \alpha_t^\lambda \circ P_\beta = -i\lambda \int_0^t \gamma_{t-s}^\lambda \circ (\mathrm{id} - P_\beta) \circ [H_{\Sigma E}, .] \circ P_\beta \circ \alpha_s^\gamma \circ P_\beta\,ds. \qquad (0.17)$$

A direct substitution of Equation (0.17) into Equation (0.16) then yields:

$$P_\beta \circ \alpha_t^\lambda \circ P_\beta = \gamma_t^\lambda \circ P_\beta$$
$$- \lambda^2 \int_0^t \int_0^s \gamma_{t-s}^\lambda \circ P_\beta \circ [H_{\Sigma E}, .] \circ (\mathrm{id} - P_\beta) \circ \gamma_{s-u}^\lambda \circ (\mathrm{id} - P_\beta) \circ [H_{\Sigma E}, .] \circ P_\beta \circ \alpha_u^\lambda \circ P_\beta\,du\,ds$$

---

[5] And the problem of deciding whether a quantum system has evolved according to a memoryless dynamics is NP-hard, as shown in [Cubitt et al., 2012].

Since we assumed that we were starting with a decoupled state $\rho_0 = \rho_\Sigma \otimes \rho_E^\beta$, $\mathcal{P}_t^\lambda(\rho_\Sigma) \otimes \rho_E^\beta = P_\beta \circ \alpha_t^\lambda \circ P_\beta(\rho_\Sigma \otimes \rho_E^\beta)$, and the last equation above leads to the following formulation for the reduced dynamics $(\mathcal{P}_t^\lambda)_{t \geq 0}$:

$$\mathcal{P}_t^\lambda(\rho_\Sigma) = \mathcal{P}_t^0(\rho_\Sigma) + \lambda^2 \int_0^t \int_0^s \mathcal{P}_{t-s}^0 \circ K^\lambda(s-u) \circ \mathcal{P}_u^\lambda(\rho_\Sigma) \, du \, ds \,,$$

where the *integral kernel* K^λ responsible for the memory effects of the evolution is defined as

$$K^\lambda(s) = -\operatorname{Tr}_E\left([H_{\Sigma E}, .] \circ (\operatorname{id} - P_\beta) \circ \gamma_s^\lambda \circ (\operatorname{id} - P_\beta) \circ [H_{\Sigma E}, .] \circ R^\beta\right), \qquad R^\beta : \rho_\Sigma \mapsto \rho_\Sigma \otimes \rho_E^\beta \,.$$

After differentiation, we obtained the so-called *generalized master equation* of [Nakajima, 1958, Prigonine and Resibois, 1961, Zwanzig, 1960]:

$$\frac{d}{dt}\rho_\Sigma(t) = -i\left[H_\Sigma, \rho_\Sigma(t)\right] + \lambda^2 \int_0^t K^\lambda(t-s)\rho_\Sigma(s)\, ds, \tag{0.18}$$

The analysis of such an evolution is in general very difficult to carry out due to the memory effects: for example, the existence of a stationary state has only been proved in special cases. For this reason, we look for a simplified, Markovian version of Equation (0.18). One way to achieve this goal is by taking the so-called *weak coupling limit*. As shown below, this approximation involves two limits: one first needs to go from the Schrödinger picture to the interaction picture and suppress oscillations of parts of the system Σ that are uncoupled by taking $\lambda \to 0$. Moreover, in order to make up for the slowing of the decay as $\lambda \to 0$, time needs to be rescaled in such a way that $\lambda^2 t \equiv \tau$ is kept fixed (*van Hove limit* [van Hove, 1954, Hove, 1957]). Rescaling the time $\tau := \lambda^2 t$ and going over to the interaction picture:

$$\mathcal{P}_{-\lambda^{-2}\tau}^0 \circ \mathcal{P}_{\lambda^{-2}\tau}^\lambda(\rho_\Sigma)$$
$$= \rho_\Sigma + \int_0^\tau d\tau' \left(\mathcal{P}_{-\lambda^{-2}\tau'}^0 \circ \left[\int_0^{\lambda^{-2}(\tau-\tau')} dx \, \mathcal{P}_{-x}^0 \circ K^\lambda(x)\right] \circ \mathcal{P}_{\lambda^{-2}\tau'}^0\right) \circ \mathcal{P}_{-\lambda^{-2}\tau'}^0 \circ \mathcal{P}_{\lambda^{-2}\tau'}^\lambda(\rho_\Sigma).$$

The integral kernel admits a power series expansion in the coupling constant λ. Define $\tilde{K}^0 := \mathcal{P}_{-x}^0 \circ K^0(x)$, where K^0 denotes the 0-th order term in that expansion, as well as $\tilde{K} := \int_0^\infty \tilde{K}^0(x)\, dx$. The above expression then simplifies after taking the limit $\rho_\Sigma(\tau) := \lim_{\lambda \to 0} \mathcal{P}_{-\lambda^{-2}\tau}^0 \circ \mathcal{P}_{\lambda^{-2}\tau}^\lambda(\rho_\Sigma)$,

$$\tilde{\mathcal{P}}_\tau(\rho_\Sigma) := \rho_\Sigma(\tau) = \rho_\Sigma(0) + \lim_{\lambda \to 0} \int_0^\tau \mathcal{P}_{-\lambda^{-2}\tau'}^0 \circ \tilde{K} \circ \mathcal{P}_{\lambda^{-2}\tau'}^0(\rho_\Sigma(\tau'))\, d\tau' \tag{0.19}$$

Now, since we assumed $\mathcal{H}_\Sigma$ to be finite dimensional, the following time average always exists:

$$\mathcal{L} := \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^T \mathcal{P}_{-t}^0 \circ \tilde{K} \circ \mathcal{P}_t^0 \, dt = \sum_{\omega \in \operatorname{sp}([H_\Sigma, .])} Q_\omega \circ \tilde{K} \circ Q_\omega, \tag{0.20}$$

where the operators Q_ω are the spectral projectors of $[H_\Sigma, .]$ seen as a self-adjoint operator acting on $\mathcal{T}_1(\mathcal{H}_\Sigma)$ with corresponding eigenvalue ω, that is

$$Q_\omega(T) = \sum_{\varepsilon_m - \varepsilon_n = \omega} P_m \, T \, P_n,$$

where ε_j, resp. P_j, denote the eigenvalues, resp. eigenprojectors, of H_Σ:

$$H_\Sigma := \sum_j \varepsilon_j \, P_j.$$

Since $H_{\Sigma E}$ is assumed to be bounded on $\mathcal{H}_\Sigma \otimes \mathcal{H}_E$, so are $\tilde{K}$ and $\mathcal{L}$ on $\mathcal{T}_1(\mathcal{H}_\Sigma)$. Inserting Equation (0.20) into Equation (0.19), we finally obtain the following Markovian limit of Equation (0.18):

$$\frac{d}{d\tau}\rho_\Sigma(\tau) = \mathcal{L}\rho_\Sigma(\tau). \tag{0.21}$$

This heuristic derivation was rigorously proved to hold by [Davies, 1974] under some suitable assumptions. It is unclear from the above derivation what information about the physical process $\mathcal{P}_t^\lambda \rho_\Sigma$ remains after taking the weak coupling limit $\rho_\Sigma(\tau)$, which amounts to letting $\lambda \to 0$ and $t \to \infty$ at the same time. We first make the observation that the interaction term $\lambda H_{\Sigma E}$ is expected to be small compared with the Hamiltonian H_Σ of the system for physical systems of relevance. Indeed, interaction terms should only occur at the surface between Σ and E, which represents only a small amount of particles (and hence of energy) compared to the bulk of the system. Moreover, assume that the limit $\sigma^\lambda := \lim_{t\to\infty} \mathcal{P}_t^\lambda(\rho_\Sigma)$ exists and is unique (i.e. independent of the initial state ρ_Σ), and assume that there exists a unique stationary state σ for the Markovian evolution governed by Equation (0.21). Since the Gibbs state

$$\rho_\Sigma^\beta := \frac{\mathrm{e}^{-\beta H_\Sigma}}{\mathrm{Tr}\,\mathrm{e}^{-\beta H_\Sigma}} \tag{0.22}$$

can be shown to always be stationary (see e.g. Property 1 of [Spohn and Lebowitz, 1978]), we will assume $\sigma = \rho_\Sigma^\beta$. Then, since ρ_Σ^β trivially commutes with H_Σ, this heuristically means in particular that

$$\rho_\Sigma^\beta \underset{\lambda\to 0}{\sim} \mathcal{P}_{-\lambda^{-2}\tau}^0 \circ \mathcal{P}_{\lambda^{-2}\tau}^\lambda(\rho_\Sigma^\beta) \underset{\lambda\to 0}{\sim} \mathcal{P}_{\lambda^{-2}\tau}^\lambda(\rho_\Sigma^\beta) \underset{\lambda\to 0}{\sim} \lim_{t\to\infty} \mathcal{P}_t^\lambda(\rho_\Sigma^\beta) \equiv \sigma^\lambda,$$

and hence, the invariant state of the Markovian evolution should correspond to the 0-th order term of a perturbative expansion of σ^λ.

Completely positive maps and quantum Markov semigroups In the above paragraph, the maps $(\tilde{\mathcal{P}}_\tau)_{\tau\geq 0}$ form a so-called *quantum Markov semigroup* (QMS) on the Banach space $\mathcal{T}_1(\mathcal{H}_\Sigma)$. Given a separable Hilbert space $\mathcal{H}$, a family $(\mathcal{P}_t)_{t\geq 0}$ of linear mappings $\mathcal{P}_t : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_1(\mathcal{H})$ is called a quantum Markov semigroup if[6]:

- *Semigroup property*: For all $s, t \geq 0$, $\mathcal{P}_{t+s} = \mathcal{P}_t \circ \mathcal{P}_s$;

- *Strong continuity*: For all $T \in \mathcal{T}_1(\mathcal{H})$, $\lim_{t\to 0_+} \|\mathcal{P}_t(T) - T\|_1 = 0$;

- *Complete positivity*: For all $t \geq 0$ and any $n \in \mathbb{N}$, the map $\mathcal{P}_t \otimes \mathrm{id}_{\mathbb{M}_n(\mathbb{C})}$ is positivity preserving[7];

- *Preservation of the trace*: For all $t \geq 0$ and $T \in \mathcal{T}_1(\mathcal{H})$, $\mathrm{Tr}\,(\mathcal{P}_t(T)) = \mathrm{Tr}(T)$.

In fact, the last two properties, namely preservation of the trace and complete positivity, are commonly accepted as providing the correct general notion of an evolution by quantum information theorists. Such maps are called *completely positive, trace preserving* (CPTP), and will also be referred to as *quantum channels* in the sequel. The property of trace preservation is a simple non-commutative extension of

---

[6] A more general definition in the context of von Neumann algebras will be provided in Part III.

[7] We recall that a map $\mathcal{P} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is positivity preserving when for any $X \in \mathcal{B}(\mathcal{H})$, $X \geq 0$ implies that $\mathcal{P}(X) \geq 0$.

the idea of conservation of mass of probability measures under classical stochastic evolutions. On the contrary, the notion of complete positivity is a truly non-commutative concept which classically reduces to the simple preservation of positivity (see Theorem 3.9 of [Paulsen, 2002]). The following generalization of the GNS representation of states, due to [Stinespring, 1955], provides a physical explanation behind the condition of complete positivity:

Theorem 0.1.10 (Stinespring's dilation theorem)**.** *Let $\mathcal{A}$ be a unital C^*-algebra, and let $\mathcal{P} : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be a completely positive map. Then there exists a Hilbert space $\mathcal{H}'$, a unital $*$-homomorphism $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H}')$, and a bounded operator $V : \mathcal{H} \to \mathcal{H}'$ with $\|\mathcal{P}(\mathbb{1})\| = \|V\|^2$ such that*

$$\mathcal{P}(X) = V^* \pi(X) V.$$

If $\mathcal{P}$ is unital, then V is an isometry. In this case, $\mathcal{H}$ can be identified with the subspace $V\mathcal{H}$ of $\mathcal{H}'$. Therefore, V^* becomes the projection $P_{\mathcal{H}}$ of $\mathcal{H}'$ onto $\mathcal{H}$. Therefore,

$$\mathcal{P}(X) = P_{\mathcal{H}} \pi(X)|_{\mathcal{H}}.$$

The above characterization of *completely positive unital* (CPU) maps in the Heisenberg picture (evolution of observables) can be recast in the dual Schrödinger picture (evolution of states) as follows: let $\mathcal{P}_* : \mathcal{T}_1(\mathcal{H}_1) \to \mathcal{T}_1(\mathcal{H}_2)$ a CPTP map, then there exists a Hilbert space $\mathcal{H}_3$, a pure state $|0\rangle \in \mathcal{H}_3$ and an isometry $V : \mathcal{H}_1 \to \mathcal{H}_2 \otimes \mathcal{H}_3$ such that for any state $\rho \in \mathcal{D}(\mathcal{H}_1)$

$$\mathcal{P}_*(\rho) = \mathrm{Tr}_{\mathcal{H}_3}(V\,(\rho \otimes |0\rangle\langle 0|)\,V^*).$$

In words, any quantum channel on a system Σ can be decomposed as a unitary evolution on the system and an environment followed by a reduction to Σ. Before ending this section, we mention one last practical characterization of completely positive maps due to [Kraus, 1971]:

Theorem 0.1.11. *Let $\mathcal{M}$ be a von Neumann algebra of operators on a Hilbert space $\mathcal{H}$ and let $\mathcal{K}$ be another Hilbert space. A linear map $\mathcal{P} : \mathcal{M} \to \mathcal{B}(\mathcal{K})$ is normal and completely positive if and only if it can be represented in the form:*

$$\mathcal{P}(X) = \sum_{j=1}^{\infty} V_j^* X V_j, \tag{0.23}$$

where $(V_j)_{j=1}^{\infty}$ are bounded operators from $\mathcal{K}$ to $\mathcal{H}$ such that the series $\sum_{j=1}^{\infty} V_j^* X V_j$ converges strongly.

0.1.6. Symmetries

A common yet elegant way of describing a physical system is via an analysis of its symmetry properties. For instance, consider the motion of non-relativistic point masses moving in a reference frame consisting of three spatial coordinates together with a time coordinate. Inertial frames are characterized by the property that free point masses travel rectilinearly and at constant speed. The coordinates of a particle between any two such frames are thus related by the *Galilei transformation*

$$(\vec{x}, t) \mapsto (\vec{x}', t') \equiv (\mathbf{R}\vec{x} + \vec{x}_0 + \vec{v}\,t, t + \tau), \qquad \mathbf{R} \in \mathrm{SO}(3),\ \tau \in \mathbb{R},\ \vec{x}_0,\ \vec{v} \in \mathbb{R}^3. \tag{0.24}$$

Here, the rotation matrix $\mathbf{R}$ represents the orientation of the new spatial axes with respect to the old ones, $\vec{x}_0$ is the shift between the respective origins of the two frames, $\vec{v}$ is their (constant) relative velocity, and τ can be interpreted as the time difference between the clocks of two inertial observers.

Transformations of the form of Equation (0.24) constitute the *Galilean group*, which contains the subgroups of *kinematical transformation* ($t' = t$) and the *Euclidean group* of spatial transformations ($\vec{x}' = \mathbf{R}\vec{x} + \vec{x}_0$, $t' = t$) as special cases. Galilean relativity is then a statement about the equivalence of the laws of mechanics in all inertial frames. More precisely, the equations of motion of a free particle should be invariant under any Galilei transformation. When adding forces (e.g. by means of an electromagnetic field), the equations of motion become invariant under a smaller group, which translates the incurred loss of symmetry.

Introducing symmetries in the quantum realm, where the notion of point particles is closely related to the measurement process, is a bit more intricate. One natural way to proceed is by restating the Galilean relativity principle as follows [Holevo, 2011]: *the statistics of an experiment is the same in any inertial frame of reference.* More precisely, since both preparation and measurement devices are macroscopic objects, they can be associated with a reference frame $(\vec{x}, t)$. Denoting the prepared state by $\omega \in \mathcal{S}$ and the measurement by the POVM $\mathbf{M} = \{M(B) : B \in \mathcal{B}(\mathbb{R}^4)\}$, a change of reference frame $g := (\mathbf{R}, \vec{x}_0, \vec{v}, \vec{\tau})$ applied to both apparatus and state results in a new state $g.\omega$ and a new measurement $g.\mathbf{M}$ such that for any Borel set $B \in \mathcal{B}(\mathbb{R}^4)$

$$g.\omega\,(g.M(B)) = \omega\,(M(B)). \tag{0.25}$$

It is natural to assume associativity of the actions of the group on the sets of states and POVMs: for any state ω and POVM $\mathbf{M}$,

$$g_1.(g_2.\omega) = (g_1 g_2).\omega, \qquad \forall B \in \mathcal{B}(\mathbb{R}^4), \ \ g_1.(g_2.M(B)) = (g_1 g_2).M(B), \tag{0.26}$$

where $g_1 g_2$ denotes the natural product operation of the group under consideration. By equivalence of all inertial frames, the maps $\omega \mapsto g.\omega$ and $\mathbf{M} \mapsto g.\mathbf{M}$ shall be considered as isomorphic. The map $\omega \mapsto g.\omega$ is affine: given a convex combination $\omega = \sum_j p_j \omega_j$ of states ω_j, for any POVM $\mathbf{M}$ and any $B \in \mathcal{B}(\mathbb{R}^4)$,

$$g.\omega\,(g.M(B)) = \omega\,(M(B)) = \sum_j p_j \omega_j\,(M(B)) = \sum_j p_j\, g.\omega_j\,(g.M(B)).$$

Since observables separate states, we conclude that

$$g.\omega = \sum_j p_j\, g.\omega_j.$$

Hence, the group of Galilei transformations acts as a group of automorphisms on the set of states, and as a group of one-to-one transformations of the set of measurements.

Of course, the logic described above generalizes to any group G acting on both states and measurements in the dual manner described in Equation (0.25). The next theorem, due to Wigner, reveals the structure of automorphisms of the set of density operators on a Hilbert space $\mathcal{H}$:

Theorem 0.1.12. *Any automorphism of the set $\mathcal{D}(\mathcal{H})$ of density operators on a separable Hilbert space $\mathcal{H}$ has the form*

$$\rho \mapsto V \rho V^*,$$

where V is a unitary or anti-unitary operators in the Hilbert space $\mathcal{H}$.

By Theorem 0.1.12 and in view of the present discussion, to any $g \in G$ one can associate either

a unitary or an antiunitary operator V_g such that for any density operator $\rho \in \mathcal{D}(\mathcal{H})$

$$g.\rho = V_g \rho V_g^*.$$

Moreover, by the associativity property (0.26), for any such ρ,

$$V_{g_2} V_{g_1} \rho V_{g_1}^* V_{g_2}^* = V_{g_2 g_1} \rho V_{g_2 g_1}^*, \qquad g_1, g_2 \in G.$$

Therefore, there is a complex-valued function $\phi(g_1, g_2)$ with $|\phi(g_1, g_2)| = 1$ such that

$$V_{g_2} V_{g_1} = \phi(g_2, g_1) V_{g_2 g_1}, \qquad g_1, g_2 \in G. \tag{0.27}$$

In what follows, we will be interested in the case when G is a connected topological group, that is, when the group is given a topology with respect to which operations of composition and inversion are continuous. Let $V_e = \mathbb{1}$, i.e. V maps the unit element e of G in the identity operators $\mathbb{1}$ on $\mathcal{H}$. Assuming moreover that the map $g \mapsto V_g$ is weakly continuous, that is for any $\varphi, \psi \in \mathcal{H}$, $g \mapsto \langle \varphi, V_g \psi \rangle$ is continuous, one can easily derive that the operators V_g, $g \in G$ will all be unitary. Such a map is called a *projective unitary representation* of G in $\mathcal{H}$. If $\phi = 1$, it is simply called a *unitary representation*.

Any projective unitary representation of the additive group $\mathbb{R}$ of one dimensional shifts reduces to a unitary one (see Proposition 3.2.1 of [Holevo, 2011]). We will see in Section 0.2 that this does not hold true in the multidimensional setting.

We end this section with a comment: we call a representation an *irreducible representation* if the only invariant closed subspaces of all operators $\{V_g : g \in G\}$ are $\{0\}$ and $\mathcal{H}$, which means that the representation is in a sense minimal. Under some regularity assumptions (which are always satisfied in finite dimensions), any representation can be decomposed into the (possibly continuous) sum of irreducible representations. This is what lead Wigner to interpret the latter as describing elementary systems ("particles").

0.2. Example: the algebra of canonical commutation relations

In this section, we introduce an example of quantum systems that will be investigated in the remaining parts of this thesis, namely bosonic systems. These are described by representations of a *symplectic group* $(\mathcal{Z}, +, \{., .\})$, that is an Abelian group $(\mathcal{Z}, +)$ equipped with a *symplectic form* $\{., .\} : \mathcal{Z} \times \mathcal{Z} \to \mathbb{R}$: for all $z, z', z'' \in \mathcal{Z}$,

$$\{0, z\} = 0$$
$$\{z, z'\} = -\{z', z\},$$
$$\{z, z' + z''\} = \{z, z'\} + \{z, z''\}.$$

The symplectic form $\{., .\}$ is said to be non-degenerate if $\{z, z'\} = 0$ for all $z' \in \mathcal{Z}$ implies that $z = 0$. Then, for any $z \in \mathcal{Z}$, define the abstract symbols W_z, called *Weyl operators*, which satisfy the so-called *canonical commutation relations*: for any $z, z' \in \mathcal{Z}$:

$$W_z W_{z'} = \mathrm{e}^{-i\{z, z'\}/2} W_{z+z'}. \tag{Weyl-Segal CCR}$$

Define an involution on the set of Weyl operators by setting

$$W_z^* = W_{-z}.$$

The norm closure of the complex span of Weyl operators becomes a unital C^*-algebra, with unit $\mathbb{1} = W_0$, called the *algebra of canonical commutation relations* (CCR). In what follows, we exclusively focus on the case where $\mathcal{Z}$ is a real vector space and $\{.,.\}$ satisfies the following additional axiom

$$\{\lambda z, z'\} = \lambda \{z, z'\}, \qquad \lambda \in \mathbb{R}, \quad z, z' \in \mathcal{Z}.$$

For instance, take $\mathcal{Z}$ to be the *phase space* $\mathbb{R}^{2d}$ (typically $d = 3$ in the case of a point particle) of point particles which is clearly in one-to-one correspondence with its associated group of *phase space translations*:

$$(\vec{y}, t) \mapsto (\vec{y}', t') := (\vec{y} + \vec{v}t + \vec{x}, t), \qquad \vec{x}, \vec{v} \in \mathbb{R}^3.$$

The next proposition provides the reason for introducing Weyl-Segal CCR in this context (see Proposition 3.3.1 of [Holevo, 2011]).

Proposition 0.2.1. *Any projective unitary representation of the group of kinematical transformations can be described by a family of unitary operators satisfying Weyl-Segal CCR.*

The following famous theorem, whose original proof was given in [Von Neumann, 2018], characterizes all the representations of the CCR algebra of a finite dimensional symplectic space. Let us first recall that any symplectic space necessarily has even dimension $2d$. Moreover, given a symplectic form $\{.,.\}$, there always exists a basis $\{e_1, ..., e_d, f_1, ..., f_d\}$ in which it takes the form

$$\{z, z'\} = \sum_{i=1}^{d} x_i y_i' - x_i' y_i, \qquad z = (x_1, ..., x_d, y_1, ..., y_d), \ z' = (x_1', ..., x_d', y_1', ..., y_d') \in \mathcal{Z}.$$

Then,

Theorem 0.2.2 (Stone-von Neumann uniqueness theorem)**.** *In the case when $\mathcal{Z}$ is $2d$-dimensional, any strongly continuous irreducible representation of the CCR algebra is unitarily equivalent to the representation $\mathbb{L}_2(\mathbb{R}^d)$. Moreover, for each $z = (x_1, ..., x_d, y_1, ..., y_d) \in \mathcal{Z}$, the operator W_z can be written as a product of commuting operators:*

$$W_z = \prod_{j=1}^{d} W_{x_j, y_j} = \prod_{j=1}^{d} \mathrm{e}^{ix_j y_j/2} V_{x_j} U_{y_j},$$

where for each coordinate j, $(U_{x_j})_{x_j \in \mathbb{R}}$ and $(V_{y_j})_{y_j \in \mathbb{R}}$ denote two one-parameter groups of unitary operators on $\mathbb{L}_2(\mathbb{R}^d)$ of associated self-adjoint generators $Q_j : \psi \mapsto (\mathbf{x} \mapsto x_j \psi(z))$, $P_j : \psi \mapsto \frac{1}{i} \partial_{x_i} \psi$, densely defined for example on the subspace of Schwartz functions, so that

$$V_{x_j} = \mathrm{e}^{ix_j P_j}, \qquad x_j \in \mathbb{R},$$
$$U_{v_j} = \mathrm{e}^{-iy_j Q_j}, \qquad v_j \in \mathbb{R}.$$

On the dense subspace of Schwartz functions, one can show that the following relations hold:

$$[Q_j, P_j] = i\delta_{jk} \mathbb{1}_{\mathbb{L}_2(\mathbb{R}^{2d})}, \qquad [P_j, P_k] = 0, \qquad [Q_j, Q_k] = 0; \qquad \text{(Heisenberg CCR)}$$

$$W_z = \exp\left(i \sum_{j=1}^{d} (x_j P_j - y_j Q_j)\right), \qquad z = (x_1, y_1, ..., x_d, y_d) \in \mathcal{Z} \qquad (0.28)$$

For each j, we say that the operators P_j and Q_j which satisfy Heisenberg CCR are *canonically conjugate observables.* The operators W_z are also called Weyl *displacement operators*, the latter name being justified by the relation

$$W_z R_j W_{-z} = R_j + z_j \mathbb{1} \quad \forall j \in \{1, 2, \ldots, 2d\}, \qquad (0.29)$$

where the operators R_j form the coordinates of the vector

$$\mathbf{R} := (Q_1, P_1, ..., Q_d, P_d). \qquad (0.30)$$

A natural Hamiltonian corresponding to the system we have just defined is the *quantum harmonic oscillator* $H = \hbar \sum_{j=1}^{d} \omega_j (N_j + \frac{1}{2})$, where $\hbar$ is the reduced Planck constant, ω_j is the angular frequency of mode j and N_j is the so-called *number operator* that is defined on its own domain, containing the space of Schwartz functions, as:

$$N_j := \frac{1}{2}(Q_j - iP_j)(Q_j + iP_j).$$

The term P_j^2 in the above sum corresponds to the contribution to the kinetic energy of a quantum particle in the direction j, whereas the terms Q_j^2 represent its potential energy. As a matter of fact, the Hamiltonian H also describes the energy of a quantum optical system: the unbounded *creation and anihilation operators* are densely defined as follows:

$$a_j = \frac{1}{\sqrt{2}} (Q_j + iP_j), \qquad a_j^* = \frac{1}{\sqrt{2}} (Q_j - iP_j). \qquad (0.31)$$

In terms of a_j and a_j^*, the Heisenberg CCR takes the form

$$[a_j, a_k^*] = \delta_{jk} \mathbb{1}. \qquad (0.32)$$

Using this relation, the number operator N_j takes the form

$$N_j = a_j^* a_j \qquad\qquad N = \sum_{j=1}^{d} N_j.$$

Using Equation (0.32), one shows the existence of an orthonormal basis $\{|n_j\rangle\}_{n_j \in \mathbb{N}}$ of $\mathbb{L}_2(\mathbb{R})$ such that for any $n_j \in \mathbb{N}$,

$$a_j |n_j\rangle = \sqrt{n_j} |n_j - 1\rangle, \qquad a_j^* |n_j\rangle = \sqrt{n_j + 1} |n_j + 1\rangle.$$

This basis diagonalizes the number operator N_j, showing therefore that N_j has discrete spectrum:

$$N_j |n_j\rangle = n_j |n_j\rangle.$$

For sake of simplicity, we will work in the von Neumann algebra generated by the Weyl operators represented on the separable Hilbert space $\mathcal{H} := \mathbb{L}_2(\mathbb{R}^d)$. As a matter of fact, it is known to be the

whole algebra $\mathcal{B}(\mathcal{H})$ of bounded linear operators on $\mathbb{L}_2(\mathbb{R}^d)$.

Quantum characteristic function   A quantum state $\rho \in \mathcal{D}(\mathcal{H})$ of the system is uniquely defined on $\mathcal{Z}$ through its *quantum characteristic function*:

$$\mathcal{F}_\rho^{\mathrm{q}}(z) \coloneqq \mathrm{Tr}\left(\rho\, W_z\right). \tag{0.33}$$

The following theorem provides a justification of the above definition (see Theorem 5.3.3 of [Holevo, 2011]):

Theorem 0.2.3 (Non-commutative Parseval's relation). *The function given by (0.33) extends uniquely to a map $T \mapsto \mathcal{F}_T^{\mathrm{q}}$ from the Hilbert space $\mathcal{T}_2(\mathcal{H})$ onto $\mathbb{L}_2(\mathbb{R}^{2d})$, so that*

$$\mathrm{Tr}(T_1^* T_2) = (2\pi)^{-d} \int_{\mathbb{R}^{2d}} \overline{\mathcal{F}_{T_1}^{\mathrm{q}}(z)} \mathcal{F}_{T_2}^{\mathrm{q}}(z) dz, \tag{0.34}$$

(where the notation $\overline{f(z)}$ denotes the complex conjugate of $f(z)$).

The inverse Fourier transform of the characteristic function

$$\mathcal{F}_{\mathcal{F}_\rho^{\mathrm{q}}}^{-1}(u) \coloneqq (2\pi)^{-d} \int_{\mathbb{R}^{2d}} \mathrm{e}^{-i(u,z)} \, \mathcal{F}_\rho^{\mathrm{q}}(z)\, dz \tag{0.35}$$

is called the *Wigner function* of the state[8]. Under this convention the integral of the Wigner function is equal to $(2\pi)^d$. Contrary to the classical case, a Wigner function can in general be negative, and therefore cannot be interpreted as a probability density associated to a state.

A quantum state ρ is said to have finite k^{th} moment, for $k \in \mathbb{N}$, if

$$\mathrm{Tr}\left(\rho\, |R_j|^k\right) < \infty\,, \quad \forall\, j \in \{1, 2, \ldots, 2d\}\,,$$

where the operator R_j denotes the jth entry of the vector $\mathbf{R}$.

Gaussian states   An important class of states for a continuous variable quantum system are the Gaussian states. They play a fundamental role in the study of various quantum information processing tasks involving continuous variable quantum systems which are relevant in quantum computation and quantum cryptography. Important examples of Gaussian states arise in quantum optics and include coherent states, squeezed states and thermal states. A Gaussian state ρ is defined via its characteristic function

$$\mathcal{F}_\rho^{\mathrm{q}}(z) = \mathrm{e}^{i(\mu,z) - \frac{1}{2}(z, \Sigma z)} \qquad z \in \mathcal{Z}, \tag{0.36}$$

where μ is the mean vector of the state ρ, with elements

$$\mu_i \coloneqq \mathrm{Tr}(\rho\, R_i), \quad i = 1, 2, \ldots, 2d, \tag{0.37}$$

and Σ is its $2n \times 2n$ (real, symmetric) covariance matrix, with elements

$$\Sigma_{ij} \coloneqq \frac{1}{2}\,\mathrm{Tr}(\rho\,(R_i^c R_j^c + R_j^c R_i^c)), \tag{0.38}$$

---

[8]Note that the above definition differs from the usual one by a factor of $(2\pi)^{-d}$.

where $R_i^c \coloneqq R_i - \mathrm{Tr}(\rho\, R_i)$. The characteristic function given by Equation (0.36) is exactly the one of a $2d$-dimensional Gaussian vector of mean μ and covariance Σ. The Gaussian state is said to be *centered* if $\mu = 0$. Gaussian states have some very useful properties (see [De Palma et al., 2015a, Holevo, 2012]). For example, the linear span of all Gaussian states is dense in the Banach space, $\mathcal{T}_1(\mathcal{H})$, of trace class operators. Moreover, any Gaussian state can be written in a rather simple way as follows (see Lemma 12.21 and Theorem 12.22 of [Holevo, 2011]).

Theorem 0.2.4. *A Gaussian state ρ with covariance matrix Σ is invertible if and only if*

$$\det\left(\Sigma + \frac{i}{2}\Omega\right) \neq 0. \tag{0.39}$$

Moreover, any Gaussian state $\rho \equiv \rho_{\mu,\Sigma}$ of mean μ and covariance matrix Σ can be written as

$$\rho_{\mu,\Sigma} = W_\mu\, \rho_{0,\Sigma}\, W_{-\mu}, \tag{0.40}$$

where $\rho_{0,\Sigma}$ is a centered Gaussian state of covariance matrix Σ, and W_z denotes the Weyl operator given by Equation (0.28). Further, a centered, invertible Gaussian state can be expressed as follows:

$$\rho_{0,\Sigma} = C\, e^{-\mathbf{R}^T \Gamma \mathbf{R}}, \tag{0.41}$$

where

$$C = \left[\det\left(\Sigma + \frac{i}{2}\Omega\right)\right]^{-1/2}, \tag{0.42}$$

where Ω is the matrix representing the symplectic form $\{.,.\}$ in the basis $\{e_1, ..., e_d, f_1, ..., f_d\}$:

$$\Omega = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}^{\oplus n}. \tag{0.43}$$

The matrix Γ is then defined through the relation

$$2\,\Omega^{-1}\Sigma = \cot(\Gamma\Omega). \tag{0.44}$$

Quantum Schwartz operators   In this section we introduce the framework of Schwartz operators in the CCR algebra as recently defined in [Keyl et al., 2016] and give some of their useful properties. We first recall that a Schwartz function on $\mathbb{R}^n$ is a smooth function $\varphi : \mathbb{R}^n \mapsto \mathbb{C}$ for which

$$\sup_{x \in \mathbb{R}^n} \left| x_1^{\alpha_1} \dots x_n^{\alpha_n} \frac{\partial^{\beta_1}}{\partial x_1^{\beta_1}} \dots \frac{\partial^{\beta_n}}{\partial x_n^{\beta_n}} \varphi(x) \right| < \infty \tag{0.45}$$

for all $\alpha, \beta \in \mathbb{N}^n$. The set of Schwartz functions is denoted by $\mathcal{S}(\mathbb{R}^n)$. In analogy with the theory of Schwartz functions, the authors of [Keyl et al., 2016] defined a *Schwartz operator* on $\mathcal{H} \coloneqq \mathbb{L}_2(\mathbb{R}^n)$ to be a Hilbert Schmidt operator T such that for any $\alpha, \alpha', \beta, \beta' \in \mathbb{N}^n$,

$$\sup\left\{ |\langle P_1^{\beta_1} \dots P_n^{\beta_n} Q_1^{\alpha_1} \dots Q_n^{\alpha_n} \psi, T P_1^{\beta_1'} \dots P_n^{\beta_n'} Q_1^{\alpha_1'} \dots Q_n^{\alpha_n'} \varphi \rangle| \mid \|\psi\|, \|\varphi\| \leq 1 \right\} < \infty, \tag{0.46}$$

where $\langle .,. \rangle$ denotes the inner product on $\mathcal{H}$, and the sup is taken over square integrable Schwartz functions φ, ψ. The set of Schwartz operators is denoted by $\mathcal{S}(\mathcal{H})$. A Hilbert Schmidt operator T belongs to $\mathcal{S}(\mathcal{H})$ if and only if its characteristic function χ_T is in $\mathcal{S}(\mathbb{R}^{2n})$ (see Proposition 3.18 of [Keyl et al., 2016]). As a consequence, any Gaussian state is a Schwartz operator.

Next, an operator A is said to be *polynomially bounded* if $\mathrm{dom}(N) \subset \mathrm{dom}(A), \mathrm{dom}(A^*)$ and $\|A(1+N)^{-1}\|_\infty, \|A^*(1+N)^{-1}\|_\infty < \infty$.

Quantum channels on the CCR algebra   Given the algebra $\mathcal{B}(\mathcal{H})$, $\mathcal{H} = \mathbb{L}_2(\mathbb{R}^n)$, corresponding to a quantum bosonic system of n modes, a *Gaussian operation* is a CPTP map $\mathcal{P}_*$ on $\mathcal{T}_1(\mathcal{H})$ that preserves the set of Gaussian states. Any such operation is characterized by its action on the Fourier side, that is on characteristic functions: given a state $\rho \in \mathcal{D}(\mathcal{H})$ of characteristic function $\mathcal{F}_\rho^{\mathrm{q}}$:

$$\mathcal{F}_{\mathcal{P}_*(\rho)}^{\mathrm{q}} : z \mapsto \mathcal{F}_\rho^{\mathrm{q}}(Xz)\, \mathrm{e}^{i(\mu', z) - \frac{1}{2}(z, Yz)},$$

where $\mu' \in \mathbb{R}^{2n}$, and X, Y are two $2n \times 2n$ real matrices acting on the phase space such that $Y^T = Y$ and $Y + i(\Omega - X\Omega X^T) \geq 0$. The action of $\mathcal{P}_*$ on a Gaussian state ρ_G is then characterized by its action on the mean vector μ and covariance matrix Σ of ρ_G:

$$\Sigma \mapsto X\Sigma X^T + Y$$
$$\mu \mapsto X\mu + \mu'$$

Any such operation admits a Stinespring dilation of the following form: for any $\rho \in \mathcal{D}(\mathcal{H})$:

$$\mathcal{P}_*(\rho) = \mathrm{Tr}_E\left[U_G\left(\rho \otimes \rho_G\right)U_G^*\right],$$

where ρ_G is a Gaussian state on an "environment" system $\mathcal{H}_E = \mathbb{L}_2(\mathbb{R}^m)$, $m \in \mathbb{N}$, and where U_G is a Gaussian unitary acting on $\mathcal{H} \otimes \mathcal{H}_E$. Such a unitary is determined by $Y = 0$ and $X \equiv S$ is a symplectic matrix. Among the class of Gaussian unitaries, the *beamsplitter* U_λ plays an important role. It is defined on a $2n$ mode quantum bosonic system through its symplectic matrix: given $0 < \lambda < 1$,

$$S_\lambda := \begin{pmatrix} \sqrt{\lambda}\,\mathbb{1}_n & \sqrt{1-\lambda}\,\mathbb{1}_n \\ \sqrt{1-\lambda}\,\mathbb{1}_n & -\sqrt{\lambda}\,\mathbb{1}_n \end{pmatrix} \otimes \mathbb{1}_2. \tag{0.47}$$

0.3. Interlude: von Neumann algebras

We start by recalling that a (concrete) *von Neumann algebra*[9] $\mathcal{M}$ on a Hilbert $\mathcal{H}$ is a self-adjoint, weakly closed subalgebra of the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on $\mathcal{H}$, which contains the identity operator $\mathbb{1}$. A functional ω on a von Neumann algebra $\mathcal{M}$ is said to be *normal* if for any increasing net $\{X_\alpha\}$ of positive operators in $\mathcal{M}$, $\omega(\mathrm{l.u.b.}\,X_\alpha) = \mathrm{l.u.b.}\,\omega(X_\alpha)$, where $\mathrm{l.u.b.}$ stands for the least upper bound of a net. Normal functionals form a subspace of the space of functionals $\mathcal{M}^*$ on $\mathcal{M}$, called the *predual* of $\mathcal{M}$, and denoted by $\mathcal{M}_*$. The subset of positive normal functionals is denoted by $\mathcal{M}_*^+$. A *state* on the von Neumann algebra $\mathcal{M}$ is a positive linear functional $\omega : \mathcal{M} \to \mathbb{C}$ such that $\omega(\mathbb{1}) = 1$. A normal state ω is characterized by the existence of a density operator ρ, i.e. a non-negative trace-class operator ρ on $\mathcal{H}$ with $\mathrm{Tr}(\rho) = 1$, such that (see Theorem 2.4.21. of [Bratteli and Robinson, 1979])

$$\omega(X) = \mathrm{Tr}(\rho X), \quad X \in \mathcal{M}. \tag{0.48}$$

A normal state is said to be *faithful* if for any positive element $X \in \mathcal{M}$, $\omega(X) = 0$ implies that $X = 0$. A von Neumann algebra in *standard form* is a quadruple $(\mathcal{M}, \mathcal{K}, J, \mathcal{K}^+)$ where $\mathcal{K}$ is a Hilbert space, $\mathcal{M} \subset \mathcal{B}(\mathcal{K})$ is a von Neumann algebra, J is an anti-unitary involution on $\mathcal{K}$ and $\mathcal{K}^+$ is a cone in $\mathcal{K}$ such

---

[9]A more abstract (yet equivalent) definition which does not require to mention any Hilbert space can be given by saying that $\mathcal{M}$ is the dual of a Banach space (see e.g. [Bratteli and Robinson, 1979].

that:

(i) $\mathcal{K}^+$ is self-dual, i.e. $\mathcal{K}^+ = \{x \in \mathcal{K}; \langle y, x \rangle \geq 0 \ \forall y \in \mathcal{K}^+\}$;

(ii) $J\mathcal{M}J = \mathcal{M}'$;

(iii) $JXJ = X^*$ for $X \in \mathcal{M} \cap \mathcal{M}'$;

(iv) $J\psi = \psi$ for $\psi \in \mathcal{K}^+$;

(v) $XJXK^+ \subset \mathcal{K}^+$ for $X \in \mathcal{M}$;

where in (ii), $\mathcal{M}'$ denotes the commutant of $\mathcal{M}$ in $\mathcal{B}(\mathcal{H})$. A quadruple $(\pi, \mathcal{K}, J, \mathcal{K}^+)$ is a *standard representation* of the von Neumann algebra $\mathcal{M}$ if $\pi : \mathcal{M} \to \mathcal{B}(\mathcal{K})$ is a faithful representation and $(\pi(\mathcal{M}), \mathcal{K}, J, \mathcal{K}^+)$ is in standard form. It is a celebrated result in operator algebras that a standard representation always exists, which means that any von Neumann algebra can be seen as a standard von Neumann algebra acting on a Hilbert space $\mathcal{K}$, up to some isomorphism π. Moreover, if $(\pi_1, \mathcal{H}_1, J_1, \mathcal{H}_1^+)$ and $(\pi_2, \mathcal{K}_2, J_2, \mathcal{K}_2^+)$ are two standard representations of $\mathcal{M}$, then there exists a unique isometry $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that $U\pi_1(X)U^* = \pi_2(X)$ for all $X \in \mathcal{M}$, $UJ_1U^* = J_2$, and $U\mathcal{K}_1^+ = \mathcal{K}_2^+$. The following basic theorem of operator algebra will prove useful for the definition of the relative entropy in Section 1.2.

Theorem 0.3.1. *Let $\mathcal{M}$ be a von Neumann algebra and $(\pi(\mathcal{M}), \mathcal{K}, J, \mathcal{K}^+)$ a standard representation of $\mathcal{M}$. For any positive normal functional ω on $\mathcal{M}$, there exists a unique $\Omega_\omega \in \mathcal{K}^+$ such that for all $X \in \mathcal{M}$,*

$$\omega(X) = \langle \Omega_\omega, \pi(X) \Omega_\omega \rangle. \tag{0.49}$$

Moreover, the map $\mathcal{M}_*^+ \ni \omega \mapsto \Omega_\omega \in \mathcal{H}^+$ is a bijection.

Example 0.3.2 (Classical probability theory). Let $(\Omega, \mathcal{F}, \mu)$ be a σ-finite measure space. Define the von Neumann algebra $\mathbb{L}_\infty(\Omega, \mathcal{F}, \mu)$ of bounded, measurable functions acting by pointwise left multiplication on the Hilbert space $\mathbb{L}_2(\Omega, \mathcal{F}, \mu)$ of square integrable functions. This means that to any $f \in \mathbb{L}_\infty(\Omega, \mathcal{F}, \mu)$, one can associate the operator $L_f : \mathbb{L}_2(\Omega, \mathcal{F}, \mu) \to \mathbb{L}_2(\Omega, \mathcal{F}, \mu)$ defined by

$$L_f(g)(x) = f(x)g(x), \quad \forall g \in \mathbb{L}_2(\Omega, \mathcal{F}, \mu), \quad x \in \Omega. \tag{0.50}$$

The map $L : \mathbb{L}_\infty(\Omega, \mathcal{F}, \mu) \ni f \mapsto L_f$ is a faithful representation, so that the quadruple $(L(\mathbb{L}_\infty(\Omega, \mathcal{F}, \mu)), \mathbb{L}_2(\Omega, \mathcal{F}, \mu), \bar{}, \mathbb{L}_2(\Omega, \mathcal{F}, \mu)_+)$ is in standard form, where $^-$ is the usual complex conjugation and $\mathbb{L}_2(\Omega, \mathcal{F}, \mu)_+$ denotes the set of positive square-integrable functions with respect to μ. We use the same notation for normal states and their associated probability measures, which is justified by the above mentioned isometry. Now any function $f \in \mathbb{L}_1(\Omega, \mathcal{F}, \mu)$ such that $\|f\|_1 = 1$ represents a state ω_f on $\mathbb{L}_\infty(\Omega, \mathcal{F}, \mu)$ via the relation:

$$\omega_f(h) := \int_\Omega f(x)h(x)\mu(dx), \quad \forall h \in \mathbb{L}_\infty(\Omega, \mathcal{F}, \mu). \tag{0.51}$$

Indeed, one has $\int_\Omega f(x)1\mu(dx) = \|f\|_1 = 1$, 1 encoding for the identity element in $\mathbb{L}_\infty(\Omega, \mathcal{F}, \mu)$, and one easily checks that

$$\omega_f(h) = \langle \sqrt{f}, L_h \sqrt{f} \rangle, \tag{0.52}$$

so that ω_f is indeed normal. Actually any normal functional can be written in this way, so that $\mathbb{L}_\infty(\Omega, \mathcal{F}, \mu)_* \cong \mathbb{L}_1(\Omega, \mathcal{F}, \mu)$. This is the space of (complex) measures that are absolutely continuous with respect to μ.

Example 0.3.3 (Quantum systems)**.** Take $\mathcal{M} \coloneqq \mathcal{B}(\mathcal{H})$ to be the von Neumann algebra of all bounded operators on a finite dimensional Hilbert space $\mathcal{H}$. Then any density operator ρ provides a state via Equation (0.48). Then $\mathcal{M}_*$ is identified with the space of trace-class operators acting on $\mathcal{H}$. The map $\pi : \mathcal{B}(\mathcal{H}) \ni A \mapsto L_A$, where for any $A \in \mathcal{B}(\mathcal{H})$ $L_A : \mathcal{T}_2(\mathcal{H}) \mapsto \mathcal{T}_2(\mathcal{H})$ is the operator of left multiplication by A, is a faithful representation, turning $(\pi(\mathcal{B}(\mathcal{H})), \mathcal{T}_2(\mathcal{H}), {}^*, \mathcal{T}_2(\mathcal{H})_+)$ into a standard form, where * is the usual adjoint, and $\mathcal{T}_2(\mathcal{H})_+$ is the space of non-negative Hilbert Schmidt operators on $\mathcal{H}$. As already discussed in Equation (0.48) any positive, normal functional ω on $\mathcal{B}(\mathcal{H})$ can be associated with a trace-class operator ρ so that for any $A \in \mathcal{B}(\mathcal{H})$,

$$\omega(A) = \mathrm{Tr}(\rho A) = \mathrm{Tr}(\sqrt{\rho}A\sqrt{\rho}) = \langle \sqrt{\rho}, L_A(\sqrt{\rho})\rangle_{\mathrm{HS}}, \tag{0.53}$$

where $\langle .,.\rangle_{\mathrm{HS}}$ denotes the Hilbert-Schmidt inner product on $\mathcal{T}_2(\mathcal{H})$. Hence, one identifies Ω_ω with $\sqrt{\rho}$.

Chapter 1.

Norms, relative entropies, Fisher metrics

In order to make any quantitative statement, mathematicians are usually lead to the introduction of certain distance measures. In the specific study of quantum tasks and evolutions, numerous measures can be defined, some of which admitting an operational interpretation (see Part V). In this chapter, we introduce all the distance measures that will be used throughout this thesis. They are roughly organized in three categories- namely *norms*, *relative entropies* and *Fisher information metrics*- that are deeply connected to one another.

Layout of the chapter: In Section 1.1, we introduce and give basic properties of three important classes of non-commutative Banach spaces, namely Shatten, $\mathbb{L}_p$ and amalgamated $\mathbb{L}_p$ spaces. We show how the latter can be used to define the so-called notion of a completely bounded $\mathbb{L}_p$ norm that will prove very useful in Chapter 10. Section 1.2 deals with the closely connected notion of relative entropies. Finally, the Hessians of the latter, also known as Fisher information metrics, are introduced in Section 1.3 and shown to provide the set of quantum states with a Riemannian structure. Similarly to the commutative case, this framework turns out to be useful in parameter estimation of quantum states (see Section 13.2).

1.1. Norms

1.1.1. Operator norms

Schatten norms   In Chapter 0, we introduced the two dual concepts of states and observables associated to a (quantum) system. For each of these notions, we have already seen a natural notion of distance, that we recall here: let $\mathcal{M}$ be a von Neumann subalgebra of the set $\mathcal{B}(\mathcal{H})$ of bounded operators on a given separable Hilbert space $(\mathcal{H}, \langle ., . \rangle)$. The norm of an operator $X \in \mathcal{M}$ is defined as

$$\|X\|_\infty := \sup_{\psi \in \mathcal{H} \backslash \{0\}} \frac{\|X\psi\|}{\|\psi\|}, \tag{1.1}$$

where $\|.\|$ denotes the norm associated to the inner product $\langle ., . \rangle$ on $\mathcal{H}$. The norm $\|.\|_\infty$ quantifies the amplitude of observables evolving in the so-called *Heisenberg picture*. In the dual *Schrödinger picture*,

the norm of a normal linear functional ω is then defined as:

$$\|\omega\|_1 := \sup_{0 \neq X \in \mathcal{M}} \frac{|\omega(X)|}{\|X\|_\infty}, \qquad \omega \in \mathcal{M}_*, \tag{1.2}$$

which provides $\mathcal{M}_*$ with a Banach space structure.

In the case when the von Neumann algebra $\mathcal{M}$ possesses a semi-finite trace τ, we define, for $p \geq 1$, the so-called noncommutative $\mathbb{L}_p(\mathcal{M})$ space as the completion of $\{X \in \mathcal{M} : \tau(|X|^p) < \infty\}$ in the following norm:[1]

$$\|X\|_{\mathbb{L}_p(\mathcal{M})} := \left[\tau(|X|^p)\right]^{\frac{1}{p}} .$$

It can be shown that for any element $\omega \in \mathcal{M}_*$, there exists a density $\rho \in \mathbb{L}_1(\tau)$ such that $\omega(.) = \tau(\rho.)$. Moreover, in this case we have (see Proposition 1.1 of [Davies and Lindsay, 1992])

$$\|\rho\|_{\mathbb{L}_1(\mathcal{M})} = \|\omega\|_1 . \tag{1.3}$$

For this reason $\|.\|_1$ is commonly referred to as the *trace norm*. Its associated trace distance has the appropriate operational interpretation of a measure of indistinguishability between two states for an external observer allowed to perform any measurement on the system [Fuchs and van de Graaf, 2006, Audenaert et al., 2008, Jaksic et al., 2012]: assume given one of two faithful normal states $\omega, \nu \in \mathcal{M}_*$, and consider the total error probability of guessing it, given a test, i.e. a two outcomes POVM $(T, \mathbb{1} - T)$, $T \in \mathcal{M}$, $0 \leq T \leq \mathbb{1}$:

$$P_{\mathrm{err}}(\nu, \omega, T) := \nu(T) + \omega(\mathbb{1} - T) .$$

The *quantum Neyman-Pearson lemma* states that the minimum total error probability is given by

$$\inf_T P_{\mathrm{err}}(\nu, \omega, T) = 1 - \frac{1}{2}\|\omega - \nu\|_1.$$

Moreover, the $\mathbb{L}_p(\mathcal{M})$ spaces interpolate between $\mathbb{L}_1(\mathcal{M})$ and $\mathcal{M}$.

In the case when $\mathcal{M} = \mathcal{B}(\mathcal{H})$ and $\tau = \mathrm{Tr}$ denotes the usual trace on $\mathcal{B}(\mathcal{H})$, each $\mathbb{L}_p(\mathcal{M})$ space is more commonly referred to as the *Schatten class* of order p and denoted by $\mathcal{T}_p(\mathcal{H})$. Denoting by $\mathcal{F}(\mathcal{H})$, resp. $\mathcal{K}(\mathcal{H})$, the space of finite rank, resp. *compact operators* on $\mathcal{H}$, the following hierarchy of spaces is satisfied: for any $1 \leq p \leq q \leq \infty$

$$\|T\|_q \leq \|T\|_p \quad \Rightarrow \quad \mathcal{F}(\mathcal{H}) \subset \mathcal{T}_1(\mathcal{H}) \subset \mathcal{T}_p(\mathcal{H}) \subset \mathcal{T}_q(\mathcal{H}) \subset \mathcal{K}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H}) .$$

The case $p = 2$ plays a special role. Indeed, $\mathcal{T}_2(\mathcal{H})$ is a Hilbert space called the space of *Hilbert Schmidt* operators, with corresponding Hilbert Schmidt inner product defined as

$$\forall A, B \in \mathcal{T}_2(\mathcal{H}), \qquad \langle A, B \rangle_{\mathrm{HS}} := \mathrm{Tr}(A^* B) . \tag{1.4}$$

Amalgamated $\mathbb{L}_p$ norms Let $(\mathcal{M}, \tau)$ be a finite von Neumann algebra with an associated normalized trace τ ($\tau(\mathbb{1}_{\mathcal{M}}) = 1$). Then, given a von Neumann subalgebra $\mathcal{N} \subset \mathcal{M}$, the following *amalgamated $\mathbb{L}_p$ norms* were defined in [Junge and Parcet, 2010] (see also [Gao et al., 2017]): given

---

[1] The study of non-commutative $\mathbb{L}_p$ spaces with respect to a trace on general von Neumann algebras was first developed in [Dixmier, 1953, Segal, 1953, Kunze, 1958, Gohberg and Kreĭn, 1969, Simon, 2010, Segal, 1953].

$1 \le p, q, r \le \infty$ with $\frac{1}{r} = |\frac{1}{p} - \frac{1}{q}|$, define

$$\|X\|_{\mathbb{L}_p^q(\mathcal{N} \subset \mathcal{M})} = \begin{cases} \inf_{X = AYB} \|A\|_{\mathbb{L}_{2r}(\mathcal{N})} \|Y\|_{\mathbb{L}_q(\mathcal{M})} \|B\|_{\mathbb{L}_{2r}(\mathcal{N})} & p \le q\,, \\ \sup_{\|A\|_{\mathbb{L}_{2r}(\mathcal{N})} = \|B\|_{\mathbb{L}_{2r}(\mathcal{N})} \le 1} \|AXB\|_{\mathbb{L}_q(\mathcal{M})} & q \le p\,, \end{cases} \tag{1.5}$$

where A, B are elements in $\mathbb{L}_{2r}(\mathcal{N})$. For $\mathcal{N} = \mathcal{M}$, we just find another description $\mathbb{L}_p^q(\mathcal{N} \subset \mathcal{M}) = \mathbb{L}_p(\mathcal{M})$. Note that for a selfadjoint element X, we may assume $A = A^*$ in Equation (1.5). By Hölder's inequality, $\mathbb{L}_p^q(\mathcal{N} \subset \mathcal{M}) \subset \mathbb{L}_p(\mathcal{M})$. The Banach spaces $\mathbb{L}_p^q(\mathcal{N} \subset \mathcal{M})$ are then completions of $\mathcal{M}$ with respect to the above norms. In the particular case when $\mathcal{M} = \mathbb{M}_k(\mathcal{N})$ is the algebra of k by k matrices with coefficients in $\mathcal{N}$, the spaces $\mathbb{L}_p^q(\mathcal{N} \subset \mathcal{M}) \equiv S_p^k(\mathbb{L}_q(\mathcal{N}))$ coincide with Pisier's vector-valued $\mathbb{L}_p$ spaces [Pisier, 1993].

Quantum weighted $\mathbb{L}_p$ norms   The Schatten norms introduced above should be understood as quantum extensions of the usual $\mathbb{L}_p(\mathbb{N})$ norms associated with the counting measure. In Parts III and IV, we will also deal with non-commutative extensions of the notion of weighted $\mathbb{L}_p$ norms. Typically, in the context of the convergence and contractivity properties of an ergodic Markov semigroup, the weight is provided by the invariant measure of the evolution. Luckily, quantum extensions of these norms have already extensively been studied in various degrees of generality: in the general von Neumann algebraic setting, two different, yet isomorphic, definitions for weighted $\mathbb{L}_p$ spaces were defined in [Haagerup, 1979, Kosaki, 1984]. In our simple case of the tracial von Neumann algebra $\mathcal{B}(\mathcal{H})$, given a faithful state represented by a density operator $\sigma \in \mathcal{D}(\mathcal{H})$, Kosaki's version of the *quantum weighted norm* of order p with respect to σ is defined as follows [Olkiewicz and Zegarlinski, 1999]: for $X, Y \in \mathcal{B}(\mathcal{H})$, define the *weighted inner product* (also known as KMS-*inner product*) $\langle X, Y \rangle_\sigma$ as

$$\langle X, Y \rangle_\sigma := \operatorname{Tr}\left[\sigma^{\frac{1}{2}} X^* \sigma^{\frac{1}{2}} Y\right]. \tag{1.6}$$

The above inner product has the advantage of preserving the complete positivity under the adjoint operation. It also induces a norm $\|.\|_{\mathbb{L}_2(\sigma)}$ on $\mathcal{B}(\mathcal{H})$:

$$\|X\|_{\mathbb{L}_2(\sigma)} := \sqrt{\langle X, X \rangle_\sigma}, \tag{1.7}$$

and we denote by $\mathbb{L}_2(\sigma)$ the completion of the space of bounded operators $\mathcal{B}(\mathcal{H})$ under this norm. In fact, let σ have the following eigenvalue decomposition $\sigma = \sum_{i=1}^\infty \lambda_i |e_i\rangle\langle e_i|$. Then, given any (possibly unbounded) closed operator $(X, \operatorname{dom}(X))$ on $\mathcal{H}$ such that $e_i \in \operatorname{dom}(X)$ for all $i \in \mathbb{N}$, define the $\mathbb{L}_2(\sigma)$ norm of X as follows:

$$\|X\|_{\mathbb{L}_2(\sigma)} = \left(\sum_{i,j=1}^\infty \sqrt{\lambda_i \lambda_j} \, |\langle e_i, X e_j \rangle|^2 \right)^{1/2}, \tag{1.8}$$

whenever the quantity on the right hand side is finite. In this case, X is said to be *square integrable* with respect to the state σ. One can easily verify that the above quantity coincides with Equation (1.7) whenever X is bounded. Any two operators X and Y are said to be equivalent if for all $i \in \mathbb{N}$, $X e_i = Y e_i$, and we identify $\mathbb{L}_2(\sigma)$ with the equivalence class obtained by this procedure. Indeed:

Theorem 1.1.1. *The space $\mathbb{L}_2(\sigma)$ can be identified with the equivalence class of square integrable operators with respect to σ. This means that whenever X is a square integrable operator with respect to the state σ, there exists a Cauchy sequence $\{X_n\}_{n \in \mathbb{N}}$ of bounded operators such that $\|X_n - X\|_{\mathbb{L}_2(\sigma)} \to 0$*

as $n \to \infty$. *Conversely, any Cauchy sequence* $\{X_n\}_{n \in \mathbb{N}}$ *of bounded operators converges to an operator* X *that is square integrable with respect to* σ.

Proof. Let P_n denote the following spectral projector

$$P_n = \sum_{i=1}^{n} |e_i\rangle\langle e_i| .$$

Define the sequence of bounded operators $\{X_n\}_{n \in \mathbb{N}}$ by $X_n = P_n X P_n$. This sequence converges to X in $\mathbb{L}_2(\sigma)$, since

$$\|X_n - X\|^2_{\mathbb{L}_2(\sigma)} = \sum_{i,j=1}^{\infty} \sqrt{\lambda_i \lambda_j} \, |\langle e_i, (X - P_n X P_n) e_j \rangle|^2$$

$$= \sum_{i,j=n+1}^{\infty} \sqrt{\lambda_i \lambda_j} \, |\langle e_i, X e_j \rangle|^2 \underset{n \to \infty}{\to} 0 ,$$

since $\|X\|_{\mathbb{L}_2(\sigma)} < \infty$. The converse is obvious by definition.

$\square$

Next proposition relates the space $\mathbb{L}_2(\sigma)$ to the one of Hilbert Schmidt operators on $\mathcal{H}$.

Proposition 1.1.2. *A closed, densely defined operator* X *with* $e_i \in \mathrm{dom}(X)$ *for any* $i \in \mathbb{N}$ *is in* $\mathbb{L}_2(\sigma)$ *if and only it* $\sigma^{\frac{1}{4}} X$ *extends to the range of* $\sigma^{\frac{1}{4}}$ *so that* $\sigma^{\frac{1}{4}} X \sigma^{\frac{1}{4}}$ *is Hilbert Schmidt.*

Proof. If $X \in \mathbb{L}_2(\sigma)$, the extension is given by the following expression: for any $\psi := \sum_{j=1}^{\infty} c_j e_j \in \mathcal{H}$:

$$\sigma^{\frac{1}{4}} X \left(\sum_{j=1}^{\infty} \lambda_j^{\frac{1}{4}} c_j e_j \right) = \sum_{j=1}^{\infty} \lambda_j^{\frac{1}{2}} c_j X e_j , \qquad (1.9)$$

where the convergence of the series follows by the square integrability of X. The operator $\sigma^{\frac{1}{4}} X \sigma^{\frac{1}{4}}$ is Hilbert Schmidt since

$$\mathrm{Tr}(\sigma^{\frac{1}{4}} X \sigma^{\frac{1}{4}})^* (\sigma^{\frac{1}{4}} X \sigma^{\frac{1}{4}}) = \sum_{j=1}^{\infty} \|\sigma^{\frac{1}{4}} X \sigma^{\frac{1}{4}} e_j\|^2_{\mathcal{H}} = \sum_{j=1}^{\infty} \lambda_j^{\frac{1}{2}} \|\sigma^{\frac{1}{4}} X e_j\|^2_{\mathcal{H}} = \sum_{j,k=1}^{\infty} \sqrt{\lambda_j \lambda_k} \, |\langle e_j, X e_k \rangle|^2 < \infty .$$

The converse statement is obvious.

$\square$

Similarly, we define the weighted norms $\|\cdot\|_{\mathbb{L}_p(\sigma)}$ on $\mathcal{B}(\mathcal{H})$ for all $p \geq 1$ as follows

$$\|X\|_{\mathbb{L}_p(\sigma)} := \mathrm{Tr}\left(\left| \sigma^{\frac{1}{2p}} X \sigma^{\frac{1}{2p}} \right|^p \right)^{\frac{1}{p}} .$$

and the weighted $\mathbb{L}_p(\sigma)$ spaces as the closures of $\mathcal{B}(\mathcal{H})$ in the corresponding $\mathbb{L}_p$ norm. The $\mathbb{L}_p(\sigma)$ norms are connected to the usual Schatten norms as follows: first, define the map:

$$\Gamma_\sigma : \quad X \in \mathcal{B}(\mathcal{H}) \quad \mapsto \quad \sigma^{\frac{1}{2}} X \sigma^{\frac{1}{2}} , \qquad \text{so that} \qquad \Gamma_\sigma^{\frac{1}{p}} : \quad X \in \mathcal{B}(\mathcal{H}) \quad \mapsto \quad \sigma^{\frac{1}{2p}} X \sigma^{\frac{1}{2p}} . \qquad (1.10)$$

Then one has $\|X\|_{\mathbb{L}_p(\sigma)} = \|\Gamma_\sigma^{\frac{1}{p}}(X)\|_p$.

Theorem 1.1.3. *For any* $p \geq 1$, *the map* $\Gamma_\sigma^{\frac{1}{p}} : \mathcal{B}(\mathcal{H}) \to \mathcal{T}_p(\mathcal{H})$ *extends to an isometry from* $\mathbb{L}_p(\sigma)$ *to* $\mathcal{T}_p(\mathcal{H})$.

Proof. The proof is rather standard, but given for sake of completeness. By density, for any $X \in \mathbb{L}_p(\sigma)$, there exists a sequence $\{X_n\}_{n \in \mathbb{N}}$ of bounded operators in $\mathcal{B}(\mathcal{H})$ such that $\|X_n - X\|_{\mathbb{L}_p(\sigma)} \to 0$ as $n \to \infty$.

Since $\{X_n\}$ is a Cauchy sequence in $\mathbb{L}_p(\sigma)$, this implies that $\{\Gamma_\sigma^{\frac{1}{p}}(X_n)\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{T}_p(\mathcal{H})$ and hence converges by completeness of the Schatten spaces. Denoting by T its limit, the reverse triangle inequality insures that

$$\left| \|\Gamma_\sigma^{\frac{1}{p}}(X_n)\|_p - \|T\|_p \right| \le \|\Gamma_\sigma^{\frac{1}{p}}(X_n) - T\|_p , \tag{1.11}$$

so that $\|X\|_{\mathbb{L}_p(\sigma)} = \lim_{n\to\infty} \|X_n\|_{\mathbb{L}_p(\sigma)} = \lim_{n\to\infty} \|\Gamma_\sigma^{\frac{1}{p}}(X_n)\|_p = \|T\|_p$. Assuming that $\{X_n'\}_{n\in\mathbb{N}}$ is another sequence converging to X, and denoting by T' the corresponding limit of $\Gamma_\sigma^{\frac{1}{p}}(X_n')$ in $\mathcal{T}_p(\mathcal{H})$, we get by continuity of the norm that $T' = T$, since:

$$\|T' - T\|_p \le \|T' - \Gamma_\sigma^{\frac{1}{p}}(X_n')\|_p + \|\Gamma_\sigma^{\frac{1}{p}}(X_n' - X_n)\|_p + \|\Gamma_\sigma^{\frac{1}{p}}(X_n) - T\|_p .$$

The terms on the above right hand side all converge to 0 by definition. Hence, one can defined without ambiguity $T = \Gamma_\sigma(X)$, and the result follows. $\qquad\square$

Next, we introduce the cones $\mathbb{L}_p^+(\sigma)$ of positive $\mathbb{L}_p(\sigma)$ operators on $\mathcal{H}$. First, defined for $p \ge 1$ the p-modulus map $|.|_p$ as follows: for any $X \in \mathcal{B}(\mathcal{H})$:

$$|X|_p := \sigma^{-\frac{1}{2p}} \left| \sigma^{\frac{1}{2p}} X \sigma^{\frac{1}{2p}} \right| \sigma^{-\frac{1}{2p}} .$$

It was shown p. 249 of [Olkiewicz and Zegarlinski, 1999] that the p-modulus map can be extended by continuity to an isomorphism on the entire $\mathbb{L}_p(\sigma)$:

$$\| |X|_p \|_{\mathbb{L}_p(\sigma)} = \|X\|_{\mathbb{L}_p(\sigma)} .$$

This allows us to define the following family of positive cones:

$$\mathbb{L}_p^+(\sigma) := \left\{ X \in \mathbb{L}_p(\sigma) : |X|_p = X \right\} .$$

One can finally define a useful family of bijective maps $\mathrm{I}_{q,p}$ between $\mathbb{L}_p(\sigma)$ and $\mathbb{L}_q(\sigma)$ for each $p, q \ge 1$, defined for all $X \in \mathcal{B}(\mathcal{H})$ by:

$$\mathrm{I}_{q,p}(X) := \Gamma_\sigma^{-\frac{1}{q}} \left| \Gamma_\sigma^{\frac{1}{p}}(X)^{\frac{p}{q}} \right| = \sigma^{-\frac{1}{2q}} \left| \sigma^{\frac{1}{2p}} X \sigma^{\frac{1}{2p}} \right|^{\frac{p}{q}} \sigma^{-\frac{1}{2q}} . \tag{1.12}$$

Weighted $\mathbb{L}_p(\sigma)$ norms enjoy the following properties [Ball et al., 1994, Olkiewicz and Zegarlinski, 1999]:

Proposition 1.1.4. *Let $1 \le p \le \infty$. Then*

(i) *Minkowski inequality: For any $X, Y \in \mathbb{L}_p(\sigma)$,*

$$\|X + Y\|_{\mathbb{L}_p(\sigma)} \le \|X\|_{\mathbb{L}_p(\sigma)} + \|Y\|_{\mathbb{L}_p(\sigma)}.$$

(ii) *Hölder's inequality: Let $\hat{p}$ denote the Hölder conjugate of p, i.e. $p^{-1} + \hat{p}^{-1} = 1$. Then, for any $X \in \mathbb{L}_p(\sigma)$ and $Y \in \mathbb{L}_{\hat{p}}(\sigma)$, $\Gamma_\sigma^{\frac{1}{p}}(X)\Gamma_\sigma^{\frac{1}{2\hat{p}}}(Y) \in \mathcal{T}_1(\mathcal{H})$ and*

$$\langle X, Y \rangle_\sigma := \mathrm{Tr}\left(\Gamma_\sigma^{\frac{1}{p}}(X) \Gamma_\sigma^{\frac{1}{\hat{p}}}(Y) \right) \le \|X\|_{\mathbb{L}_p(\sigma)} \|Y\|_{\mathbb{L}_{\hat{p}}(\sigma)}.$$

(iii) <u>*Duality of weighted norms:*</u> *For any* $X \in \mathbb{L}_p(\sigma)$,

$$\|X\|_{\mathbb{L}_p(\sigma)} = \sup_{Y \in \mathbb{L}_{\hat{p}}(\sigma)} \frac{|\langle Y, X \rangle_\sigma|}{\|Y\|_{\mathbb{L}_{\hat{p}}(\sigma)}}.$$

(iv) *For any* $q \geq 1$ *and any* $X \in \mathbb{L}_q^+(\sigma)$, $\mathrm{I}_{\hat{q},q}(X)$ *is the only operator* $Y \in \mathbb{L}_{\hat{q}}^+(\sigma)$ *such that*

$$\|X\|_{\mathbb{L}_q(\sigma)}^q = \langle Y, X \rangle_\sigma. \tag{1.13}$$

(v) <u>*Uniform convexity*</u> *for any* $p \geq 1$, *the space* $\mathbb{L}_p(\sigma)$ *is uniformly convex. That is, for any* $\varepsilon > 0$, *there exists* $\delta > 0$ *such that for any two unit vectors* $X, Y \in \mathbb{L}_p(\sigma)$,

$$\|X - Y\|_{\mathbb{L}_p(\sigma)} \geq \varepsilon \qquad \Rightarrow \qquad \frac{1}{2}\|X + Y\| \leq 1 - \delta.$$

In particular, the uniform convexity of $\mathbb{L}_p(\sigma)$ spaces implies their *strict convexity*: for any $X, Y \in \mathbb{L}_p(\sigma) \backslash \{0\}$,

$$\|X + Y\|_{\mathbb{L}_p(\sigma)} = \|X\|_{\mathbb{L}_p(\sigma)} + \|Y\|_{\mathbb{L}_p(\sigma)} \qquad \Leftrightarrow \qquad \exists\, t > 0 : X = t\, Y.$$

By Hölder's inequality, one can easily find the following hierarchy of non-commutative weighted $\mathbb{L}_p$ norms: for any $1 \leq p \leq q \leq \infty$:

$$\|X\|_{\mathbb{L}_p(\sigma)} \leq \|X\|_{\mathbb{L}_q(\sigma)} \qquad \Rightarrow \qquad \mathcal{B}(\mathcal{H}) \subset \mathbb{L}_q(\sigma) \subset \mathbb{L}_p(\sigma) \subset \mathbb{L}_1(\sigma).$$

Weighted $\mathbb{L}_p(\sigma)$ norms also satisfy the following useful interpolation theorem as a consequence of Hadamard's three lines lemma:

Theorem 1.1.5. *Weighted* $\mathbb{L}_p$ *spaces form an* interpolation family *of Banach spaces* $(\mathbb{L}_p(\sigma))_{p \in [1,\infty]}$: *for any* $1 \leq p_0 \leq p_1 \leq \infty$, $0 < \theta < 1$, *and* $X \in \mathbb{L}_{p_0}(\sigma)$

$$\|X\|_{\mathbb{L}_{p_\theta}(\sigma)} \leq \|X\|_{\mathbb{L}_{p_0}(\sigma)}^{1-\theta} \|X\|_{\mathbb{L}_{p_1}(\sigma)}^{\theta}, \tag{1.14}$$

where p_θ *is defined through the following equation:*

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}. \tag{1.15}$$

The following properties can be found in [Olkiewicz and Zegarlinski, 1999]:

Proposition 1.1.6. *The maps* $\mathrm{I}_{q,p}$ *extend to continuous bijective maps from* $\mathbb{L}_q(\sigma)$ *to* $\mathbb{L}_p(\sigma)$ *and satisfy the following properties:*

(i) $\|\mathrm{I}_{q,p}(X)\|_{\mathbb{L}_q(\sigma)}^q = \|X\|_{\mathbb{L}_p(\sigma)}^p.$

(ii) $\mathrm{I}_{q,r} \circ \mathrm{I}_{r,p} = \mathrm{I}_{q,p}$ *and* $\mathrm{I}_{p,p} = \mathrm{id}_{\mathbb{L}_p(\sigma)}.$

(iii) *For any* $p \in [1, \infty)$, $\mathbb{L}_p^+(\sigma)$ *is a closed strict cone in* $\mathbb{L}_p(\sigma)$ *such that the cone of positive semidefinite bounded operators is dense in it. In the case* $p = \infty$, *this remains true in the weak* sense.*

Weighted amalgamated $\mathbb{L}_p$ norms   The quantum weighted $\mathbb{L}_p(\sigma)$ norms introduced in the last paragraph will prove useful in the analysis of the convergence properties of *primitive* QMS, that is, QMS converging towards the faithful state σ. In the Heisenberg picture, this is equivalent to asking the observables to converge to a multiple of the identity:

$$\mathcal{P}_t(X) \underset{t \to \infty}{\to} \operatorname{Tr}(\sigma X)\, \mathbb{1}_{\mathcal{H}}.$$

We will see in Chapter 6 that, at least under the assumption of existence of a faithful stationary state σ, a QMS more generally converges towards a C*-algebra $\mathcal{N}$ called its *decoherence-free subalgebra*. Chapters 8 and 9 are devoted to the study of the speed of convergence of such evolutions. This analysis will in particular rely on the study of the contractivity properties of the QMS in question under a family of weighted amalgamated $\mathbb{L}_p$ norms related to the algebra $\mathcal{N}$. These so-called *amalgamated $\mathbb{L}_p$ norms* were introduced in [Junge and Parcet, 2010] in the framework of operator space theory.

Here, we assume that $\mathcal{H}$ is finite dimensional. In Chapter 6, we show that the conditional expectation $E_{\mathcal{N}}$ defined in Equation (0.11) is the orthogonal projection onto $\mathcal{N}$ for the KMS inner product (cf. [Bardet, 2017]): for all $X, Y \in \mathcal{B}(\mathcal{H})$:

$$\langle X, E_{\mathcal{N}}[Y] \rangle_{\sigma_{\mathrm{Tr}}} = \langle E_{\mathcal{N}}[X], Y \rangle_{\sigma_{\mathrm{Tr}}} = \langle E_{\mathcal{N}}[X], E_{\mathcal{N}}[Y] \rangle_{\sigma_{\mathrm{Tr}}}, \tag{1.16}$$

where $\sigma_{\mathrm{Tr}} := E_{\mathcal{N}*}\left(\frac{\mathbb{1}_{\mathcal{H}}}{d_{\mathcal{H}}}\right)$. This in particular implies that:

$$\sigma_{\mathrm{Tr}}^{\frac{1}{2}} E_{\mathcal{N}}[X] \sigma_{\mathrm{Tr}}^{\frac{1}{2}} = E_{\mathcal{N}*}(\sigma_{\mathrm{Tr}}^{\frac{1}{2}} X \sigma_{\mathrm{Tr}}^{\frac{1}{2}}). \tag{1.17}$$

For $1 \le q \le p \le +\infty$, and $\frac{1}{r} = \frac{1}{q} - \frac{1}{p}$, define[2]

$$\|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} := \inf_{\substack{A, B \in \mathcal{N}, Y \in \mathcal{B}(\mathcal{H}) \\ X = AYB}} \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|Y\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})}, \tag{1.18}$$

$$\|Y\|_{\mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}}))} := \sup_{A, B \in \mathcal{N}} \frac{\|A Y B\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}{\|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})}}. \tag{1.19}$$

For any $1 \le q, p \le +\infty$, we denote the space $\mathcal{B}(\mathcal{H})$ endowed with the norms $\|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))}$ by $\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))$. In the case when $\sigma_{\mathrm{Tr}} = d_{\mathcal{H}}^{-1} \mathbb{1}_{\mathcal{H}}$, these norms reduce to the amalgamated $\mathbb{L}_p$ norms defined through Equation (1.5).

Remark 1.1.7. Classically, these norms reduce to the following: let $(\Omega, \mathcal{F}, \mu)$ a probability space, and let $\mathcal{G}$ a sub-sigma algebra of $\mathcal{F}$. Then for any $\mathcal{F}$-measurable random variable $X : \Omega \to \mathbb{R}$, and any $1 \le q, p \le \infty$:

$$\|X\|_{\mathbb{L}_q(\mathbb{L}_\infty(\Omega, \mathcal{G}, \mu), \mathbb{L}_p(\mu))} = \|X\|_{\mathbb{L}_q^p(\mathbb{L}_\infty(\Omega, \mathcal{G}, \mu) \subset \mathbb{L}_\infty(\Omega, \mathcal{F}, \mu))} = \mathbb{E}\left[\left(\mathbb{E}[|X|^p \,|\, \mathcal{G}]\right)^{\frac{q}{p}}\right]^{\frac{1}{q}} = \left\|\mathbb{E}[|X|^p \,|\, \mathcal{G}]^{\frac{1}{p}}\right\|_{\mathbb{L}_q(\mu)}.$$

This follows easily from applying Hölder's inequality. One reason behind this seemingly complicated generalization is because it can be shown to define a norm, as opposed to the arguably more natural form $\|E_{\mathcal{N}}[|X|^p]^{\frac{1}{p}}\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}$.

In the next proposition, we gathered some basic properties of these norms.

Proposition 1.1.8. *Let $1 \le q, p \le +\infty$ together with their Hölder conjugates $\hat{q}, \hat{p}$, i.e. such that $\frac{1}{p} + \frac{1}{\hat{p}} = 1$, and $\frac{1}{q} + \frac{1}{\hat{q}} = 1$. Moreover, let $\mathcal{N}$ be a subalgebra of $\mathcal{B}(\mathcal{H})$ with corresponding conditional*

---

[2]In [Junge and Parcet, 2010], these norms were defined with respect to any state σ on $\mathcal{H}$ with respect to which $E_{\mathcal{N}}$ is a conditional expectation. However, we will see that the choice σ_{Tr} is very convenient in Chapter 8.

expectation $E_{\mathcal{N}}$ and $\sigma_{\mathrm{Tr}} \coloneqq E_{\mathcal{N}*}(d_{\mathcal{H}}^{-1}\mathbb{1})$. Then the following holds:

(i) *Hölder's inequality: For any* $X \in \mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}}))$, $Y \in \mathbb{L}_{\hat{p}}(\mathcal{N}, \mathbb{L}_{\hat{q}}(\sigma_{\mathrm{Tr}}))$,

$$|\langle X, Y \rangle_{\sigma_{\mathrm{Tr}}}| \leq \|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} \|Y\|_{\mathbb{L}_{\hat{q}}(\mathcal{N}, \mathbb{L}_{\hat{p}}(\sigma_{\mathrm{Tr}}))}.$$

(ii) *Duality: For any* $X \in \mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))$,

$$\|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} = \sup\left\{|\langle X, Y \rangle_{\sigma_{\mathrm{Tr}}}| : \ \|Y\|_{\mathbb{L}_{\hat{q}}(\mathcal{N}, \mathbb{L}_{\hat{p}}(\sigma_{\mathrm{Tr}}))} = 1\right\}.$$

(iii) *Relation with* $\mathbb{L}_p(\sigma_{\mathrm{Tr}})$ *norms: if* $q \leq p$, *then for any* $X \in \mathbb{L}_p(\sigma_{\mathrm{Tr}})$,

$$\|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} \leq \|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} \leq \|X\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})}, \tag{1.20}$$

$$\|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} \leq \|X\|_{\mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}}))} \leq \|X\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})}. \tag{1.21}$$

and, in both cases, equality holds for all X *if* $p = q$. *This last statement is usually referred to as Fubini's Theorem.*

(iv) *The hierarchy of norms: for* $1 \leq q_1 \leq q_2, p_1 \leq p_2 \leq +\infty$, *and any* $X \in \mathcal{B}(\mathcal{H})$,

$$\|X\|_{\mathbb{L}_{q_1}(\mathcal{N}, \mathbb{L}_{p_1}(\sigma_{\mathrm{Tr}}))} \leq \|X\|_{\mathbb{L}_{q_2}(\mathcal{N}, \mathbb{L}_{p_2}(\sigma_{\mathrm{Tr}}))}.$$

(v) *When* $1 \leq q \leq p \leq +\infty$, *the* sup *on the right hand side of Equation* (1.19) *may be restricted to the set of positive semidefinite operators* $A, B \geq 0$. *Furthermore, for all positive semidefinite* X,

$$\|X\|_{\mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}}))} = \sup_{A \in \mathcal{N}(\mathcal{P}), \ A > 0, \ \|A\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1} \left\|A^{1/2r} X A^{1/2r}\right\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}, \tag{1.22}$$

(vi) *Similarly, the* inf *on the right hand side of Equation* (1.18) *may be restricted to the set of positive semidefinite operators* $A, B \geq 0$. *Furthermore, for all positive semidefinite* X,

$$\|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} = \inf_{A \in \mathcal{N}(\mathcal{P}), \ A > 0, \ \|A\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1} \left\|A^{-1/2r} X A^{-1/2r}\right\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})}. \tag{1.23}$$

(vii) $\left\{\|.\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))}\right\}_{1 \leq q \leq p \leq +\infty}$ *defines a family of interpolating norms (cf. Section 1.1.2).*

(viii) *For all* $1 \leq q \leq p \leq +\infty$, $\|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} = \|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}$ *whenever* $X \in \mathcal{N}$.

Proof.

(i) Hölder's inequality follows directly from Hölder's inequality for the $\mathbb{L}_p(\sigma_{\mathrm{Tr}})$ norms (cf. Proposition 1.1.4): without loss of generality, assume that $p \leq q$, so that $\hat{q} \leq \hat{p}$. Consider any decomposition of Y of the form $Y = AZB$, with $A, B \in \mathcal{N}$ and $Z \in \mathcal{B}(\mathcal{H})$. Then,

$$\begin{aligned}
|\langle X, Y \rangle_{\sigma_{\mathrm{Tr}}}| = |\langle X, AZB \rangle_{\sigma_{\mathrm{Tr}}}| &= |\langle A^* X B^*, Z \rangle_{\sigma_{\mathrm{Tr}}}| \\
&\leq \|A^* X B^*\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})} \|Z\|_{\mathbb{L}_{\hat{p}}(\sigma_{\mathrm{Tr}})} \\
&\leq \|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|Z\|_{\mathbb{L}_{\hat{p}}(\sigma_{\mathrm{Tr}})}.
\end{aligned}$$

We conclude by taking the infimum over the operators A, B, Z.

(ii) Assume without loss of generality that $1 \leq q \leq p \leq +\infty$. Then

$$
\begin{aligned}
\|X\|_{\mathbb{L}_p(\mathcal{N},\mathbb{L}_q(\sigma_{\mathrm{Tr}}))} &= \sup_{A,B\in\mathcal{N}} \{\|AXB\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} : \ \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})}\|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \leq 1\} \\
&= \sup_{A,B\in\mathcal{N},\, Z\in\mathcal{B}(\mathcal{H})} \{|\langle AXB, Z\rangle_{\sigma_{\mathrm{Tr}}}| : \ \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})}\|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \leq 1, \|Z\|_{\mathbb{L}_{\hat{q}}(\sigma_{\mathrm{Tr}})} \leq 1\} \\
&\leq \sup_{A,B\in\mathcal{N},\, Z\in\mathcal{B}(\mathcal{H})} \{|\langle X, A^*ZB^*\rangle_{\sigma_{\mathrm{Tr}}}| : \ \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})}\|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})}\|Z\|_{\mathbb{L}_{\hat{q}}(\sigma_{\mathrm{Tr}})} \leq 1\} \\
&= \sup_{A,B\in\mathcal{N},\, W,Z\in\mathcal{B}(\mathcal{H})} \{|\langle X, W\rangle_{\sigma_{\mathrm{Tr}}}| : \ W = A^*ZB^*,\ \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})}\|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})}\|Z\|_{\mathbb{L}_{\hat{q}}(\sigma_{\mathrm{Tr}})} \leq 1\} \\
&\leq \sup_{W\in\mathcal{B}(\mathcal{H})} \{|\langle X, W\rangle_{\sigma_{\mathrm{Tr}}}| : \ \|W\|_{\mathbb{L}_{\hat{p}}(\mathcal{N},\mathbb{L}_{\hat{q}}(\sigma_{\mathrm{Tr}}))} \leq 1\},
\end{aligned}
$$

where in the second line, we used the duality of $\mathbb{L}_p(\sigma_{\mathrm{Tr}})$ norms, in the third line we used that for $A, B \in \mathcal{N}$, $[A, \sigma_{\mathrm{Tr}}] = [B, \sigma_{\mathrm{Tr}}] = 0$, and in the last line we used that $\frac{1}{r} = \frac{1}{\hat{p}} - \frac{1}{\hat{q}}$. Using Hölder's inequality (i), the condition $\|W\|_{\mathbb{L}_{\hat{p}}(\mathcal{N},\mathbb{L}_{\hat{q}}(\sigma_{\mathrm{Tr}}))} \leq 1$ implies

$$
|\langle X, W\rangle_{\sigma_{\mathrm{Tr}}}| \leq \|X\|_{\mathbb{L}_p(\mathcal{N},\mathbb{L}_q(\sigma_{\mathrm{Tr}}))}\|W\|_{\mathbb{L}_{\hat{p}}(\mathcal{N},\mathbb{L}_{\hat{q}}(\sigma_{\mathrm{Tr}}))} \leq \|X\|_{(p,q),\mathcal{N}}.
$$

Therefore, the supremum is attained. This shows that the Banach space $\mathbb{L}_{\hat{p}}(\mathcal{N},\mathbb{L}_{\hat{q}}(\sigma_{\mathrm{Tr}}))$ is the dual of $\mathbb{L}_p(\mathcal{N},\mathbb{L}_q(\sigma_{\mathrm{Tr}}))$. As these spaces are finite dimensional, the converse holds.

(iii) The second inequality in (1.20) and the first inequality in (1.21) are obvious by definition. The second inequality in (1.21) and the first inequality in (1.20) are proved by a use of Hölder's inequality for $\mathbb{L}_p(\sigma_{\mathrm{Tr}})$ norms.

(iv) By convexity of the inverse function, $\frac{1}{r_1} \equiv \frac{1}{q_1} - \frac{1}{p_1} \geq \frac{1}{q_2} - \frac{1}{p_2} \equiv \frac{1}{r_2}$, so that

$$
\begin{aligned}
\|X\|_{\mathbb{L}_{q_1}(\mathcal{N},\mathbb{L}_{p_1}(\sigma_{\mathrm{Tr}}))} &= \inf_{A,B\in\mathcal{N},\, Y\in\mathcal{B}(\mathcal{H}),\, X=AYB} \|A\|_{\mathbb{L}_{2r_1}(\sigma_{\mathrm{Tr}})} \|B\|_{\mathbb{L}_{2r_1}(\sigma_{\mathrm{Tr}})} \|Y\|_{\mathbb{L}_{p_1}(\sigma_{\mathrm{Tr}})} \\
&\leq \inf_{A,B\in\mathcal{N},\, Y\in\mathcal{B}(\mathcal{H}),\, X=AYB} \|A\|_{\mathbb{L}_{2r_2}(\sigma_{\mathrm{Tr}})} \|B\|_{\mathbb{L}_{2r_2}(\sigma_{\mathrm{Tr}})} \|Y\|_{\mathbb{L}_{p_2}(\sigma_{\mathrm{Tr}})} \\
&= \|X\|_{\mathbb{L}_{q_2}(\mathcal{N},\mathbb{L}_{p_2}(\sigma_{\mathrm{Tr}}))},
\end{aligned}
$$

where in the second line we used the hierarchy of the $\|.\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})}$ norms: for $p \leq p'$, $\|X\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})} \leq \|X\|_{\mathbb{L}_{p'}(\sigma_{\mathrm{Tr}})}$.

(v) The first claim follows directly from invariance of $\mathcal{N}$ under $A \mapsto |A| \equiv \sqrt{A^*A}$, polar decomposition, as well as invariance of the $\mathbb{L}_p(\sigma_{\mathrm{Tr}})$ norms under unitary transformations $U \in \mathcal{N}$. Assume now that $X \geq 0$. Then, by Hölder's inequality for the Schatten norms,

$$
\|AXB\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} \leq \sqrt{\|AXA^*\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}\|BXB^*\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}} \leq \max\left\{\|AXA^*\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}, \|BXB^*\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}\right\},
$$

where we also used that $\Gamma_{\sigma_{\mathrm{Tr}}}(AXB^*) = A\Gamma_{\sigma_{\mathrm{Tr}}}(X)B^*$. Moreover, equality holds when $A = B$. Since positive definite operators are dense in the set of positive semidefinite operators, we conclude that for all positive semidefinite X,

$$
\|X\|_{\mathbb{L}_p(\mathcal{N},\mathbb{L}_q(\sigma_{\mathrm{Tr}}))} = \sup_{A\in\mathcal{N},\ A>0,\ \|A\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}=1} \left\|A^{1/2r}\, X\, A^{1/2r}\right\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}.
$$

(vi) This properties is more difficult to prove than the previous one. We refer to point (iv) of Proposition 4.1.5 in [Xu, 2007].

(vii) This is proved in [Junge and Parcet, 2010].

(viii) From the first inequality of Equation (1.20), we only need to find $A, B \in \mathcal{N}$ and $Y \in \mathcal{B}(\mathcal{H})$ such that $X = AYB$, and $\|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} = \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|Y\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})}$. This works by taking $A = B = X^{\frac{q}{2r}}$ and $Y = X^{\frac{q}{p}}$. Indeed, in this case,

$$\|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} = \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} = \left(\mathrm{Tr}(\sigma_{\mathrm{Tr}} X^q)\right)^{\frac{1}{2r}} = \|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{\frac{q}{2r}}, \qquad \|Y\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})} = \mathrm{Tr}(\sigma_{\mathrm{Tr}} X^q)^{\frac{1}{p}} = \|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{\frac{q}{p}},$$

and the claim follows from the fact that $\frac{1}{r} + \frac{1}{p} = \frac{1}{q}$.

Remark 1.1.9. We will see in Chapter 8 that the weighted amalgamated $\mathbb{L}_p$ norms are uniformly convex only in the case when $\mathcal{N}$ is trivial, that is in the special case when they coincide with the non-commutative weighted $\mathbb{L}_p$ norms introduced in the previous paragraph.

1.1.2. Superoperator norms

Here, we shift our attention to the problem of distinguishing quantum channels. The norms of last section provide the sets of states and observables with different Banach space structures. Next, given a bounded linear map $\Phi : \mathcal{B}_1 \to \mathcal{B}_2$, between any two such Banach spaces $(\mathcal{B}_1, \|.\|_{\mathcal{B}_1})$ and $(\mathcal{B}_2, \|.\|_{\mathcal{B}_2})$, the *operator norm* of Φ is defined as follows

$$\|\Phi : \mathcal{B}_1 \to \mathcal{B}_2\| := \sup_{X \in \mathcal{B}_1 \backslash \{0\}} \frac{\|\Phi(X)\|_{\mathcal{B}_2}}{\|X\|_{\mathcal{B}_1}}.$$

Superoperator norms induced by Schatten norms   In the Schrödinger picture, one is mainly interested in the distinguishability of CPTP maps. There, $\mathcal{B}_1$ and $\mathcal{B}_2$ are both taken to be the Banach spaces $\mathcal{T}_1(\mathcal{H})$ of trace-class operators on a given separable Hilbert space $\mathcal{H}$. Then, it was shown in [Watrous, 2005] that in the case of a completely positive map $\mathcal{P}$

$$\|\mathcal{P} : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_1(\mathcal{H})\| = \sup_{\rho \in \mathcal{D}_+(\mathcal{H})} \|\mathcal{P}(\rho)\|_1.$$

We saw in Section 1.1.1 that the trace norm is related to the minimum total error made in the task of state discrimination. This operational interpretation can be directly translated to the discrimination of quantum channels: Assume that Alice sends a state ρ to Bob over a noisy channels $\mathcal{P}_i \in \{\mathcal{P}_1, \mathcal{P}_2\}$. Bob's task is then to infer which channel the initial state ρ went through by performing a two-outcome POVM $\mathbf{M} = \{M_1, M_2\}$ where $M_1 = \mathbb{1} - M_2$. Similarly to the usual hypothesis testing problem on states, two errors are associated to this test: $\mathrm{Tr}(M_i \mathcal{P}_j(\rho))$ represents the error made when inferring that the channel $\mathcal{P}_i$ was used when it actually was $\mathcal{P}_j$. Assume further that channel $\mathcal{P}_1$, resp. $\mathcal{P}_2$, has probability $p \in (0,1)$, resp. $(1-p)$, of being used. Then, for a fixed initial state ρ, the minimum total error probability is given by

$$P_{\mathrm{err}}(\mathcal{P}_1, \mathcal{P}_2, p, \rho) = \min_{\mathbf{M}} p \, \mathrm{Tr}(M_2 \, \mathcal{P}_1(\rho)) + (1-p) \, \mathrm{Tr}(M_1 \, \mathcal{P}_2(\rho)) = \frac{1}{2} \left(1 - \|p\,\mathcal{P}_1(\rho) + (1-p)\,\mathcal{P}_2(\rho)\|_1\right).$$

$$(1.24)$$

After optimization over the input state ρ, Equation (1.24) leads to the following minimum error probability for the task of discrimination between the two channels $\mathcal{P}_1$ and $\mathcal{P}_2$:

$$P_{\mathrm{err}}(\mathcal{P}_1, \mathcal{P}_2, p) = \min_{\rho} P_{\mathrm{err}}(\mathcal{P}_1, \mathcal{P}_2, p, \rho) = \frac{1}{2}\left(1 - \|p\mathcal{P}_1 + (1-p)\mathcal{P}_2 : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_1(\mathcal{H})\|\right). \qquad (1.25)$$

More generally, one can define the superoperator norm induced by Schatten norms: for $1 \leq p, q \leq \infty$ and a linear map $\Phi : \mathcal{T}_q(\mathcal{H}) \to \mathcal{T}_p(\mathcal{H})$, we define

$$\|\Phi : \mathcal{T}_q(\mathcal{H}) \to \mathcal{T}_p(\mathcal{H})\| = \sup_{X \in \mathcal{T}_q(\mathcal{H}) \backslash \{0\}} \frac{\|\Phi(X)\|_{\mathcal{T}_p(\mathcal{H})}}{\|X\|_{\mathcal{T}_q(\mathcal{H})}}.$$

Once again, the supremum can be restricted to the set $\mathcal{T}_q^+(\mathcal{H})$ of positive semidefinite operators in $\mathcal{T}_q(\mathcal{H})$.

These superoperator norms are related to the ones for maps between weighted $\mathbb{L}_p(\sigma)$ spaces, for a given faithful normal state represented by the density operator $\sigma \in \mathcal{D}_+(\mathcal{H})$: given $\Phi : \mathbb{L}_q(\sigma) \to \mathbb{L}_p(\sigma)$:

$$
\begin{aligned}
\|\Phi : \mathbb{L}_q(\sigma) \to \mathbb{L}_p(\sigma)\| &:= \sup_{X \in \mathbb{L}_q(\sigma) \backslash \{0\}} \frac{\|\Phi(X)\|_{\mathbb{L}_p(\sigma)}}{\|X\|_{\mathbb{L}_q(\sigma)}} \\
&= \sup_{Y \in \mathcal{T}_q(\mathcal{H}) \backslash \{0\}} \frac{\|\Gamma_\sigma^{\frac{1}{p}} \circ \Phi \circ \Gamma_\sigma^{-\frac{1}{q}}(Y)\|_{\mathcal{T}_p(\mathcal{H})}}{\|Y\|_{\mathcal{T}_q(\mathcal{H})}} \\
&= \|\Gamma_\sigma^{\frac{1}{p}} \circ \Phi \circ \Gamma_\sigma^{-\frac{1}{q}} : \mathcal{T}_q(\mathcal{H}) \to \mathcal{T}_p(\mathcal{H})\|, \quad (1.26)
\end{aligned}
$$

where the second identity follows from Theorem 1.1.3.

Complex Interpolation   We saw in Section 1.1.1 that weighted $\mathbb{L}_p$ norms as well as amalgamated $\mathbb{L}_p$ space defined a family of complex interpolating spaces. Here, we briefly review this notion which we will use in different situations throughout this thesis: Two Banach spaces $\mathcal{X}_0$ and $\mathcal{X}_1$ are said to be compatible if there exists a Hausdorff topological vector space $\mathcal{X}$ such that $\mathcal{X}_0, \mathcal{X}_1 \subset \mathcal{X}$. This can always be done by considering the sum space $\mathcal{X}_0 + \mathcal{X}_1 := \{\psi_0 + \psi_1 : \psi_0 \in \mathcal{X}_0, \psi_1 \in \mathcal{X}_1\}$, which is a Banach space equipped with the norm $\|\psi\|_{\mathcal{X}_0 + \mathcal{X}_1} := \inf_{\psi = \psi_0 + \psi_1} \|\psi_0\|_{\mathcal{X}_0} + \|\psi_1\|_{\mathcal{X}_1}$. Next, we denote the strip $S := \{z \in \mathbb{C} : 0 \leq \operatorname{Re} z \leq 1\}$ and let S_0 denote its open interior. We denote by $\mathcal{F}(\mathcal{X}_0, \mathcal{X}_1)$ the space of all functions $f : S \to \mathcal{X}_0 + \mathcal{X}_1$, which are bounded and continuous on S and analytic on S_0, for which $\{f(it) : t \in \mathbb{R}\} \subset \mathcal{X}_0$ and $\{f(1 + it) : t \in \mathbb{R}\} \subset \mathcal{X}_1$. $\mathcal{F}(\mathcal{X}_0, \mathcal{X}_1)$ is still a Banach space equipped with the norm $\|f\|_{\mathcal{F}} := \max\{\sup_t \|f(it)\|_{\mathcal{X}_0}, \sup_t \|f(1 + it)\|_{\mathcal{X}_1}\}$. Next, we denote the *complex interpolation space* $(\mathcal{X}_0, \mathcal{X}_1)_\theta$, $0 < \theta < 1$, as the quotient space of $\mathcal{F}(\mathcal{X}_0, \mathcal{X}_1)$ given as

$$(\mathcal{X}_0, \mathcal{X}_1)_\theta = \{\psi \in \mathcal{X}_0 + \mathcal{X}_1 : \psi = f(\theta), f \in \mathcal{F}(\mathcal{X}_0, \mathcal{X}_1)\}$$

equipped with the quotient norm

$$\|\psi\|_\theta := \inf \{\|f\|_{\mathcal{F}}; f(\theta) = \psi\}.$$

Noncommutative $\mathbb{L}_p$ spaces form a complex interpolation family of Banach spaces: $(\mathbb{L}_\infty(\sigma), \mathbb{L}_1(\sigma))_{\frac{1}{p}} = \mathbb{L}_p(\sigma)$. This extends to the amalgamated norms as follows: for any $1 \leq p, q \leq \infty$, $\mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\operatorname{Tr}})) = (\mathbb{L}_\infty(\mathcal{N}, \mathbb{L}_q(\sigma_{\operatorname{Tr}})), \mathbb{L}_1(\mathcal{N}, \mathbb{L}_q(\sigma_{\operatorname{Tr}}))_{\frac{1}{p}}$ and $\mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\operatorname{Tr}})) = (\mathbb{L}_p(\mathcal{N}, \mathbb{L}_\infty(\sigma_{\operatorname{Tr}})), \mathbb{L}_p(\mathcal{N}, \mathbb{L}_1(\sigma_{\operatorname{Tr}})))_{\frac{1}{q}}$.

We will need a generalization of the standard Riezs Thorin theorem, which allows the operator Φ itself to vary analytically. This is the well-known Stein interpolation theorem [Bergh and Löfström, 2012]:

Theorem 1.1.10. *Let $(\mathcal{X}_0, \mathcal{X}_1)$ and $(\mathcal{Y}_0, \mathcal{Y}_1)$ be two compatible couples of Banach spaces. Let $\Phi_z : z \in S \subset \mathcal{B}(\mathcal{X}_0 + \mathcal{X}_1, \mathcal{Y}_0 + \mathcal{Y}_1)$ be a bounded analytic family of maps such that*

$$\{\Phi_{it} : t \in \mathbb{R}\} \subset \mathcal{B}(\mathcal{X}_0, \mathcal{Y}_0), \quad \{\Phi_{1+it} : t \in \mathbb{R}\} \subset \mathcal{B}(\mathcal{X}_1, \mathcal{Y}_1).$$

Suppose that $\Lambda_0 = sup_t\{\|\Phi_{it} : \mathcal{X}_0 \to \mathcal{Y}_0\|\}$ and $\Lambda_1 = sup_t\{\|\Phi_{1+it} : \mathcal{X}_1 \to \mathcal{Y}_1\}$ are both finite, then for $0 < \theta < 1$, Φ_θ is a bounded linear map from $(\mathcal{X}_0, \mathcal{X}_1)_\theta$ to $(\mathcal{Y}_0, \mathcal{Y}_1)_\theta$ and

$$\|\Phi_\theta : (\mathcal{X}_0, \mathcal{X}_1)_\theta \to (\mathcal{Y}_0, \mathcal{Y}_1)_\theta\| \le \Lambda_0^{1-\theta} \Lambda_1^\theta .$$

Some norm estimates In this paragraph, we provide some important estimates on several norms of the identity map: consider a subalgebra $\mathcal{N}$ of $\mathcal{B}(\mathcal{H})$ for some finite dimensional Hilbert space $\mathcal{H}$ and let $E_\mathcal{N}$ be a conditional expectation from $\mathcal{B}(\mathcal{H})$ to $\mathcal{N}$. Once again, define σ_{Tr} as in Equation (0.10) and subsequently the norms $\|\cdot\|_{\mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}}))}$ as in Equations (1.18) and (1.19). First, for any $p \ge 2$:

$$\|\mathrm{id} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))\| \le \|\mathrm{id} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_p(\sigma_{\mathrm{Tr}})\| = \left\|\sigma_{\mathrm{Tr}}^{-1}\right\|_\infty^{\frac{1}{2}-\frac{1}{p}} . \tag{1.27}$$

Using the Riesz-Thorin interpolation theorem together with the fact that id is contractive for $\|\cdot\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}}))}$:

Lemma 1.1.11. *For any $2 \le p \le q \le \infty$,*

$$\|\mathrm{id} : \mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}})) \to \mathbb{L}_2(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}}))\| \le \left\|\sigma_{\mathrm{Tr}}^{-1}\right\|_\infty^{\frac{1}{p}-\frac{1}{q}} . \tag{1.28}$$

In general, the bound given by (1.27) can be very bad. In the bipartite scenario where $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, $\mathcal{N} = \mathcal{B}(\mathcal{H}_A) \otimes \mathbb{1}_{\mathcal{H}_B}$ and $\sigma_{\mathrm{Tr}} = \frac{\mathbb{1}_{\mathcal{H}_A}}{d_{\mathcal{H}_A}} \otimes \sigma$ for some full-rank density matrix σ, one can get the better bound:

Lemma 1.1.12. *For any $2 \le p \le \infty$,*

$$\left\|\mathrm{id}_{\mathcal{B}(\mathcal{H})} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))\right\| \le \left\|\mathrm{id}_{\mathcal{B}(\mathcal{H}_B)} : \mathbb{L}_2(\sigma) \to \mathbb{L}_p(\sigma)\right\|_{\mathrm{cb}} \le \left\|\sigma^{-1}\right\|_\infty^{\frac{1}{2}-\frac{1}{p}} . \tag{1.29}$$

In particular, the outer bound does not depend on $\mathcal{H}_A$.

Proof. The first inequality is obvious by definition of the weighted CB norms (see next paragraph). For the second inequality, it is enough to prove that for all $\mathcal{H}_A$,

$$\left\|\mathrm{id}_{\mathcal{B}(\mathcal{H})} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))\right\| \le \left\|\sigma^{-1}\right\|_\infty^{\frac{1}{2}-\frac{1}{p}} \tag{1.30}$$

Now, for any fixed $\mathcal{H}_A$ and $\frac{1}{r} = \frac{1}{2} - \frac{1}{p}$,

$$
\begin{aligned}
\left\|\mathrm{id}_{\mathcal{B}(\mathcal{H})}\right. & \left. : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))\right\| \\
&= \sup_{X \in \mathcal{B}(\mathcal{H})} \frac{\|X\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(d_{\mathcal{H}_A}^{-1} \mathbb{1}_{\mathcal{H}_A} \otimes \sigma))}}{\|X\|_{\mathbb{L}_2(\mathbb{1}_{\mathcal{H}_A} d_{\mathcal{H}_A}^{-1} \otimes \sigma)}} \\
&= \sup_{X \in \mathcal{B}(\mathcal{H})} \inf_{A \in \mathcal{B}(\mathcal{H}_A)} \frac{\|(A \otimes \mathbb{1}_{\mathcal{H}_B})^{-1} X (A \otimes \mathbb{1}_{\mathcal{H}_B})^{-1}\|_{\mathbb{L}_p(\mathbb{1}_{\mathcal{H}_A} d_{\mathcal{H}_A}^{-1} \otimes \sigma)} \|A\|_{\mathbb{L}_{2r}(\mathbb{1}_{\mathcal{H}_A} d_{\mathcal{H}_A}^{-1})}^2}{\|X\|_{\mathbb{L}_2(\mathbb{1}_{\mathcal{H}_A} d_{\mathcal{H}_A}^{-1} \otimes \sigma)}} \\
&= \sup_{X \in \mathcal{B}(\mathcal{H})} \inf_{A \in \mathcal{B}(\mathcal{H}_A)} \frac{\|(A \otimes \mathbb{1}_{\mathcal{H}_B})^{-1} X (A \otimes \mathbb{1}_{\mathcal{H}_B})^{-1}\|_{\mathbb{L}_p(\mathbb{1}_{\mathcal{H}_A} \otimes \sigma)} \|A\|_{2r}^2}{\|X\|_{\mathbb{L}_2(\mathbb{1}_{\mathcal{H}_A} \otimes \sigma)}} .
\end{aligned}
$$

Assuming $p = \infty$, the above right hand side is bounded by

$$
\sup_{X \in \mathcal{B}(\mathcal{H})} \frac{\|X\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_\infty(d_{\mathcal{H}_A}^{-1} \mathbb{1}_{\mathcal{H}_A} \otimes \sigma))}}{\|X\|_{\mathbb{L}_2(\mathbb{1}_{\mathcal{H}_A} d_{\mathcal{H}_A}^{-1} \otimes \sigma)}} = \left\|\sigma^{-1}\right\|_\infty^{-1/2} \sup_X \frac{1}{\|X\|_2} \inf_{A \in \mathcal{B}(\mathcal{H}_A)} \left\|(A \otimes \mathbb{1}_{\mathcal{H}_B})^{-1} X (A \otimes \mathbb{1}_{\mathcal{H}_B})^{-1}\right\|_\infty \|A\|_{2r}
$$

$$
= \left\|\sigma^{-1}\right\|_\infty^{-1/2} \sup_{X \in \mathcal{B}(\mathcal{H})} \frac{1}{\|X\|_2} \|X\|_{(2,\infty)}
$$

$$
= \left\|\sigma^{-1}\right\|_\infty^{-1/2} \left\|\operatorname{id} : \mathcal{T}_2(\mathcal{H}) \to \mathcal{B}(\mathcal{H})\right\|_{\mathrm{cb}}
$$

$$
\leq \left\|\sigma^{-1}\right\|_\infty^{-1/2},
$$

where $\|X\|_{(2,\infty)}$ denotes the (unnormalized) $(2, \infty)$ norm of Pisier [Pisier, 1993]. We conclude by interpolating for fixed $\mathcal{H}_A$ at the level of Equation (1.30), since $\left\|\operatorname{id}_{\mathcal{B}(\mathcal{H})} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\sigma_{\mathrm{Tr}})\right\| = 1$. $\qquad\square$

So far we only focused on the norm $\|\operatorname{id} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))\|$ for different value of p. In Section 8.6, however, we need another kind of estimate:

Proposition 1.1.13. *For all $1 \leq p \leq q$, we have*

$$
\|\operatorname{id} : \mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}})) \to \mathbb{L}_q(\sigma_{\mathrm{Tr}})\| = \left(\max_{i \in I} d_{\mathcal{H}_i}\right)^{\frac{1}{p} - \frac{1}{q}}, \tag{1.31}
$$

where the $d_{\mathcal{H}_i}$ are the dimensions of the spaces $\mathcal{H}_i$ taking place in the decomposition of $\mathcal{N}$ given by (0.9). For $p = 2$ and $q = \infty$, this yields

$$
\|\operatorname{id} : \mathbb{L}_2(\mathcal{N}, \mathbb{L}_\infty(\sigma_{\mathrm{Tr}})) \to \mathbb{L}_\infty(\sigma_{\mathrm{Tr}})\| \leq \max_{i \in I} \sqrt{d_{\mathcal{H}_i}}. \tag{1.32}
$$

Proof. Because of the two following trivial norm estimates

$$
\|\operatorname{id} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})\| \leq 1,
$$

$$
\|\operatorname{id} : \mathbb{L}_1(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_1(\sigma_{\mathrm{Tr}})\| \leq 1
$$

and by applying twice the Riesz-Thorin interpolation Theorem (one for the first parameter and then one for the second), it is enough to prove

$$
\|\operatorname{id} : \mathbb{L}_1(\mathcal{N}, \mathbb{L}_\infty(\sigma_{\mathrm{Tr}})) \to \mathbb{L}_\infty(\sigma_{\mathrm{Tr}})\| \leq \max_{i \in I} d_{\mathcal{H}_i}.
$$

But by duality, this is the same as

$$
\|\operatorname{id} : \mathbb{L}_1(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_\infty(\mathcal{N}, \mathbb{L}_1(\sigma_{\mathrm{Tr}}))\| \leq \max_{i \in I} d_{\mathcal{H}_i}.
$$

Let $X \in \mathcal{B}(\mathcal{H})$ be positive semi-definite and fix $\varepsilon > 0$. Then there exists a positive definite $A \in \mathcal{N}$ with $\|A\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1$ such that:

$$
\|X\|_{\mathbb{L}_\infty(\mathcal{N}, \mathbb{L}_1(\sigma_{\mathrm{Tr}}))} \leq \left\|A^{\frac{1}{2}} X A^{\frac{1}{2}}\right\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} + \varepsilon
$$

$$
= \operatorname{Tr}\left(A \sigma_{\mathrm{Tr}} X\right) + \varepsilon
$$

$$
\leq \|A\|_\infty \|X\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} + \varepsilon,
$$

73

where in the last line we use Hölder's inequality. Then we have

$$\|A\|_\infty = \sum_{i \in I} \|A_i\|_\infty$$

$$\leq \sum_{i \in I} d_{\mathcal{H}_i} \|A_i\|_{\mathbb{L}_1(\mathbb{1}_{\mathcal{H}_i} d_{\mathcal{H}_i}^{-1})}$$

$$\leq \max_{i \in I} d_{\mathcal{H}_i},$$

where in the last line we use that $\|A\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1$. This concludes the proof. $\qquad\square$

In Chapter 9, we will also be concerned with norms of other linear maps between amalgamated $\mathbb{L}_p$ spaces: given a finite von Neumann algebra $(\mathcal{M}, \tau)$ with normalized trace τ, and a von Neumann subalgebra $\mathcal{N}$ of $\mathcal{M}$, a map $\Phi : \mathbb{L}_p(\mathcal{M}) \to \mathbb{L}_q(\mathcal{M})$ is called a $\mathcal{N}$-*bimodule map* if, for any $A, B \in \mathcal{N}$ and any $X \in \mathcal{M}$:

$$\Phi(AXB) = A\Phi(X)B.$$

Given an $\mathcal{N}$-bimodule map Φ and $p \leq q$, the following was proved in Lemma 3.12 of [Gao et al., 2018b], generalizing an earlier statement for vector valued $\mathbb{L}_p$ norms (see Lemma 1.7 of [Pisier, 1993]): for any $s \geq 1$:

$$\|\Phi : \mathbb{L}_\infty^p(\mathcal{N} \subset \mathcal{M}) \to \mathbb{L}_\infty^q(\mathcal{N} \subset \mathcal{M})\| = \|\Phi : \mathbb{L}_s^p(\mathcal{N} \subset \mathcal{M}) \to \mathbb{L}_s^q(\mathcal{N} \subset \mathcal{M})\|. \qquad (1.33)$$

Tensorization, CB and diamond norms The norms defined in the previous paragraph were examples of operator norms of linear maps between Banach spaces. In the commutative case, all the norms defined above share the obvious, yet crucial, so-called *tensorization property*: given $1 \leq p, q \leq \infty$, two measure spaces $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$, and two maps $\Phi_1 : \mathbb{L}_q(\mu_1) \to \mathbb{L}_p(\mu_1)$, $\Phi_2 : \mathbb{L}_q(\mu_2) \to \mathbb{L}_p(\mu_2)$,

$$\|\Phi_1 \otimes \Phi_2 : \mathbb{L}_q(\mu_1 \otimes \mu_2) \to \mathbb{L}_p(\mu_1 \otimes \mu_2)\| = \|\Phi_1 : \mathbb{L}_q(\mu_1) \to \mathbb{L}_p(\mu_1)\| \, \|\Phi_2 : \mathbb{L}_q(\mu_2) \to \mathbb{L}_p(\mu_2)\|. \quad (1.34)$$

This fundamental property plays a key role in establishing speed of convergence of evolutions occurring on infinite dimensional classical systems starting from the study of evolutions defined on two-point sample spaces (see e.g. [Bakry, 1994]). It is also closely related to the notion of additivity of the capacity of classical channels. However, an analogue of Equation (1.34) for the non-commutative $\mathbb{L}_p$ norms defined above is known to fail in general: for example, in the case when $1 = q \leq p$, [Werner and Holevo, 2002] showed that the operator norm $\| . : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_p(\mathcal{H})\|$ does not tensorize when considering two copies of the following CPTP map, nowadays referred to as the *Werner-Holevo channel*:

$$\Phi(\rho) = \frac{1}{d_{\mathcal{H}} - 1} \left(\mathrm{Tr}(\rho) \, \mathbb{1}_{\mathcal{H}} - \rho^T \right),$$

where ρ^T denotes the transpose of ρ, for $d_{\mathcal{H}} = 3$.

One possible way to recover the property of tensorization in the quantum realm is to modify the usual definition of the norm of a superoperator as follows: first, we call a linear map $\Phi : \mathcal{A} \to \mathcal{B}$ between two C*-algebras $\mathcal{A}$ and $\mathcal{B}$ *completely bounded* (CB in short) if $\sup_n \| \mathrm{id}_{\mathbb{M}_n(\mathbb{C})} \otimes \Phi : \mathbb{M}_n(\mathbb{C}) \otimes \mathcal{A} \to \mathbb{M}_n(\mathbb{C}) \otimes \mathcal{B}\| < \infty$[3]. In this case, the *completely bounded norm* of Φ is defined as

$$\|\Phi : \mathcal{A} \to \mathcal{B}\|_{\mathrm{cb}} := \sup_{n \in \mathbb{N}} \| \mathrm{id}_{\mathbb{M}_n(\mathbb{C})} \otimes \Phi : \mathbb{M}_n(\mathbb{C}) \otimes \mathcal{A} \to \mathbb{M}_n(\mathbb{C}) \otimes \mathcal{B}\|. \qquad (1.35)$$

---

[3] We refer to the first chapter of [Paulsen, 2002] for a natural construction of the unique norm on $\mathcal{A} \otimes \mathbb{M}_n(\mathbb{C})$, resp. $\mathcal{B} \otimes \mathbb{M}_n(\mathbb{C})$, turning them into C*-algebras.

This definition is closely related to the notion of complete positivity introduced in Section 0.1.5. In fact, it is a well-known fact that any CP map is CB (see Proposition 3.6 of [Paulsen, 2002]).

The dual of the CB norm is usually referred to as *diamond norm* in the quantum information community: assume for sake of simplicity that $\mathcal{A} = \mathcal{B}(\mathcal{H}_A)$ and $\mathcal{B} = \mathcal{B}(\mathcal{H}_B)$, given two separable Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. Then, the diamond norm of $\Psi : \mathcal{T}_1(\mathcal{H}_A) \to \mathcal{T}_1(\mathcal{H}_B)$ is defined as

$$\|\Psi\|_\diamond := \sup_{n \in \mathbb{N}} \|\Psi \otimes \mathrm{id}_{\mathbb{M}_n(\mathbb{C})} : \mathcal{T}_1(\mathcal{H}_A) \otimes \mathcal{T}_1(\mathbb{C}^n) \to \mathcal{T}_1(\mathcal{H}_B) \otimes \mathcal{T}_1(\mathbb{C}^n)\|, \tag{1.36}$$

wherever the quantity on the right hand side of Equation (1.36) is finite. By duality, for every CB map $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$,

$$\|\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)\|_{\mathrm{cb}} = \|\Phi_* : \mathcal{T}_1(\mathcal{H}_A) \to \mathcal{T}_1(\mathcal{H}_B)\|_\diamond.$$

In comparison with the more traditional operator norm $\|. : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_1(\mathcal{H})\|$ defined in the last paragraph, the diamond norm is perhaps physically better motivated: imagine that in the hypothesis testing experiment leading to Equation (1.25), Alice is allowed to send any bipartite state $\rho \in \mathcal{D}(\mathbb{C}^n \otimes \mathcal{H}_A)$ and Bob is allowed to perform any two outcome POVM $\mathbf{M}^n = \{M_1^n, M_2^n\}$ on $\mathbb{C}^n \otimes \mathcal{H}_B$, where the dimension n of the second system can be arbitrarily large. Assuming that the noisy channels $\mathcal{P}_1, \mathcal{P}_2 : \mathcal{T}_1(\mathcal{H}_A) \to \mathcal{T}_1(\mathcal{H}_B)$ only act on the subsystem $\mathcal{H}_A$, Equation (1.25) is then replaced by

$$P_{\mathrm{err},\diamond}(\mathcal{P}_1, \mathcal{P}_2, p)$$
$$:= \inf_{n \in \mathbb{N}} \inf_{\rho_n \in \mathcal{D}(\mathbb{C}^n \otimes \mathcal{H}_A)} \inf_{\mathbf{M}^n} p \operatorname{Tr}(M_2^n (\mathrm{id}_{\mathbb{M}_n(\mathbb{C})} \otimes \mathcal{P}_1)(\rho_n)) + (1-p) \operatorname{Tr}(M_1^n (\mathrm{id}_{\mathbb{M}_n(\mathbb{C})} \otimes \mathcal{P}_2)(\rho_n))$$
$$= \frac{1}{2}(1 - \|p\,\mathcal{P}_1 + (1-p)\,\mathcal{P}_2\|_\diamond).$$

Since in general $\|\Psi\|_\diamond \geq \|\Psi : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_1(\mathcal{H})\|$, $P_{\mathrm{err},\diamond}(\mathcal{P}_1, \mathcal{P}_2, p) \leq P_{\mathrm{err}}(\mathcal{P}_1, \mathcal{P}_2, p)$. This in particular means that more information about the channels $\mathcal{P}_1$ and $\mathcal{P}_2$ is being retrieved by allowing bipartite input states. In the same way as for superoperator norms induced by Schatten norms, one can actually define a whole family of completely bounded $\mathcal{T}_p$ norms interpolating between the completely bounded norm defined in Equation (1.35) and the diamond norm defined in Equation (1.36).

More generally, a complex Banach space $(\mathcal{X}, \|\cdot\|_\mathcal{X})$ is called an *operator space* if it can be embedded isometrically in some $\mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$. Ruan proved in [Ruan, 1988] that this is equivalent to the existence of a family of norms $\|\cdot\|_{k,\mathcal{X}}$ on $\mathbb{M}_k(\mathcal{X})$ for all $k \geq 1$ with the following properties:

(i) $\|\cdot\|_{1,\mathcal{X}} = \|\cdot\|_\mathcal{X}$.

(ii) For all $X \in \mathbb{M}_k(\mathcal{X})$ and all $Y \in \mathbb{M}_l(\mathcal{X})$,

$$\|X \oplus Y\|_{k+l,\mathcal{X}} = \|X\|_{k,\mathcal{X}} + \|Y\|_{l,\mathcal{X}}. \tag{1.37}$$

In particular, if $X \in \mathcal{X}$ and writing $\tilde{X} = |e\rangle\langle e| \otimes X \in \mathbb{M}_k(\mathcal{X})$ for some norm 1 vector e in $\mathbb{C}^k$,

$$\|\tilde{X}\|_{k,\mathcal{X}} = \|X\|_\mathcal{X}. \tag{1.38}$$

(iii) For any $\alpha, \beta \in \mathbb{M}_k(\mathbb{C})$ and $X \in \mathcal{X}$,

$$\|\alpha \cdot X \cdot \beta\|_{k,\mathcal{X}} \leq \|\alpha\|_\infty \|X\|_\mathcal{X} \|\beta\|_\infty.$$

The space $\mathcal{X}$ is said to be given an *operator space structure*. Now, given a map $\Phi : \mathcal{X} \to \mathcal{Y}$ between two operator spaces $\mathcal{X}, \mathcal{Y}$, the *completely bounded norm* of Φ is:

$$\|\Phi : (\mathcal{X}, \{\|.\|_{k,\mathcal{X}}\}_k) \to (\mathcal{Y}, \{\|.\|_{k,\mathcal{Y}}\}_k)\|_{\mathrm{cb}} := \sup_m \|\mathrm{id}_{\mathbb{M}_m} \otimes \Phi : (\mathbb{M}_m(\mathcal{X}), \|.\|_{m,\mathcal{X}}) \to (\mathbb{M}_m(\mathcal{Y}), \|.\|_{m,\mathcal{Y}})\|.$$

In fact, given a Banach space $\mathcal{X}$, there always are many operator space structures possible on $\mathcal{X}$ [Pisier, 2003]. Any of the commutative, non-commutative or amalgamated $\mathbb{L}_p$ spaces defined above has a *natural* abstract operator space structure, that we do not recall here. We refer to [Junge and Parcet, 2010] for more information, and simply recall some important properties of these norms: first, a CB version of Equation (1.33) holds: for $\mathcal{N} \subset \mathcal{M}$, given a $\mathcal{N}$-bimodule map $\Phi : \mathcal{M} \to \mathcal{M}$:

$$\|\Phi : \mathbb{L}_\infty^p(\mathcal{N} \subset \mathcal{M}) \to \mathbb{L}_\infty^q(\mathcal{N} \subset \mathcal{M})\|_{\mathrm{cb}} = \|\Phi : \mathbb{L}_s^p(\mathcal{N} \subset \mathcal{M}) \to \mathbb{L}_s^q(\mathcal{N} \subset \mathcal{M})\|_{\mathrm{cb}}. \qquad (1.39)$$

Next we recall a result on the CB norm in commutative C*-algebras (see e.g. Theorem 3.9 of [Paulsen, 2002]) that will prove very useful in Chapter 9:

Lemma 1.1.14. *Let $\Phi : \mathbb{L}_p(E, \mathcal{F}, \mu) \to \mathbb{L}_q(E, \mathcal{F}, \mu)$, and assume that either $p = 1$ or $q = +\infty$. Then*

$$\|\Phi : \mathbb{L}_p(E, \mathcal{F}, \mu) \to \mathbb{L}_q(E, \mathcal{F}, \mu)\|_{\mathrm{cb}} = \|\Phi : \mathbb{L}_p(E, \mathcal{F}, \mu) \to \mathbb{L}_q(E, \mathcal{F}, \mu)\|.$$

We end this section by mentioning a simple weighted extension of a result from [Devetak et al., 2006]:

Proposition 1.1.15. *Let $\mathcal{H}_1$, $\mathcal{H}_2$, $\mathcal{K}_1$ and $\mathcal{K}_2$ be four finite dimensional Hilbert spaces, and let $\sigma_i \in \mathcal{D}(\mathcal{H}_i)$, $\rho_i \in \mathcal{D}(\mathcal{K}_i)$ be four full-rank density operators. Then, for any $1 \le p, q \le \infty$, and any two completely positive maps $\Phi_i : \mathbb{L}_q(\sigma_i) \to \mathbb{L}_p(\rho_i)$, $i = 1, 2$, the following holds:*

$$\|\Phi_1 \otimes \Phi_2 : \mathbb{L}_q(\sigma_1 \otimes \sigma_2) \to \mathbb{L}_p(\rho_1 \otimes \rho_2)\|_{\mathrm{cb}} = \|\Phi_1 : \mathbb{L}_q(\sigma_1) \to \mathbb{L}_p(\rho_1)\|_{\mathrm{cb}} \|\Phi_2 : \mathbb{L}_q(\sigma_2) \to \mathbb{L}_p(\rho_2)\|_{\mathrm{cb}}.$$

1.2. Entropies

In this section, we introduce a class of measures between two states of upmost importance in quantum information theory, namely entropies. Given two probability measures μ, ν on a measurable space $(\Omega, \mathcal{F})$ with $\mu \ll \nu$, the relative entropy between μ and ν is defined as follows:

$$D(\mu\|\nu) := \int \frac{d\mu}{d\nu} \ln \frac{d\mu}{d\nu} \, d\nu. \qquad (1.40)$$

In the case when ν is replaced by the (unnormalized) Lebesgue measure μ_{Leb} on $\mathbb{R}^n$, the relative entropy is directly related to the Shannon entropy of a random variable X on $\mathbb{R}^n$ whose law μ admits a probability density function f with respect to μ_{Leb}:

$$S(X) \equiv S(f) := -\int f \ln f \, d\mu_{\mathrm{Leb}} \equiv -D(\mu\|\mu_{\mathrm{Leb}}). \qquad (1.41)$$

In Chapter 11, we also use the following notation when $\mu, \nu \ll \mu_{\mathrm{Leb}}$ with associated densities f_1, f_2:

$$D(f_1\|f_2) := D(\mu\|\nu) = \int f_1 \ln \frac{f_1}{f_2} \, d\mu_{\mathrm{Leb}}.$$

In order to generalize the above definitions to the non-commutative setting, one first needs to have an associated notion of Radon-Nikodym derivative at hand. This role is played by the so-called

relative modular operator. Relative modular operators were introduced originally by Araki in order to extend the notion of relative entropy to arbitrary states on a C*-algebra (see [Araki, 1976, Araki, 1977, Ohya and Petz, 1993], we also refer to Petz's papers [Petz, 1985] and [Petz, 1986] for a discussion on the relation between the relative modular operator and Rényi divergences).

Modular theory   Let $\mathcal{M}$ be a von Neumann algebra in a standard representation $(\pi, \mathcal{H}, J, \mathcal{H}^+)$, and let $\omega, \nu \in \mathcal{M}_*^+$. Then, define the linear map $S_{\nu|\omega}$ on the domain $\mathcal{M}\Omega_\omega + (\mathcal{M}\Omega_\omega)^\perp$ by

$$S_{\nu|\omega}(A\Omega_\omega + \Theta) = P_\omega A^* \Omega_\nu, \tag{1.42}$$

for all $A \in \mathcal{M}$, $\Theta \in (\mathcal{M}\Omega_\omega)^\perp$, and where P_ω is the support of ω, namely the projection defined by

$$P_\omega := \inf\{P \in \mathcal{M} | P \text{ is a projection and } \omega(\mathbb{1} - P) = 0\} \tag{1.43}$$

$S_{\nu|\omega}$ is a densely defined anti-linear operator. It is closable and we denote its closure by the same symbol. The positive operator

$$\Delta_{\nu|\omega} := S_{\nu|\omega}^* S_{\nu|\omega} \tag{1.44}$$

is called the *relative modular operator*. In the case $\omega = \nu$, $\Delta_\omega := \Delta_{\omega|\omega}$ reduces to the *modular operator* associated with the pair $(\mathcal{M}, \Omega_\omega)$ already encountered in Section 0.1.5.

Example 1.2.1 (Classical probability theory, continued)**.** With the notations of Example 0.3.2, for any measure $\nu \ll \mu$, the vector Ω_ν defined in 0.3.1 reduces to

$$\Omega_\nu = \left(\frac{d\nu}{d\mu}\right)^{1/2}. \tag{1.45}$$

Using (1.42) and (1.44), one then finds that the relative modular operator of ν with respect to μ is defined by

$$\Delta_{\nu|\mu}(f) := \frac{d\nu}{d\mu} f, \qquad f \in \mathbb{L}_2(\Omega, \mathcal{F}, \mu). \tag{1.46}$$

Example 1.2.2 (Quantum systems, continued)**.** In the framework of Example 0.3.3, given any two positive functionals ω, ν with associated positive, trace-class operators ρ, σ:

$$\Delta_{\rho|\sigma}(A) := \Delta_{\omega|\nu}(A) = \rho A \sigma^{-1}. \tag{1.47}$$

As a linear operator on $\mathcal{B}(\mathcal{H})$, $\Delta_{\rho|\sigma}$ is positive and its spectrum $\mathrm{sp}(\Delta_{\rho|\sigma})$ consists of the ratios of eigenvalues λ/μ, $\lambda \in \mathrm{sp}(\rho)$, $\mu \in \mathrm{sp}(\sigma)$. For any $x \in \mathrm{sp}(\Delta_{\rho|\sigma})$, the corresponding spectral projection is the map

$$P_x(\Delta_{\rho|\sigma}) : \begin{cases} \mathcal{A} & \to & \mathcal{A} \\ A & \mapsto & \sum_{\lambda \in \mathrm{sp}(\rho), \mu \in \mathrm{sp}(\sigma) : \lambda/\mu = x} P_\lambda(\rho) A P_\mu(\sigma) \end{cases} \tag{1.48}$$

In addition, let $\mu_{\rho|\sigma}$ denote the spectral measure for $-\ln\Delta_{\rho|\sigma}$ with respect to $\Omega_\sigma := \sigma^{1/2}$, i.e. the probability measure such that for any bounded measurable function f,

$$\langle \Omega_\sigma, f(-\ln\Delta_{\rho|\sigma})\Omega_\sigma \rangle = \int f(x)\,d\mu_{\rho|\sigma}(x) \equiv \mathbb{E}[f(X)], \tag{1.49}$$

where X is a random variable of law $\mu_{\rho|\sigma}$ (see e.g. Sections VII and VIII of [Reed and Simon, 1972]).

The measure $\mu_{\rho|\sigma}$ can be related to the well-known Nussbaum-Skoła distributions: For two density matrices $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ with spectral decompositions

$$\rho = \sum_\lambda \lambda P_\lambda(\rho), \quad \sigma = \sum_\mu \mu P_\mu(\sigma),$$

these distributions are given by $(p_{\lambda,\mu})_{\lambda,\mu}$ and $(q_{\lambda,\mu})_{\lambda,\mu}$, where

$$p_{\lambda,\mu} = \lambda \operatorname{Tr}\left(P_\lambda(\rho) P_\mu(\sigma)\right), \quad q_{\lambda,\mu} = \mu \operatorname{Tr}\left(P_\lambda(\rho) P_\mu(\sigma)\right).$$

There is of course a connection between the Nussbaum-Szkoła distributions and relative modular operators. Assume for simplicity that all ratios λ/μ are distinct and consider a random variable Z which takes values λ/μ with probability $q_{\lambda,\mu}$. Then using (1.48) one can easily verify that

$$\begin{aligned}
\mathbb{P}(Z = \lambda/\mu) = q_{\lambda,\mu} &= \mu \operatorname{Tr}\left(P_\lambda(\rho) P_\mu(\sigma)\right) \\
&= \langle \Omega_\sigma, P_{\lambda/\mu}(\Delta_{\rho|\sigma}) \Omega_\sigma \rangle \\
&= \langle \Omega_\sigma, \mathbf{1}_{\{\lambda/\mu\}}(\Delta_{\rho|\sigma}) \Omega_\sigma \rangle,
\end{aligned} \tag{1.50}$$

where $\mathbf{1}_{\{\lambda/\mu\}}$ denotes the indicator function on the singleton $\{\lambda/\mu\}$, *i.e.* $\mathbf{1}_{\{\lambda/\mu\}}(x)$ is equal to 1 when $x = \lambda/\mu$ and equal to 0 else. This follows from the fact that, since $\Delta_{\rho|\sigma}$ is self-adjoint, the spectral theorem implies that $\mathbf{1}_{\{\lambda/\mu\}}(\Delta_{\rho|\sigma}) = \sum_{x \in \operatorname{sp}(\Delta_{\rho|\sigma})} \mathbf{1}_{\{\lambda/\mu\}}(x) P_x(\Delta_{\rho|\sigma})$. Equation (1.50) implies that for any bounded measurable function f,

$$\mathbb{E}\left[f(Z)\right] = \langle \Omega_\sigma, f(\Delta_{\rho|\sigma}) \Omega_\sigma \rangle$$

and hence the law of Z is the law of $\Delta_{\rho|\sigma}$ with respect to $\Omega_\sigma := \sigma^{1/2}$. This in turn implies that for any bounded measurable function f,

$$\mathbb{E}\left[f(-\ln Z)\right] = \langle \Omega_\sigma, f(-\ln \Delta_{\rho|\sigma}) \Omega_\sigma \rangle.$$

Hence the law of $-\ln Z$ is precisely $\mu_{\rho|\sigma}$.

Araki's relative entropy   We are now ready to introduce Araki's general definition of the relative entropy. This relative entropy reduces to the classical Kullback-Leibler divergence in the case of equivalent probability measures, and to [Umegaki, 1962]'s quantum relative entropy in the case of states defined through density operators ρ, σ on a fixed separable Hilbert space, where $\operatorname{supp} \rho \subset \operatorname{supp} \sigma$. We invite the interested reader to have a look at [Jaksic et al., 2012, Takesaki, 2003, Ohya and Petz, 1993] for further details.

For any two positive normal functionals on a von Neumann algebra $\mathcal{M}$, we denote by $\mu_{\nu|\omega}$ the spectral measure for $-\ln \Delta_{\nu|\omega}$ with respect to the state ω. This means that it is the only probability measure on the spectrum of $\Delta_{\nu|\omega}$ such that for any bounded measurable function f on $\operatorname{sp}(\Delta_{\nu|\omega})$,

$$\langle \Omega_\omega, f(-\ln \Delta_{\nu|\omega})(\Omega_\omega) \rangle = \int_{\operatorname{sp}(\Delta_{\nu|\omega})} f(x) \, \mu_{\mu|\nu}(dx). \tag{1.51}$$

Then *Araki's relative entropy* of ω with respect to ν is defined by

$$
\text{Ent}(\omega|\nu) := \begin{cases} -\langle \Omega_\omega, \ln(\Delta_{\nu|\omega})\Omega_\omega \rangle = \displaystyle\int_{\text{sp}(\Delta_{\nu|\omega})} x\, \mu_{\nu|\omega}(dx) & \omega << \nu \\ +\infty & \text{otherwise.} \end{cases} \tag{1.52}
$$

where $\omega << \nu$ means that ω is normal with respect to ν, i.e. that $P_\omega \le P_\nu$. Roughly speaking, the above quantity can be interpreted as a measure of distance between two positive normal functionals. As promised, Araki's relative entropy reduces to the classical Kullback-Leibler divergence in the case of classical probability distributions absolutely continuous with respect to a given measure, and to Umegaki's quantum relative entropy in the case of density operators ρ, σ on a Hilbert space $\mathcal{H}$ such that $\text{supp}\,\rho \subset \text{supp}\,\sigma$, which already makes it very attractive from an abstract point of vue:

Example 1.2.3 (Classical probability theory, continued). Let $\mu << \nu$ two probability measures on $(\Omega, \mathcal{F})$, of associate states ω_μ and ω_ν. Then $\text{Ent}(\omega_\mu|\omega_\nu)$ reduces to the relative entropy $D(\mu\|\nu)$ defined in Equation (1.40).

Example 1.2.4 (Quantum systems, continued). In the case of the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on a finite dimensional (or separable) Hilbert space $\mathcal{H}$, Araki's relative entropy reduces to Umegaki's definition:

$$
D(\rho\|\sigma) := \text{Ent}(\omega_\rho|\omega_\sigma) = \text{Tr}(\rho\,(\ln\rho - \ln\sigma)) \tag{1.53}
$$

for normal states $\omega_\rho << \omega_\sigma$ on $\mathcal{B}(\mathcal{H})$ with associated positive density operators ρ, σ.

Araki's relative entropy satisfies the following useful properties (see Theorems 5.3 and 5.20 of [Ohya and Petz, 1993]):

Theorem 1.2.5. *Let $\mathcal{M}$, $\mathcal{N}$, $\mathcal{M}_1$ and $\mathcal{M}_2$ be four von Neumann algebras. Then, the following properties hold:*

- *For any normal state ω on $\mathcal{M}$,*

$$
\text{Ent}(\omega\|\omega) = 0.
$$

- *Additivity: for any normal states ω_1, σ_1 on $\mathcal{M}_1$ and ω_2, σ_2 on $\mathcal{M}_2$,[4]*

$$
\text{Ent}(\omega_1 \otimes \omega_2|\sigma_1 \otimes \sigma_2) = \text{Ent}(\omega_1|\sigma_1) + \text{Ent}(\omega_2|\sigma_2). \tag{1.54}
$$

- *Monotonicity: For any Schwarz map[5] $\Phi: \mathcal{M} \to \mathcal{N}$, and any two normal states ω, ν on $\mathcal{N}$,*

$$
\text{Ent}(\omega \circ \Phi|\nu \circ \Phi) \le \text{Ent}(\omega|\nu). \tag{1.56}
$$

In the case when $\mathcal{M} = \mathcal{B}(\mathcal{H})$ and $\mathcal{N} = \mathcal{B}(\mathcal{H}')$, for two given separable Hilbert spaces $\mathcal{H}, \mathcal{H}'$, the monotonicity property (1.56) is usually stated in the dual Schrödinger picture: For any two states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, $\rho << \sigma$, and any CPTP map $\mathcal{P}$,

$$
D(\mathcal{P}(\rho)\|\mathcal{P}(\sigma)) \le D(\rho\|\sigma). \tag{DPI}
$$

---

[4]For an exposition of tensor products of von Neumann algebras and tensor products of states, see [Kadison and Ringrose, 1983].

[5]A unital map $\Phi: \mathcal{M} \to \mathcal{N}$ between two von Neumann algebras $\mathcal{M}$ and $\mathcal{N}$ is called a *Schwarz map* if for any $X \in \mathcal{M}$,

$$
\Phi(X^*X) \ge \Phi(X)^*\Phi(X). \tag{1.55}
$$

In particular, any CP unital map is a Schwarz map.

This is called the *data processing inequality* (DPI). In words, it states that quantum channels tend to reduce the distinguishability between two states (cf. Section 13.1).

In fact, the properties listed in Theorem 1.2.5 can be used in defining an axiomatic approach to the quantum relative entropies [Matsumoto, 2010, Capel et al., 2018, Wilming et al., 2017]: any function f defined on pairs of quantum states acting on finite dimensional Hilbert spaces of same dimension satisfying the following four properties:

- *Continuity*: For any full-rank state σ, $\rho \mapsto f(\rho, \sigma)$ is continuous on $\mathcal{D}(\mathcal{H})$;

- *Monotonicity*: For any CPTP map $\mathcal{P}$, $f(\mathcal{P}(.), \mathcal{P}(.)) \leq f(.,.)$;

- *Additivity*: For any for states $\rho_i << \sigma_i \in \mathcal{D}(\mathcal{H}_i)$, $f(\rho_1 \otimes \rho_2, \sigma_1 \otimes \sigma_2) = f(\rho_1, \sigma_1) + f(\rho_2, \sigma_2)$;

- *Superadditivity*: For any bipartite state $\rho_{12} \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ of marginals ρ_i and any $\sigma_i \in \mathcal{D}_+(\mathcal{H}_i)$, $f(\rho_{12}, \sigma_1 \otimes \sigma_2) \geq f(\rho_1, \sigma_1) + f(\rho_2, \sigma_2)$;

is proportional to the relative entropy function: $f(.,.) = C\,D(.\|.)$, for some constant $C > 0$. Finally, we mention that the quantum relative entropy and the trace distance satisfy the so-called *quantum Pinsker inequality*:

$$\|\rho - \sigma\|_1 \leq \sqrt{2\,D(\rho\|\sigma)}\,. \tag{1.57}$$

The quantum Pinsker inequality is in particular very useful in getting bounds on the speed of convergence of a QMS from a certain type of quantum functional inequality called the *modified logarithmic Sobolev inequality* (see Part IV).

f-divergences   In the context of discrete measurable spaces, the relative entropy defined in Equation (1.40) can be extended as follows: given a convex function f on $(0, \infty)$ such that $f(1) = 0$, and any two probability vectors $p = (p_i)$ and $q = (q_i)$, define the f-*divergence* between p and q as

$$D_f(p\|q) := \sum_i p_i\, f\left(\frac{q_i}{p_i}\right). \tag{1.58}$$

The usual relative entropy is then retrieved by taking $f(x) = -\ln(x)$. For a review of classical f-divergences, see [Raginsky, 2016].

There exist different generalizations of f-divergences in the quantum setting (see the survey article [Hiai and Mosonyi, 2017], where the authors studied the ability of these divergences to detect reversibility of quantum operations, and references therein for more details). Here, we focus on f-divergences as first defined and studied by Petz in [Petz, 1985] in the context of von Neumann algebras, and [Petz, 1986] (see also [Lesniewski and Ruskai, 1999, Hiai et al., 2011]) in the finite dimensional setting. A function $f : \mathbb{R} \to \mathbb{R}$ is called *operator convex* if for any $0 \leq t \leq 1$ and any bounded, hermitian operators A, B on any Hilbert space $\mathcal{H}$:

$$f\left(tA + (1-t)B\right) \leq t\,f(A) + (1-t)\,f(B)\,, \tag{1.59}$$

Then f is called *operator concave* if $-f$ is operator convex. The concept of operator convexity is closely related to the one of operator monotonicity: a function $f : \mathbb{R} \to \mathbb{R}$ is called *operator monotone* if for any separable Hilbert space $\mathcal{H}$ and any two bounded hermitian operators $A, B \in \mathcal{B}(\mathcal{H})$, $A \leq B$ implies that $f(A) \leq f(B)$. In fact, the following result can be found in [Bhatia, 1997]:

Theorem 1.2.6. *Let $f : [0, \infty) \to [0, \infty)$ a continuous function. Then f is operator monotone if and only if it is operator concave. Moreover, let $g : (0, \infty) \to (0, \infty)$. Then g is operator monotone if and only if $1/g$ is operator convex.*

The well-known *Loewner theorem* states that any operator convex function $f : \mathbb{R}_+ \to \mathbb{R}$ such that $f(1) = 0$ satisfies the following integral representation:

$$f(u) = \alpha(u-1) + \beta(u-1)^2 + \gamma \frac{(u-1)^2}{u} + \int_0^\infty \frac{(u-1)^2}{u+s} \nu(ds), \tag{1.60}$$

where $\alpha \in \mathbb{R}$, $\beta, \gamma \geq 0$ and ν is a positive measure on $\mathbb{R}_+$ satisfying $\int_0^\infty (1+t^2)^{-1} \nu(dt) < \infty$ (see [Bhatia, 1997]).

For sake of simplicity, we reduce ourselves to the case where $\mathcal{M} = \mathcal{B}(\mathcal{H})$ is the algebra of bounded operators on a finite dimensional Hilbert space $\mathcal{H}$. Then, for any operator convex function f, and any two states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, the *$f$-divergence* of ρ with respect to σ is defined as

$$D_f(\rho\|\sigma) = \begin{cases} \operatorname{Tr}\left(\rho^{1/2} f(\Delta_{\sigma|\rho})(\rho^{1/2})\right) - f(1) & \operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma) \\ \infty & \text{otherwise.} \end{cases} \tag{1.61}$$

In finite dimensions, f-divergences are differentiable monotone relative entropy distances in the sense defined in [Lesniewski and Ruskai, 1999]. In particular, they all are monotonous under CPTP maps.

Example 1.2.7 (α-Rényi divergences)**.** After the relative entropy, the *α-Rényi divergences* constitute perhaps the most important family of f-divergences: for any full-rank states ρ, σ and any $\alpha \in \mathbb{R}$, $\Delta_{\rho|\sigma}^\alpha(A) = \rho^\alpha A \sigma^{-\alpha}$. Then, given $\alpha \in (0, \infty) \backslash \{1\}$, define

$$D_\alpha(\rho\|\sigma) = \begin{cases} \dfrac{1}{\alpha - 1} \ln \langle \Omega_\sigma, \Delta_{\rho|\sigma}^\alpha (\Omega_\sigma) \rangle_{\mathrm{HS}} = \dfrac{1}{\alpha - 1} \ln \operatorname{Tr}(\rho^\alpha \sigma^{1-\alpha}) & \operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma) \\ \infty & \text{otherwise} \end{cases}$$

It is a well-known fact that $\alpha \mapsto D_\alpha(\rho\|\sigma)$ is monotonically increasing and that $D_\alpha \to D$ as $\alpha \to 1$. This definition can easily be extended to the case of a separable Hilbert space as long as the quantity on the right-hand side is well-defined.

Sandwiched Rényi divergences In the previous example, we introduced a quantum generalization of the α-Rényi divergence. Another useful extension was rather recently independently introduced in [Müller-Lennert et al., 2013, Wilde et al., 2014] (also see [Jenčová, 2018, Berta et al., 2018] for extensions to arbitrary von Neumann algebras). Here, we assume once again that $\mathcal{M} = \mathcal{B}(\mathcal{H})$ for a given separable Hilbert space $\mathcal{H}$. Then, given any two states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ and $\alpha \in (0, \infty) \backslash \{1\}$, the *sandwiched α-Rényi relative entropy* of ρ with respect to σ is defined as

$$\tilde{D}_\alpha(\rho\|\sigma) = \begin{cases} \dfrac{1}{\alpha - 1} \ln \left\{ \operatorname{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \, \sigma^{\frac{1-\alpha}{2\alpha}} \right)^\alpha \right\} & \operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma) \text{ or } (\alpha \in (0,1) \text{ and } \rho \perp \sigma) \\ \infty & \text{otherwise.} \end{cases} \tag{1.62}$$

By definition, the α-Rényi entropy is directly related to the weighted $\mathbb{L}_\alpha(\sigma)$ as defined in Section 1.1.1: assuming that σ is faithful,

$$\tilde{D}_\alpha(\rho\|\sigma) = \frac{\alpha}{\alpha - 1} \ln \|\Gamma_\sigma^{-1}(\rho)\|_{\mathbb{L}_\alpha(\sigma)}.$$

Here, $\Gamma_\sigma^{-1}(\rho)$ replaces the relative modular operator in the definition of the relative entropy, and can hence be interpreted as another type of non-commutative generalization of the Radon-Nikodym derivative. Sandwiched α-Rényi divergences were shown to satisfy the data processing inequality [Frank and Lieb, 2013, Beigi, 2013]: given $\alpha \in [1/2, \infty)$, for any CPTP map $\mathcal{P}$,

$$\tilde{D}_\alpha(\mathcal{P}(\rho)\|\mathcal{P}(\sigma)) \leq \tilde{D}_\alpha(\rho\|\sigma).$$

Moreover, there exist counterexamples for the range $\alpha \in (0, 1/2)$ [Berta et al., 2017]. Just as in the case of the α-Rényi divergences, sandwiched α-Rényi divergences are monotonically increasing and satisfy $\tilde{D}_\alpha \to D$ as $\alpha \to 1$. Moreover both Petz and sandwiched Rényi divergences converge to the so-called *max relative entropy* as $\alpha \to \infty$ [Datta, 2009, Müller-Lennert et al., 2013]:

$$D_{\max}(\rho\|\sigma) \equiv D_\infty(\rho\|\sigma) := \lim_{\alpha \to \infty} D_\alpha(\rho\|\sigma) = \lim_{\alpha \to \infty} \tilde{D}_\alpha(\rho\|\sigma) = \min\left\{\gamma : \rho \leq e^\gamma \sigma\right\}. \tag{1.63}$$

The max relative entropy is additive and *quasi-convex*: for any two mixtures $\rho = \sum_i \lambda_i \rho_i$ and $\sigma := \sum_i \mu_i \sigma_i$ of states $\{\rho_i\}$ and $\{\sigma_i\}$:

$$D_{\max}(\rho\|\sigma) \leq \max_i D_{\max}(\rho_i\|\sigma_i). \tag{1.64}$$

It was observed in [Datta and Leditzky, 2014, Wilde et al., 2014] that the relation

$$D_\alpha(\rho\|\sigma) \geq \tilde{D}_\alpha(\rho\|\sigma)$$

follows for $\alpha \in (0, 1) \cup (1, \infty)$ from the Araki-Lieb-Thirring inequality [Lieb and Thirring, 1991, Araki, 1990]:

Lemma 1.2.8 (Araki-Lieb-Thirring inequality)**.** *For any $A, B \in \mathcal{P}(\mathcal{H})$, and $r \in [0, 1]$,*

$$\mathrm{Tr}(B^{r/2} A^r B^{r/2}) \leq \mathrm{Tr}(B^{1/2} A B^{1/2})^r.$$

Maximal f-divergences We introduce a second quantum extension of the f-divergence of Equation (1.58). This version of f-divergences was formerly treated in [Petz and Ruskai, 1998], and was more recently studied in detail by Matsumoto [Matsumoto, 2013] (see also [Hiai and Mosonyi, 2017]): let $\mathcal{H}$ be a finite dimensional Hilbert space and $f : (0, \infty) \to \mathbb{R}$ an operator convex function. Then for any full-rank states ρ, σ, the *maximal f-divergence* of ρ with respect to σ is defined as follows:

$$\hat{D}_f(\rho\|\sigma) := \mathrm{Tr}\left(\sigma f(\Gamma_\sigma^{-1}(\rho))\right). \tag{1.65}$$

The denomination "maximal" comes from the fact that (see Proposition 4.1 of [Matsumoto, 2013]):

$$D_f(\rho\|\sigma) \leq \hat{D}_f(\rho\|\sigma). \tag{1.66}$$

In Section 12.7, we use the maximal entropy associated to the operator convex function $f(x) = x \ln x$, which was originally studied by [Belavkin, 1982] as yet another quantum extension of the Kullback-Leibler divergence (1.40). In this case,

$$\hat{D}(\rho\|\sigma) := \hat{D}_{x \mapsto x \ln x}(\rho\|\sigma) = \mathrm{Tr}\left(\Delta_\sigma^{\frac{1}{2}}(\rho) \ln \Gamma_\sigma^{-1}(\rho)\right). \tag{1.67}$$

Strong subadditivity of the von Neumann entropy As we will see in Chapter 13, many fundamental rates occurring in quantum Shannon theory can be related to one of the quantum divergences introduced above. Here, we simply introduce the quantum entropy, which characterizes the rate at which quantum information can be compressed [Schumacher, 1995, Jozsa and Schumacher, 1994, Lo, 1995]: given a quantum state $\rho \in \mathcal{D}(\mathcal{H})$ on a finite dimensional Hilbert space $\mathcal{H}$, the *von Neumann entropy* of ρ is defined as

$$S(\rho) = -\operatorname{Tr}(\rho \ln \rho).$$

It is easy to verify that the von Neumann entropy is related to the quantum relative entropy as follows:

$$S(\rho) = -D(\rho \| \mathbb{1}_{\mathcal{H}}),$$

where we naturally extended the definition of $D(. \|.)$ to non-normalized positive operators. The von Neumann entropy satisfies what is usually considered as the most important inequality in the field of quantum information theory, the so-called *strong subadditivity* (SSA): given a tripartite system $\mathcal{H} := \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{B}_C$, and any state $\rho \in \mathcal{D}(\mathcal{H})$

$$S(\rho_{ABC}) + S(\rho_B) \le S(\rho_{AB}) + S(\rho_{BC}). \tag{1.68}$$

By now, there exist many proofs of the above inequality. It was first proved by [Lieb and Ruskai, 1973]. The extension to general von Neumann algebras was done by [Narnhofer and Thirring, 1985]. It is known to be equivalent to the data processing inequality (cf. Equation (1.56)).

Remark 1.2.9. In Theorem 8.3.3, letting $X(t) \equiv X$ for all t and $p(t) = q + t$ provides an functional analytic justification of the term entropy, as it yields:

$$\frac{d}{dp} \| \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{-1}{q}}(\rho^{\frac{1}{q}}) \|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} \bigg|_{p=q} = \frac{1}{q^2} D(\rho \| E_{\mathcal{N}*}(\rho)). \tag{1.69}$$

We see here the tight relationship between the amalgamated $\mathbb{L}_p$ norms and entropic quantities that appear in quantum information theory. This link was recently exploited in [Gao et al., 2017] to prove a generalisation of the celebrated SSA inequality. At the core of the proof lies the complex interpolation property of the amalgamated $\mathbb{L}_p$ norms.

1.3. Fisher information

In this last section, we introduce yet another type of distance on quantum states, namely *quantum Fisher information metrics*. They have already proved to be very useful in the context of quantum parameter estimation (see Section 13.2). Moreover, they turn out to be closely related to the entropy production of a quantum Markovian evolution as described in Chapter 12. For more information, we refer the reader to [Hayashi, 2016, Hayashi, 2005, Frieden, 1998].

The notion of *Fisher information* appeared in the 1920's in the field of parameter statistical estimation to measure the difficulty of reconstructing an unknown parameter: Let $(f_\theta(x))_{\theta \in \Theta}$ be a family of probability density functions on $\mathbb{R}$, depending smoothly on an unknown parameter θ. The *Fisher information* of this family is defined as

$$\mathrm{I}(\theta) := \frac{d^2}{d\alpha^2} \bigg|_{\alpha=\theta} D(f_\theta \| f_\alpha) = -\mathbb{E}_\theta \left[\frac{\partial^2}{\partial \theta^2} \ln f_\theta(x) \right] = -\frac{\partial}{\partial \alpha} \frac{\partial}{\partial \beta} D(f_\alpha \| f_\beta) \bigg|_{\alpha=\beta=\theta}, \tag{1.70}$$

where $\mathbb{E}_\theta$ denotes the expected valued with respect to the probability measure $\mu_\theta(dx) = f_\theta(x)\,dx$. Assume one has access to n observations of a real random variable $X : \Omega \to \mathbb{R}$ sampled from a distribution f_{θ_0} of true parameter $\theta_0 \in \Theta$. In order to infer the value of the true parameter θ_0 from the accessible data, one constructs a *statistical estimator*, that is, a random variable $\hat\theta(X_1, ..., X_n) : \Omega \to \Theta$, where the i.i.d. random variables $X_1, ..., X_n$ of law f_{θ_0} correspond to the n observations of the random variable X. We moreover assume that the estimator $\hat\theta$ is unbiased, i.e. $\mathbb{E}[\hat\theta] = \theta_0$. Next theorem is a cornerstone of the field of parameter estimation:

Theorem 1.3.1 (Cramér-Rao bound). *For any unbiased statistical estimator $\hat\theta$,*

$$\mathrm{Var}(\hat\theta) \geq \frac{1}{\mathrm{I}(\theta_0)}\,.$$

Hence, the (inverse) Fisher information imposes a lower bound on the accuracy of the estimator $\hat\theta$. Asymptotically (when the number of observations goes to infinity), the bound is achieved by the maximum likelihood estimator $\hat\theta = \arg\max_\theta \frac{1}{n} \sum_{i=1}^n f_\theta(X_i)$.

Now, assume that

$$f_\theta(x) := f(x - \theta)\,,$$

for some given smooth probability density function f on $\mathbb{R}$. In this model, we are interested in centering a distribution whose profile is known. Then,

$$\mathrm{I}(\theta) = -\int \frac{\partial^2}{\partial x^2} \ln f_\theta(x)\, f_\theta(x)dx = \int \frac{\partial}{\partial x} \ln f(x - \theta)\, \frac{\partial}{\partial x} f(x - \theta)dx = \int \frac{|\partial_x f|^2}{f}\, dx \equiv \mathrm{I}(f) \quad (1.71)$$

is independent of θ. This functional will reappear in Part II in the study of functional inequalities for Markovian dynamics. For the time being, we mention that the Fisher information of the density f in this case is related to the Shannon differential entropy of its associated random variable X. The following identity was proved with higher degrees of generality. Its is of crucial use in the derivation of the information theoretic inequalities of Section 4.5, and was also used in [Barron, 1986] to get a strengthened entropic version of the central limit theorem. It was first derived by [a.J. Stam, 1959] who gave credit to de Bruijn for it (see also [Blachman, 1965]). The version provided here is due to [Barron, 1984]:

Lemma 1.3.2 (de Bruijin's identity). *Let X be a random variable on $\mathbb{R}^n$ of finite variance, and G an independent standard Gaussian random variable. Then, for any $\sigma > 0$:*

$$\frac{dS(X + \sigma G)}{d\sigma^2} = \frac{1}{2} \mathrm{I}(f_{X+\sigma G})\,, \tag{1.72}$$

where $f_{X+\sigma G}(y) := \mathbb{E}[g_\sigma(y - X)]$ denotes the density of the random variable $X + \sigma G$, g_σ being the density of the Gaussian random variable σG of variance σ^2.

So far, we have introduced the Fisher information associated to the probability density function of a continuous random variable. In Chapter 12 and section 13.2, we will mostly deal with quantum extensions of Fisher informations in the setting of discrete probability spaces. These are the object of the remaining of this section.

Classical information geometry   The main idea behind the field of classical information geometry is to provide a Riemannian structure, also known as *statistical manifold*, to sets of probability distributions over a fixed set Ω. Information geometry has proven to be a very fruitful framework for the study of statistical parameter estimation. Here, the role of the metric is played by the

Fisher information metric, the construction of which is briefly recalled here: given a strictly positive probability mass function p on a finite set Ω^6, and two random variables $X, Y : \Omega \to \mathbb{R}$, the covariance of X and Y is given by

$$\mathrm{Cov}_p(X, Y) = \sum_{\omega \in \Omega} (X(\omega) - \mathbb{E}_p[X])(Y(\omega) - \mathbb{E}_p[Y]) \, p(\omega) \,,$$

where we used the symbol $\mathbb{E}_p$ in order to refer to the expectation of a random variable under the measure corresponding to the function p. If X is independent of Y, their covariance is equal to 0. On the contrary, if $X = Y$, their covariance is equal to the variance $\mathrm{Var}_p(X)$ of X. This definition of the covariance justifies the introduction of the following very useful *weighted*-inner product

$$\langle X, Y \rangle_p := \sum_{\omega \in \Omega} X(\omega) Y(\omega) \, p(\omega) \,. \tag{1.73}$$

Suppose now given an element p_{θ_0} belonging to a smooth family $\mathcal{M}_\Theta := \{ p_\theta | \ \theta \in \Theta \}$ of strictly positive probability mass functions indexed by a parameter $\theta := (\theta_1, ..., \theta_d)$ lying in a submanifold Θ of $\mathbb{R}^d$, where the local coordinates θ_i of θ are defined with respect to a prefixed canonical basis $(e_1, ..., e_d)$ of $\mathbb{R}^d$, and assume that for any $\omega \in \Omega$, $\theta \mapsto p_\theta(\omega)$ is differentiable. The goal of *parameter estimation* is to determine the value of the parameter θ_0 (see Section 13.2 for further details). In order to do so, first introduce the *logarithmic derivative* at $\theta_0 \in \Theta$ as follows: for any $i = 1, ..., d$,

$$l_i^{\theta_0}(\omega) := \left. \frac{\partial \ln p_\theta(\omega)}{\partial \theta_i} \right|_{\theta = \theta_0} \equiv \frac{1}{p_{\theta_0}(\omega)} \left. \frac{\partial p_\theta(\omega)}{\partial \theta_i} \right|_{\theta = \theta_0} . \tag{chain rule}$$

The logarithmic derivative can be regarded as a random vector under the law p_{θ_0}. The *Fisher information matrix*

$$J_{ij}^\theta := \langle l_i^\theta, l_j^\theta \rangle_{p_\theta} \tag{1.74}$$

hence represents the amount of variation in the probability distribution due to the variations in the parameter. In other words, it indicates how much information on the true parameter θ_0 one can extract from the family $\mathcal{M}_\Theta$. By the chain rule,

$$J_{ij}^\theta \equiv \sum_{\omega \in \Omega} \frac{1}{p_\theta(\omega)} \frac{\partial p_\theta(\omega)}{\partial \theta_i} \frac{\partial p_\theta(\omega)}{\partial \theta_j} = \sum_{\omega \in \Omega} p_\theta(\omega) \frac{\partial \ln p_\theta(\omega)}{\partial \theta_i} \frac{\partial \ln p_\theta(\omega)}{\partial \theta_j} \equiv \langle \partial_{\theta_i} p_\theta, \partial_{\theta_j} p_\theta \rangle_{1/p_\theta} \,,$$

We can hence distinguish between two types of inner products, the so-called *exponential inner product*, denoted by $\langle X, Y \rangle_p^{(e)}$ and whose expression was already given in Equation (1.73), and the so-called *mixture inner product* defined as

$$\langle X, Y \rangle_p^{(m)} := \langle X, Y \rangle_{1/p}^{(e)}. \tag{1.75}$$

These inner products are in some sense dual, and as we will see later, their quantum extensions play a major role in the theory of quantum information geometry.

The tangent space $\mathcal{T}_{p_\theta} \mathcal{M}_\Theta$ of $\mathcal{M}_\Theta$ at p_θ can be represented in two different ways, each related to one of the two inner products defined above. In the so-called *e-representation*, a tangent vector $X := (x_1, ..., x_d)$ is represented by $X^{(e)} = \sum_{i=1}^d x_i \, \partial_{\theta_i} p_\theta$, and we denote the corresponding tangent space at p_θ by $\mathcal{T}_{p_\theta}^{(e)} \mathcal{M}_\Theta$, whereas in the so-called *m-representation*, X is represented by $X^{(m)} := \sum_{i=1}^d x_i \, l_i^\theta$, and we denote the corresponding tangent space at p_θ by $\mathcal{T}_{p_\theta}^{(m)} \mathcal{M}_\Theta$. No matter the representation, the

---

^{6}For extensions to more general measure spaces, see [Itoh and Shishido, 2008].

Fisher information metric g_θ^{F} at p_θ is then defined as follows: for each $\theta \in \Theta$, and any two tangent vectors $X, Y \in \mathcal{T}_{p_\theta}\mathcal{M}_\Theta$,

$$g_{p_\theta}^{\mathrm{F}}(X,Y) = \langle X^{(m)}, Y^{(m)} \rangle_{p_\theta}^{(m)} \equiv \langle X^{(e)}, Y^{(e)} \rangle_{p_\theta}^{(e)} \equiv \sum_{i,j=1}^{d} x_i \, J_{ij}^\theta \, y_j,$$

where the equivalence between the two representations comes from the chain rule.

Obviously, these concepts generalize to the basis-free, nonparametric case, where the distribution to estimate lies in (a submanifold of) the set of all full support probability mass functions on Ω. Chentsov showed in [Cencov, 2000] that the Fisher information metric is the unique monotone metric on the set of faithful states. It is also the Hessian of any f-divergence:

$$g_p^{\mathrm{F}}(X,Y) = -\frac{\partial^2}{\partial\alpha\partial\beta} D_f(p + \alpha X \| p + \beta Y), \qquad (1.76)$$

where D_f is defined in Equation (1.58).

Quantum weighted inner products The classical weighted inner product $\langle X, Y \rangle_p$ does not possess a unique canonical quantum extension. This is due to the fact that there is an infinite amount of ways of extending Equation (1.73) to the quantum realm. Let $\mathcal{H}$ be a finite dimensional Hilbert space. An inner product $\langle ., . \rangle$ on $\mathcal{B}(\mathcal{H})$ is said to be *compatible* with a full-rank state ρ if for all $X \in \mathcal{B}(\mathcal{H})$, $\mathrm{Tr}(\rho X) = \langle \mathbb{1}_\mathcal{H}, X \rangle$. Analogously to the classical setting, for any completely positive, unital map Φ that is self-adjoint with respect to an inner product $\langle ., . \rangle$ compatible with ρ, ρ is invariant with respect to Φ. Indeed, using the fact that Φ is unital, we get:

$$\mathrm{Tr}(\rho X) = \langle \mathbb{1}, X \rangle = \langle \Phi(\mathbb{1}), X \rangle = \langle \mathbb{1}, \Phi(X) \rangle = \mathrm{Tr}(\rho \, \Phi(X)), \qquad \forall X \in \mathcal{B}(\mathcal{H}). \qquad (1.77)$$

Assume now that the inner product $\langle ., . \rangle$ is of the following form:

$$\langle Y, X \rangle = \langle Y, [\rho](X) \rangle_{\mathrm{HS}}, \qquad (1.78)$$

where $[\rho]$ is a linear map satisfying the following conditions:

$$[\rho](\mathbb{1}) = \rho,$$
$$[\rho \otimes \rho'](X \otimes X') = [\rho](X) \otimes [\rho'](X'),$$
$$[\rho](U^* X U) = U^*[\rho](X)U,$$

for an arbitrary unitary operator U. These conditions imply that $\langle ., . \rangle$ is compatible with ρ. Let's further restrict the class of inner products that we are going to work with: given any function $f : (0, \infty) \to (0, \infty)$ such that

$$\frac{1}{f} \in \mathrm{OMM} := \{k : \mathbb{R}_+ \to \mathbb{R} | -k \text{ is operator monotone }, \ k(x^{-1}) = xk(x), \ k(1) = 1\},$$

and a full-rank state ρ, one can easily verify that the following quadratic form defines a compatible inner product:

$$\langle X, Y \rangle_{f,\rho} := \langle X, R_\rho \circ f(\Delta_\rho)(Y) \rangle_{\mathrm{HS}}, \qquad (1.79)$$

where $R_\rho : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is the operator of right multiplication by ρ, and $\Delta_\rho : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is the modular operator $\Delta_\rho(A) \coloneqq \rho A \rho^{-1}$. The associated kernel is given by

$$[\rho]_f \coloneqq R_\rho \circ f(\Delta_\rho) \,.$$

The inner product (1.79) reduces to the exponential inner product (1.73) in the commutative case, up to a multiplicative constant $f(1)$. Hence, $[\rho]_f$ is the non-commutative extension of the operation of multiplication by ρ. Therefore, it constitutes a quantum extension of the classical exponential inner product defined in Equation (1.73), and will be denoted by $\langle .,. \rangle_{f,\rho}^{(e)}$. Likewise, the mixture inner product defined in Equation (1.75) is extended to the quantum setting by the dual relation

$$\langle A, B \rangle_{f,\rho}^{(m)} \coloneqq \langle A, [\rho]_f^{-1}(B) \rangle_{\text{HS}}$$

thanks to the correspondence $A = [\rho]_f(X)$. The case when ρ is not faithful can also be taken care of by reducing the analysis to the support of ρ (see p. 256 of [Hayashi, 2016]). The fact that the inner products $\langle .,. \rangle_{f,\rho}^{(e)}$ and $\langle .,. \rangle_{f,\rho}^{(m)}$ are *monotone* under the action of CPTP maps was first shown by Petz and Ruskai in [Petz and Ruskai, 1998], and another proof was given in [Lesniewski and Ruskai, 1999] based on integral representation of the function f:

Theorem 1.3.3. *For any CPTP map Φ, the following holds:*

$$\|A\|_{f,\rho}^{(m)} \ge \|\Phi(A)\|_{f,\Phi(\rho)}^{(m)}, \qquad \|X\|_{f,\rho}^{(e)} \ge \|[\Phi(\rho)]_f^{-1} \circ \Phi \circ [\rho]_f(X)\|_{f,\Phi(\rho)}^{(e)} \,.$$

Riemannian structures on quantum states A *Riemiannian metric* on the manifold $\mathcal{D}_+(\mathcal{H})$ of full-rank states acting on a finite-dimensional Hilbert space $\mathcal{H}$ is a smooth map $\rho \mapsto g_\rho$ which, to each element $\rho \in \mathcal{D}_+(\mathcal{H})$ associates an element of $\mathcal{T}_\rho^* \mathcal{D}_+(\mathcal{H}) \otimes \mathcal{T}_\rho^* \mathcal{D}_+(\mathcal{H})$, where, for each state ρ, $\mathcal{T}_\rho^* \mathcal{D}_+(\mathcal{H})$ represents the dual of $\mathcal{T}_\rho \mathcal{D}_+(\mathcal{H})$. Hence, g_ρ is a positive definite bilinear form on the tangent bundle $\mathcal{T}_\rho \mathcal{D}_+(\mathcal{H})$ such that for any constant vector field $A \in \mathcal{T}_* \mathcal{D}_+(\mathcal{H})$, the map $\rho \mapsto g_\rho(A_\rho, A_\rho)$ is smooth. We recall that the tangent bundle $\mathcal{T}_* \mathcal{D}_+(\mathcal{H}) = \cup_{\rho \in \mathcal{D}_+(\mathcal{H})} \mathcal{T}_\rho \mathcal{D}_+(\mathcal{H})$ consists of fibers $\mathcal{T}_\rho \mathcal{D}_+(\mathcal{H})$ which are all isomorphic to the vector space of traceless, self-adjoint operators:

$$\mathcal{T}_\rho \mathcal{D}_+(\mathcal{H}) \simeq \mathcal{T}^{(m)} \mathcal{D}_+(\mathcal{H}) \coloneqq \{A \in \mathcal{B}_{sa}(\mathcal{H}) | \operatorname{Tr}(A) = 0\} \,.$$

This representation of the tangent bundle is called the *mixture representation* (or m-representation for short) of $\mathcal{T}_* \mathcal{D}_+(\mathcal{H})$. Remark that in this representation, the tangent spaces do not depend on the point ρ on the statistical manifold $\mathcal{D}_+(\mathcal{H})$. Expressed in this representation, the metric $g^{(m)}$ is *monotone* if it contracts under any CPTP map Φ: for any $\rho \in \mathcal{D}_+(\mathcal{H})$ and all $A, B \in \mathcal{T}^{(m)} \mathcal{D}_+(\mathcal{H})$,

$$g_{\Phi(\rho)}^{(m)}(\Phi(A), \Phi(B)) \le g_\rho^{(m)}(A, B) \,.$$

Another representation is the so-called *exponential representation* (or e-representation for short), and is given by

$$\mathcal{T}_\rho \mathcal{D}_+(\mathcal{H}) \simeq \mathcal{T}_\rho^{(e)} \mathcal{D}_+(\mathcal{H}) \coloneqq \{A \in \mathcal{B}_{sa}(\mathcal{H}) | \operatorname{Tr}(\rho A) = 0\}$$

Now, by Theorem 1.3.3, monotone Riemannian metrics on $\mathcal{T}_* \mathcal{D}_+(\mathcal{H})$ can be constructed in terms of either:

(i) the inner products $\langle .,. \rangle_{f,\rho}^{(m)}$ on $\mathcal{T}_\rho^{(m)} \mathcal{D}_+(\mathcal{H})$, or

(ii) the inner products $\langle .,.\rangle^{(e)}_{f,\rho}$ on $\mathcal{T}^{(e)}_{\rho}\mathcal{D}_{+}(\mathcal{H})$.

We write $g^{(m)}$ and $g^{(e)}$ the metrics derived from the inner products $\langle .,.\rangle^{(m)}$ and $\langle .,.\rangle^{(e)}$, and call them the m, respectively e, representations of the *quantum Fisher information metrics*. To go from the m-representation to the e-representation, at a particular point ρ, and for a particular extension of the inverse encoded by a function f as above, we use the map $[\rho]_f$:

$$A = [\rho]_f(X).$$

The connection with the commutative chain rule is done as follows: assume we are given a chart of local coordinates in $(\theta_1,...,\theta_d)$, and a tangent vector at the point ρ_{θ_0}, in the m-representation, given by its coordinates in the basis $\partial_{\theta_j}\rho_\theta\big|_{\theta=\theta_0}$, the corresponding e-representation of this tangent vector is given by the coordinates in the basis

$$L_{\theta_0,j,f} := [\rho_{\theta_0}]_f^{-1}(\partial_{\theta_j}\rho_\theta|_{\theta=\theta_0}). \tag{1.80}$$

Then, $L_{\theta_0,j,f}$ can be interpreted as a quantum extension of the logarithmic derivative, and Equation (1.80) as a quantum extension of the chain rule.

Monotone Riemannian metrics from f-divergences In analogy with the classical setting, quantum f-divergences can be used to characterize monotone Riemannian metrics on the manifold $\mathcal{D}_{+}(\mathcal{H})$. The big difference with the classical scenario lies in the non-uniqueness of these metrics. Their characterization was initiated by Morozova and Chentsov in [Morozova and Chentsov, 1991], but no explicit examples were given in their study. Petz and coauthors then gave a complete characterization in a series of papers [Petz, 1995, Petz and Sudár, 1999, Petz and Ruskai, 1998, Petz and Sudár, 1996] (for an extension to infinite dimensions, see the work of Jencova [Jenčová, 2018]). There they proved that there is a one-to-one correspondence between the set of monotone metrics and the set OMM. In [Lesniewski and Ruskai, 1999], the authors further showed that there is a one-to-one correspondence between the set OMM and the set of symmetric operator convex functions f^{sym} such that $f^{\mathrm{sym}}(1) = 0$ (see Theorem II.13 of [Lesniewski and Ruskai, 1999]). By symmetric we mean such that for all $x > 0$:

$$x\,f^{\mathrm{sym}}(x^{-1}) = f^{\mathrm{sym}}(x). \tag{1.81}$$

The correspondence goes as follows: for any $k \in$ OMM,

$$f^{\mathrm{sym}}(x) = (x-1)^2 k(x). \tag{1.82}$$

Moreover to any operator convex f such that $f(1) = 0$ one can associate a symmetric operator convex function f^{sym} through

$$f^{\mathrm{sym}}(x) = f(x) + x\,f(x^{-1}), \tag{1.83}$$

with associated symmetrized f-divergence[7]

$$D_{f^{\mathrm{sym}}}(\rho\|\sigma) = D_f(\rho\|\sigma) + D_f(\sigma\|\rho). \tag{1.84}$$

This implies that there is a one-to-one correspondence between the set of symmetrized f-divergences and the set of monotone Riemannian metrics. In fact the following theorem makes this link explicit

---

[7]Here, we used the fact that $D_{x\to x f(x^{-1})}(\rho\|\sigma) = D_f(\sigma\|\rho)$.

(see Theorem II.8 of [Lesniewski and Ruskai, 1999]):

Theorem 1.3.4. *For any operator convex function $f : (0, \infty) \to (0, \infty)$ such that $f(1) = 0$, and any two traceless hermitian operators A and B, the expression*

$$g_{f,\rho}^{(m)}(A, B) := -\frac{\partial^2}{\partial\alpha\partial\beta}D_f(\rho + \alpha A \| \rho + \beta B)|_{\alpha=\beta=0} \equiv \langle A, [\rho]_{\tilde{f}}^{-1}(B)\rangle_{\mathrm{HS}} \qquad \rho \in \mathcal{D}_+(\mathcal{H}), \qquad (1.85)$$

where $[\rho]_{\tilde{f}} := R_\rho \tilde{f}(\Delta_\rho)$, defines a monotone Riemannian metric on $\mathcal{D}_+(\mathcal{H})$, where

$$\tilde{f}(x) := \frac{(x-1)^2}{f(x) + x\,f(x^{-1})}\,. \qquad (1.86)$$

Moreover, any monotone Riemannian metric $g^{(m)}$ can be written in the form of Equation (1.85) for some function $\tilde{f} = 1/k$, where $k \in \mathrm{OMM}$, and its associated symmetric f-divergence D_f, where $f(x) = f^{\mathrm{sym}}(x) = (x-1)^2\,k(x)$.

The above reverse implication that every monotone Riemannian metric stems from a generalized relative entropy was first proved by Ruskai and Lesniewski in [Lesniewski and Ruskai, 1999]. This result is in contrast with the above classical case, where there exists a unique monotone Riemannian metric, Fisher information metric, on the set of probability vectors [Cencov, 2000]. In fact, reducing to diagonal operators ρ, A, B, for any f, Ω_ρ^f reduces to

$$[\rho]_f = M_\rho\,,$$

where M_ρ denotes the operation of multiplication by ρ, and all the monotone Riemannian metrics defined in Equation (1.85) reduce to the unique classical Fisher information metric:

$$g_{f,\rho}^{(m)}(A, B) = \mathrm{Tr}(\rho^{-1}AB) \equiv \sum_{\omega\in\Omega} p(\omega)^{-1} f(\omega)g(\omega)\,,$$

where $|\Omega| = \dim(\mathcal{H})$, and where we used the notations $\rho = \sum_\omega p(\omega)|\omega\rangle\langle\omega|$, $A = \sum_\omega f(\omega)|\omega\rangle\langle\omega|$, $B = \sum_\omega g(\omega)|\omega\rangle\langle\omega|$.

Finally, to each monotone Riemannian $g_f^{(m)}$, one can associate a geodesic distance $d_f^{(m)}$, which is defined as

$$d_f^{(m)}(\rho, \sigma) := \inf_\gamma \int_0^1 \|\dot\gamma(s)\|_{g_{f,\gamma(s)}^{(m)}}\,ds,$$

where the infimum is taken over all smooth paths $\gamma : [0, 1] \to \mathcal{D}_+(\mathcal{H})$, with $\gamma(0) = \rho$ and $\gamma(1) = \sigma$, and $\|.\|_{g_f^{(m)}}$ is the metric norm defined as

$$\|A\|_{g_{f,\rho}^{(m)}} := g_{f,\rho}^{(m)}(A, A)^{1/2}, \qquad A \in \mathcal{T}^{(m)}\mathcal{D}_+(\mathcal{H})\,.$$

Properties of the Riemannian metrics $d_f^{(m)}$, as well as examples of Fisher information metrics can be found in [Lesniewski and Ruskai, 1999]. For more details on geodesics and parallel transport in the mixture and exponential representations, see Section 6.3 of [Hayashi, 2016]. We close this section by the following useful partial ordering of inner products provided in [Lesniewski and Ruskai, 1999] (see also [Kubo, 1979, Hiai and Kosaki, 1999, Perez-Garcia et al., 2006, Hiai et al., 2013, Temme et al.,

2010]): first note that the following holds for any $s \in [0,1]$ and $x \in \mathbb{R}_+$:

$$\frac{2}{x+1} \leq \frac{1+s}{2}\left(\frac{1}{s+x} + \frac{1}{sx+1}\right) \leq \frac{x+1}{2x}.$$

Moreover, any $k \in \mathrm{OMM}$ admits the following integral representation [Lesniewski and Ruskai, 1999]:

$$k(x) = \int_0^1 \left(\frac{1}{s+x} + \frac{1}{sx+1}\right)\sigma_k(s)\,ds \tag{1.87}$$

where σ_k is normalized in such a way that $t \mapsto 2\sigma_k(t)/(t+1)$ is a probability density function on $[0,1]$. Therefore, by the above integral representation, for any operator $T > 0$

$$2(R_T + L_T)^{-1} \leq R_T^{-1}\,k(\Delta_T) \leq \frac{R_T^{-1} + L_T^{-1}}{2}, \tag{1.88}$$

where R_T, resp. L_T, denotes the operation of left, resp. right, multiplication by T, and $\Delta_T(.) := T(.)T^{-1}$. Therefore,

$$[\rho]_{f_{\mathrm{Bures}}}^{-1} \equiv 2\left(R_\rho + L_\rho\right)^{-1} \leq [\rho]_f^{-1} \leq \frac{L_\rho^{-1} + R_\rho^{-1}}{2}, \tag{1.89}$$

where $f_{\mathrm{Bures}} := 2(1+x)^{-1}$. The metric associated to f_{Bures} is also known as the *Bures* metric in the quantum information geometry community. If follows directly that for all $1/f \in \mathrm{OMM}$:

$$d_f^{(m)}(\rho,\sigma) \geq d_{f_{\mathrm{Bures}}}^{(m)}(\rho,\sigma),$$

so that $d_{f_{\mathrm{Bures}}}^{(m)}$ gives the minimal monotone geodesic distance. The corresponding Fisher information for a parametrized family of states $(\rho_\theta)_{\theta \in \mathbb{R}}$ is given by:

$$\mathrm{I}_{\mathrm{SLD}}(\theta) := g_{f_{\mathrm{Bures}},\rho_\theta}^{(m)}\left(\frac{d\rho_\theta}{d\theta}, \frac{d\rho_\theta}{d\theta}\right) = \mathrm{Tr}(\rho_\theta(L_\theta^{\mathrm{SLD}})^2),$$

where L_θ^{SLD} is the so-called *symmetric logarithmic derivative* and is characterized by $\frac{d}{d\theta}\rho_\theta = \frac{1}{2}(\rho_\theta L_\theta^{\mathrm{SLD}} + L_\theta^{\mathrm{SLD}}\rho_\theta)$.

The SLD-*Fisher information* has been shown to satisfy the following quantum version of the Cramér-Rao bound (see [Helstrom, 1967, Nagaoka, 2005]). A POVM $\{M(B), B \in \mathcal{B}(\mathbb{R})\}$ is said to be *unbiased* if for all $\theta_0 \in \mathbb{R}$:

$$\int_{\theta \in \mathbb{R}} \theta\,\mathrm{Tr}(\rho_{\theta_0} M(d\theta)) = \theta_0.$$

Theorem 1.3.5. *For any unbiased POVM $\{M(B), B \in \mathcal{B}(\mathbb{R})\}$, the following bound holds:*

$$\int_{\theta \in \mathbb{R}} (\theta - \theta_0)^2\,\mathrm{Tr}(\rho_{\theta_0} M(d\theta)) \geq \frac{1}{\mathrm{I}_{\mathrm{SLD}}(\theta_0)}.$$

Part II.

Classical functional inequalities

Chapter 2.

Markov chains and processes

Part II is dedicated to a review of the theory of classical functional inequalities in continuous and discrete settings and their use in the context of convergence of Markovian evolutions towards their asymptotic regimes. There are two main reasons for introducing these classical notions. First, we hope that a review of the subject leads to a better understanding of the possible ways of extending these concepts to the quantum realm. Second, some of the more advanced classical convergence results reviewed here will prove to be very useful in establishing convergence properties of quantum evolutions via the so-called transference method explained in Chapter 9.

Layout of the chapter: In this chapter, we start by introducing the basic definitions and properties of the theory of classical Markov processes. This is mainly to fix the notations that we will use in the subsequent chapters. In Section 2.1, we recall the notion of a Markov semigroup on the algebra of bounded measurable functions. This notion is then extended to the case of Markov semigroups acting on $\mathbb{L}_p$ spaces in Section 2.2. There, we also introduce the notions of invariant measures and reversibility, as well as Dirichlet forms. The latter will prove useful in order to derive mixing times from the functional inequalities described in Chapter 3. Finally Sections 2.3 and 2.4 are devoted to the introduction of the two canonical classes of classical Markovian evolutions, namely diffusion processes and Markov chains. We refer to [Bakry et al., 2014, Applebaum, 2004] for more details.

2.1. Markov processes in continuous time

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with associated filtration $(\mathcal{F}_t)_{t \geq 0}$ and $(E, \mathcal{G})$ a measurable space. A continuous adapted process $(X_t)_{t \geq 0}$ on E is called a *Markov process* in the filtration $(\mathcal{F}_t)_{t \geq 0}$ if for all $t \geq 0$, $\mathcal{F}_t$ and the sigma-algebra $\sigma(X_s, s \geq t)$ generated by $(X_s)_{s \geq t}$ are *conditionally independent*, i.e.

$$\mathbb{P}(A \cap B | X_t) = \mathbb{P}(A | X_t) \, \mathbb{P}(B | X_t) \tag{2.1}$$

for all $A \in \mathcal{F}_t$ and $B \in \sigma(X_s, s \geq t)$. Supposing $\mathcal{F}_t := \sigma(X_s, s \leq t)$, the above identity means that given the present, X_t, the past $\sigma(X_s, s \leq t)$ and the future $\sigma(X_s, s \geq t)$ are independent. This definition is equivalent to the following identity holding true:

$$\mathbb{E}\left[f(X_{u+t}) \, | \, \mathcal{F}_t \right] = \mathbb{E}\left[f(X_{u+t}) \, | \, X_t \right], \tag{2.2}$$

for all bounded measurable functions $f : E \to \mathbb{R}$ and $u, t \geq 0$.

   Next, for any $0 \leq t \leq s < \infty$, define the operator $P_{t,s}$ on the space of bounded measurable

functions by the following expression:

$$P_{t,s}(f)(x) := \mathbb{E}[f(X_s)|X_t = x], \qquad f \in \mathbb{L}_\infty(E, \mathcal{G}). \tag{2.3}$$

The family of operators $P_{t,s}$ is called a *Markov evolution* and satisfies the following properties:

- $P_{t,s}$ is a linear operator on $\mathbb{L}_\infty(E, \mathcal{G})$ for each $0 \le s \le t < \infty$;

- $P_{s,s} = \mathrm{id}$;

- $P_{r,t} \circ P_{t,s} = P_{r,s}$ whenever $0 \le r \le t \le s < \infty$;

- $f \ge 0 \Rightarrow P_{t,s}(f) \ge 0$ for all $0 \le t \le s < \infty$;

- $P_{t,s}$ is a contraction, i.e. $\|P_{t,s}\|_\infty \le 1$, for each $0 \le t \le s < \infty$;

- $P_{t,s}(\mathbb{1}_E) = \mathbb{1}_E$ for all $0 \le t \le s < \infty$.

From this Markov evolution, define the following family of maps on $E \times \mathcal{G}$:

$$\tilde{P}_{t,s}(x, A) := (P_{t,s}(\mathbb{1}_A))(x) \tag{2.4}$$

for $x \in E$, $A \in \mathcal{G}$. These maps are called *Markov transition probabilities* (or *probability kernels*) since, for each x, t, s, $P_{t,s}(x, .)$ defines a probability measure on $(E, \mathcal{G})$. Moreover, one can reverse Equation (2.4) and get for each bounded measurable function f and $x \in E$:

$$P_{t,s}(f)(x) = \int f(y)\, \tilde{P}_{t,s}(x, dy).$$

Markov processes are essentially characterized by their initial laws and transition probabilities. This statement is true if one imposes some further conditions: The Markov process is said to be *normal* if for any $0 \le t \le s < \infty$, and each $A \in \mathcal{G}$, the mapping

$$x \mapsto \tilde{P}_{t,s}(x, A)$$

is measurable. In this case, the following equation holds

$$\tilde{P}_{t,u}(x, A) = \int \tilde{P}_{t,s}(x, dy)\, \tilde{P}_{s,u}(y, A) \quad \forall x \in A. \tag{2.5}$$

This equation is called the *Chapman-Kologorov equation*. Calling μ the *initial distribution* of $(X_t)_{t \ge 0}$, so that

$$\forall A \in \mathcal{G}, \qquad \mu(A) = \mathbb{P}(X_0 \in A),$$

one can easily verify that for any bounded measurable function $f : E^n \to \mathbb{R}$,

$$\mathbb{E}[f(X_{t_1}, ..., X_{t_n})] = \int \mu(dx_0) \int \tilde{P}_{0,t_1}(x_0, dx_1) ... \int \tilde{P}_{t_{n-1},t_n}(x_{n-1}, dx_n) f(x_1, ..., x_n), \tag{2.6}$$

which implies that the finite-dimensional distributions can be are expressed in terms of the transition probabilities and the initial distribution of the process.

Let us conversely define a *normal Markov transition mapping* to be a two-parameters family of mappings $\{\tilde{P}_{t,s}, \ 0 \le t < s < \infty\}$ from $E \times \mathcal{G}$ to $[0, 1]$ such that:

- $A \mapsto \tilde{P}_{t,s}(x, A)$ is a probability measure on $\mathcal{G}$ for all $t \le s$ and $x \in E$;

- $x \mapsto \tilde{P}_{t,s}(x, A)$ is measurable for all $A \in \mathcal{G}$, $t \le s$;

- If $0 \le t < s < u$, the Chapman-Kolmogorov Equation (2.5) holds for $\tilde{P}$.

One can prove (see e.g. Theorem 3.1.7 of [Applebaum, 2004] in the case $E = \mathbb{R}^n$) via Daniell-Kolmogorov's consistency theorem that, for any probability measure μ on $(E, \mathcal{G})$ and any normal Markov transition mapping $\tilde{P}_{t,s}$ on $\mathbb{R}^n$, there exists a Markov process on E with associated transition function $\tilde{P}_{t,s}$ and initial distribution μ having finite-dimensional distributions given by Equation (2.6). The process constructed in this way is therefore a normal Markov process called the *canonical Markov process*. This situation does not easily extend to the quantum case since quantum Markov semigroups can admit either multiple, or one, or no associated process. We come back to this point in Section 5.5.3.

Back to the commutative setting, not all Markov processes are normal. Roughly speaking, being normal for a Markov process means that it is possible to define nice conditional probability distributions for the conditional probabilities $\mathbb{P}(X_s \in A | \mathcal{F}_t), 0 \le t < s, \ A \in \mathcal{G}$. A Markov process may even have more than one transition function, but in this case, calling P^1 and P^2 two such families, one has $P^1_{t,s}(x, A) = P^2_{t,s}(x, A)$ for almost all x, relative to the distribution of X_t.

In what follows, we always assume that the Markov processes we manipulate are normal. This is for instance the case of Feller processes defined on a locally compact, separable metric space (see [Kallenberg, 2006, Applebaum, 2004]). Moreover, we restrict our attention to the subclass of a homogeneous (quantum) Markov process. A *homogeneous Markov process* is one whose transition probabilities are time homogeneous, which means that there exists a function $\tilde{P}_t(x, A)$ defined for all $t > 0$ such that $\tilde{P}_{t,s}(x, A) = \tilde{P}_{s-t}(x, A)$ for all $0 \le t \le s < \infty$, $x \in E$, $A \in \mathcal{E}$. In this case, the Chapman-Kolmogorov equation reduces to:

$$P_{t+s}(x, A) = \int P_t(x, dy) P_s(y, A). \tag{2.7}$$

This implies the existence of linear maps P_t on $\mathbb{L}_\infty(E, \mathcal{E})$ such that for any $f \in \mathbb{L}_\infty(E, \mathcal{E})$ and all $x \in E$:

$$P_t(f)(x) := P_{0,t}(f)(x) = \int \tilde{P}_t(x, dy) f(y).$$

These maps define a semigroup via the Chapman-Kolmogorov equation. This means that the maps $(P_t)_{t \ge 0}$ satisfy the following properties

- *Semigroup property*: $P_t \circ P_s = P_{t+s}$ whenever $0 \le t, s < \infty$;

- *Positivity*: $\quad f \ge 0 \ \Rightarrow \ P_t(f) \ge 0$ for all $0 \le t < \infty$ and $f \in \mathbb{L}_\infty(E, \mathcal{E})$;

- *Unitality*: $P_t(\mathbb{1}_E) = \mathbb{1}_E$ for all $0 \le t < \infty$.

Such a semigroup is simply called a *Markov semigroup*. From this definition it directly follows that $(P_t)_{t \ge 0}$ is a *contraction semigroup*, i.e. for each time $t \ge 0$,

$$|P_t(f)| \le P_t(|f|) \le P_t(\mathbb{1}_E) \|f\|_\infty = \|f\|_{\mathbb{L}_\infty(E, \mathcal{G})} \quad \Rightarrow \quad \|P_t : \mathbb{L}_\infty(E, \mathcal{G}) \to \mathbb{L}_\infty(E, \mathcal{G})\| \le 1. \tag{2.8}$$

Later on, we will mostly focus on the case when the semigroup $(P_t)_{t \ge 0}$ admits a *density kernel* with respect to some reference σ-finite measure λ on $\mathcal{G}$. This means that there exists for every $t > 0$ a positive measurable function $(x, y) \mapsto p_t(x, y)$ defined $\lambda \otimes \lambda$-almost everywhere on $E \times E$ such that, for every bounded or positive measurable function $f : E \to \mathbb{R}$ and λ-almost every $x \in E$:

$$P_t(f)(x) = \int_E k_t(x, y) f(y) \, d\lambda(y). \tag{2.9}$$

In this case, the unitality ot the semigroup imposes that $\int_E k_t(x, y) \, d\lambda(y) = 1$ λ-almost everywhere. The following sufficient condition for the existence of such a density kernel can be found for instance in Proposition 1.2.5 of [Bakry et al., 2014]: there exists $M > 0$ such that for any $t \geq 0$:

$$\| P_t : \mathbb{L}_1(\lambda) \to \mathbb{L}_\infty(\lambda) \| \leq M \,,$$

and in this case, the kernels k_t are $\lambda \otimes \lambda$-almost everywhere bounded by M.

2.2. $\mathbb{L}_p$-Markov semigroups, reversibility and Dirichlet forms

$\mathbb{L}_p$ **Markov semigroups and invariant measures** Let $(E, \mathcal{G}, \mu)$ be a probability space, $(P_t)_{t \geq 0}$ a Markov semigroup on the set of bounded measurable functions on E, and assume that μ is an *invariant measure* with respect to $(P_t)_{t \geq 0}$: for any bounded positive measurable function $f : E \to \mathbb{R}$ and any $t \geq 0$:

$$\mathbb{E}_\mu[P_t(f)] = \mathbb{E}_\mu[f] \,. \tag{2.10}$$

The semigroup $(P_t)_{t \geq 0}$ can be extended by density to the sets $\mathbb{L}_p(\mu)$, $p \geq 1$, in such a way that for any $t \geq 0$:

$$\| P_t : \mathbb{L}_p(\mu) \to \mathbb{L}_p(\mu) \| \leq 1 \,.$$

In this case, Equation (2.10) also holds for any function $f \in \mathbb{L}_1(\mu)$. More generally, a semigroup of operators $(P_t)_{t \geq 0}$ on $\mathbb{L}_p(\mu)$ is called an $\mathbb{L}_p(\mu)$-*Markov semigroup* if it satisfies the conditions of unitality and positivity on the subalgebra $\mathbb{L}_\infty(\mu)$, together with the following property:

$$\forall f \in \mathbb{L}_p(\mu), \qquad \| P_t(f) - f \|_{\mathbb{L}_p(\mu)} \underset{t \to 0}{\to} 0 \,, \tag{2.11}$$

i.e. $(P_t)_{t \geq 0}$ is strongly continuous with respect to the (real or complex) Banach space $\mathbb{L}_p(\mu)$.

Let now $(P_t)_{t \geq 0}$ be an $\mathbb{L}_2(\mu)$-Markov semigroup. We know from the Hille-Yosida theorem that it is fully determined by its (possibly unbounded) generator $(L_2, \mathrm{dom}(L_2))$:

$$L_2 : \begin{cases} \mathrm{dom}(L_2) \to \mathbb{L}_2(\mu) \\ f \mapsto \lim_{\epsilon \to 0} \dfrac{P_\epsilon(f) - f}{\epsilon} \end{cases} \tag{2.12}$$

the limit being taken in the $\|.\|_{\mathbb{L}_2(\mu)}$-topology whenever it exists. However, in practice we only know L_2 on a dense subset of its domain. From now on we therefore assume that the following hypothesis holds true:

Condition 2.2.1. There exists a dense subclass $\mathcal{A} \subset \mathrm{dom}(L_2) \cap \bigcap_{p \geq 1} \mathbb{L}_p(\mu)$ that is a *core* for L_2: for any $f \in \mathrm{dom}(L_2)$, there exists a sequence $\{f_k\}_{k \in \mathbb{N}}$ converging to f such that $\{L_2(f_k)\}_{k \in \mathbb{N}}$ converges to $L_2(f)$. We also assume that $\mathcal{A}$ is stable under composition with multivariable smooth functions and that $L_2(\mathcal{A}) \subset \mathcal{A}$. Since μ is finite, we can also assume that $\mathcal{A}$ contains constant functions. Finally, for any $f, g \in \mathcal{A}$, $P_t(f)g \in \mathcal{A}$ (ideal property).

The semigroup $(P_t)_{t \geq 0}$ is said to be *reversible* for the probability measure μ if for any $f, g \in \mathbb{L}_2(\mu)$:

$$\langle P_t(f), g \rangle_{\mathbb{L}_2(\mu)} = \langle f, P_t(g) \rangle_{\mathbb{L}_2(\mu)} \,. \tag{μ-DBC}$$

One equivalently says that $(P_t)_{t\geq 0}$ is *symmetric* with respect to μ, or that the semigroups satisfies the *detailed balance condition* (μ-DBC) with respect to μ. Such a measure μ is necessarily invariant.

Following the exact same lines as in Equation (2.8), one shows that $(P_t)_{t\geq 0}$ is a contraction in $\mathbb{L}_\infty(\mu)$. Similarly for $p \in [1, \infty)$, any $\mathbb{L}_2(\mu)$-Markov semigroup $(P_t)_{t\geq 0}$ with invariant measure μ is also $\mathbb{L}_p(\mu)$ contractive:

$$\forall f \in \mathbb{L}_p(\mu), \qquad \|P_t f\|_{\mathbb{L}_p(\mu)} \leq \|f\|_{\mathbb{L}_p(\mu)}. \tag{2.13}$$

For $p = 1$, this follows by integrating the following inequality against the measure μ:

$$|P_t(f)| = |P_t(f_+ - f_-)| \leq |P_t(f_+)| + |P_t(f_-)| = P_t(f_+ + f_-) = P_t(|f|). \tag{2.14}$$

The result for any $1 \leq p \leq \infty$ follows from Riezs Thorin interpolation for commutative $\mathbb{L}_p$ spaces.

Invariant and reversible measures can be equivalently characterized by the following two conditions: μ is invariant if and only if for all $f \in \mathcal{A}$,

$$\mathbb{E}_\mu [L_2(f)] = 0. \tag{2.15}$$

Moreover, μ is reversible if and only if for any $f, g \in \mathcal{A}$

$$\langle f, L_2(g) \rangle_{\mathbb{L}_2(\mu)} = \langle L_2(f), g \rangle_{\mathbb{L}_2(\mu)}. \tag{2.16}$$

Carré du champ operator and Dirichlet forms Given some probability measure μ, let $(P_t)_{t\geq 0}$ be an $\mathbb{L}_2(\mu)$-Markov semigroup of infinitesimal generator L_2. Suppose moreover given an algebra $\mathcal{A}$ satisfying Condition 2.2.1. The *carré du champ* operator is defined for any $f, g \in \mathcal{A}$ as follows:

$$\Gamma(f, g) := \frac{1}{2} \left(L_2(fg) - f L_2(g) - L_2(f) g \right). \tag{2.17}$$

The Carré du champ operator is used to define a notion of distance and diameter associated to the operator L_2 in the abstract measurable space E as follows:

$$d_{L_2}(x, y) := \sup_{g \in \mathcal{A}, \, \Gamma(g,g) \leq 1} |g(x) - g(y)|, \qquad \mathrm{diam}(E) = \mathrm{essup}_{E \times E} \, d_{L_2}(x, y) \leq \infty. \tag{2.18}$$

One can actually further formally define a Riemannian structure associated to L_2 on E and its associated Ricci curvature. We will come back to this idea in Section 3.6. For the time being, we make the following important observations: First, one easily verifies that the carré du champ is always non-negative:

$$\forall f \in \mathcal{A}, \quad \Gamma(f, f) \geq 0. \tag{2.19}$$

The first consequence of (2.19) is that Γ can be extended to functions in $\mathrm{dom}(L_2)$ by a simple Cauchy sequence argument (see [Bakry, 1994]). Moreover, (2.19) implies that whenever μ is invariant,

$$\forall f \in \mathrm{dom}(L_2), \quad \mathcal{E}(f, f) := -\langle f, L_2(f) \rangle_{\mathbb{L}_2(\mu)} = \mathbb{E}_\mu(\Gamma(f, f)) \geq 0. \tag{2.20}$$

The mapping $\mathcal{E}(f, g) = -\langle f, L_2(g)\rangle_{\mathbb{L}_2(\mu)}$, called the *Dirichlet form* associated to $(P_t)_{t \geq 0}$, will play a major role in the rest of this thesis. For now, we observe that, whenever (2.20) holds,

$$0 \leq \mathcal{E}(f, f) = -\langle f, L_2(f)\rangle_{\mathbb{L}_2(\mu)} = -\langle \hat{L}_2(f), f\rangle_{\mathbb{L}_2(\mu)} = -\overline{\langle f, \hat{L}_2(f)\rangle_{\mathbb{L}_2(\mu)}} = -\langle f, \hat{L}_2(f)\rangle_{\mathbb{L}_2(\mu)},$$

where $\hat{L}_2$ denotes the adjoint of L_2 in $\mathbb{L}_2(\mu)$, and the last identity comes from the fact that $\mathcal{E}(f, f)$ is real. In particular

$$\mathcal{E}(f, f) = -\frac{1}{2}\langle f, (L_2 + \hat{L}_2)(f)\rangle_{\mathbb{L}_2(\mu)}. \tag{2.21}$$

Now, if $(P_t)_{t \geq 0}$ further satisfies μ-DBC, this means that it is symmetric in $\mathbb{L}_2(\mu)$, and so is its generator (see e.g. [Hille and Phillips, 1996]). Therefore, the following *integration by parts* (IBP) formula holds: for any $f, g \in \text{dom}(L_2)$:

$$\mathcal{E}(f, g) = -\langle f, L_2(g)\rangle_{\mathbb{L}_2(\mu)} = \mathbb{E}_\mu(\Gamma(f, g)) = \mathbb{E}_\mu(\Gamma(g, f)) = \mathcal{E}(g, f). \tag{IBP}$$

The reason behind this designation will become clear with the examples treated in Section 2.3. Such a self-adjoint operator admits a spectral decomposition:

$$L_2 = \int_{\mathbb{R}} \lambda \, dE_\lambda, \tag{2.22}$$

where E_λ is a resolution of the identity acting on $\mathbb{L}_2(\mu)$. By (2.20), we further know that L_2's spectrum is non-positive, which means that the integration in Equation (2.22) can be further reduced to

$$L_2 = -\int_0^\infty \lambda \, dE_\lambda \qquad P_t = \int_0^\infty \mathrm{e}^{-\lambda t} \, dE_\lambda. \tag{2.23}$$

This formula shows that when t goes to infinity, $P_t(f)$ converges in $\mathbb{L}_2(\mu)$ towards the projection $E_0(f)$ of f on the eigenspace associated to the eigenvalue 0 of L_2. We call this space the space of *fixed points* of $(P_t)_{t \geq 0}$ and denote it by $\mathcal{F}(P)$. Note that for a Markov semigroup, $\mathcal{F}(P)$ always includes the set of constant functions $\mathbb{C}\,\mathbb{1}_E$. In the case of equality ($\mathcal{F}(P) = \mathbb{C}\,\mathbb{1}_E$), the projection E_0 is given by

$$\forall f \in \mathbb{L}_2(\mu), \quad E_0(f) = \mathbb{E}_\mu[f] \quad \Rightarrow \quad \|P_t(f) - \mathbb{E}_\mu[f]\|_{\mathbb{L}_2(\mu)} \underset{t \to \infty}{\to} 0.$$

We end this chapter by introducing the two main classes of examples that we will encounter, namely diffusion processes and Markov chains.

2.3. Diffusion processes

Here, we introduce a subclass of the class of Markov processes, called Markov diffusion processes, and refer to [Bakry, 1994, Bakry et al., 2014] for further details. Given a probability measure μ, an $\mathbb{L}_2(\mu)$-Markov semigroup $(P_t)_{t \geq 0}$ is said to be a *diffusion semigroup* if for any $f_1, ..., f_k \in \mathcal{A}$, and any smooth function $\Phi : \mathbb{R}^k \to \mathbb{R}$, one has

$$L_2\Phi(f_1, ..., f_k) = \sum_i \frac{\partial \Phi}{\partial x_i}(f_1, ..., f_k) \, L_2(f_i) + \sum_{ij} \frac{\partial^2 \Phi}{\partial x_i \partial x_j}(f_1, ..., f_k) \, \Gamma(f_i, f_j). \tag{2.24}$$

This definition simply means L_2 is a second order differential operator on the algebra $\mathcal{A}$ with no constant terms. The diffusion property implies in particular that the underlying Markov process $(X_t)_{\geq 0}$ is continuous in the sense that for any $f \in \mathcal{A}$, $t \mapsto f(X_t)$ is continuous (see [Bakry, 1989]).

Now, applying Equation (2.24) to $\Phi(f,g,h) = fgh$, we get the following *Leibniz rule*:

$$\Gamma(fg,h) = f\,\Gamma(g,h) + \Gamma(f,h)\,g\,. \tag{2.25}$$

Conversely, this property of Γ enables to establish Equation (2.24) for polynomials. More generally for a semigroup of diffusion, the following *chain rule* is verified:

$$\Gamma(\Phi(f),g) = \Phi'(f)\,\Gamma(f,g)\,. \tag{chain rule}$$

For example, consider the following canonical example: assume E is a smooth Riemannian manifold $(\mathcal{M}, g)$, let μ be a probability measure that is absolutely continuous with respect to the Lebesgue measure on $\mathcal{M}$, and consider the following expression, written in a local coordinate system, for the generator L on the space $\mathcal{A} := C^{\infty}(\mathcal{M})$ of smooth functions on $\mathcal{M}$:

$$L(f)(x) = \sum_{ij} g^{ij}(x)\frac{\partial^2 f}{\partial x_i\,\partial x_j} + \sum_i b_i(x)\frac{\partial f}{\partial x_i}\,. \tag{2.26}$$

In this case, one can readily check that

$$\Gamma(f,h) = \sum_{ij} g^{ij}\frac{\partial f}{\partial x_i}\frac{\partial h}{\partial x_j}\,. \tag{2.27}$$

If we further assume μ-DBC, IBP becomes

$$\mathcal{E}(f,h) = -\int_{\mathcal{M}} f(x)\left(\sum_{ij} g^{ij}(x)\frac{\partial^2 h}{\partial x_i\,\partial x_j} + \sum_i b_i(x)\frac{\partial h}{\partial x_i}\right) d\mu(x) = \int_{\mathcal{M}} \sum_{ij} g^{ij}(x)\frac{\partial f}{\partial x_i}\frac{\partial h}{\partial x_j}\,d\mu(x)\,.$$

In the simpler case when $\mathcal{M}$ is compact, μ is the normalized Lebesgue measure on $\mathcal{M}$, and $L = \Delta$ is the Laplace-Beltrami operator, the above identity simplifies to:

$$\mathcal{E}(f,h) = -\int_{\mathcal{M}} f(x)\,\Delta\,h(x)\,dx = \int_{\mathcal{M}} \sum_{ij} g^{ij}\frac{\partial f}{\partial x_i}\frac{\partial h}{\partial x_j}\,dx\,,$$

and we recover the usual notion of integration by parts. Moreover, the chain rule implies that for all non-negative smooth function f:

$$\mathcal{E}(\ln f, f) = -\int_{\mathcal{M}} \ln f\,\Delta f d\mu = \int_{\mathcal{M}} \frac{|\nabla f|^2}{f} = 4\,\mathcal{E}(\sqrt{f},\sqrt{f})\,, \tag{2.28}$$

where $|\nabla f|^2 := \sum_{ij} g^{ij}(x)\frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}$. As we will see in Section 3.3, the term $\mathcal{E}(f,\ln f)$ is related to the decrease of the relative entropy between a probability measure evolving according to the semigroup dual to $(P_t)_{t\geq 0}$ and the invariant measure μ, and is for this reason called the *entropy production* of $(P_t)_{t\geq 0}$. The last term is the multivariate version of the Fisher information associated to a family of translations of a probability density function, as introduced in Equation (1.71). One can similarly verify that symmetric diffusion semigroups satisfy the following: for any $p > 1$ and all nonnegative $f \in \mathcal{A}$:

$$\mathcal{E}(f^{p-1}, f) = \frac{4(p-1)}{p^2}\,\mathcal{E}(f^{p/2}, f^{p/2})\,. \tag{2.29}$$

As discussed in the next section, these identities break down for Markov chains on discrete sample spaces, which makes the analysis of their contractivity properties in some sense more difficult.

2.4. Markov chains on finite sample spaces

The other important example that we will consider is when $(E, \mathcal{G}, \mu)$ is a finite probability space, with μ being strictly positive and invariant. We denote the probability mass function corresponding to μ by $p : E \to [0, 1]$. In this case, for any semigroup $(P_t)_{t \geq 0}$, $\mathcal{A}$ is the $|E|$-dimensional algebra of all the functions defined on E, and the generator L of $(P_t)_{t \geq 0}$ can be represented by a matrix $L = (L_{ij})_{(i,j) \in E \times E}$, so that for all $t \geq 0$

$$P_t = \mathrm{e}^{t\,L}\,.$$

The unitality of the semigroup is equivalent to the fact that for all $i \in E$, $\sum_j L_{ij} = 0$, whereas its positivity is equivalent to the fact that for all $i \neq j$, $L_{ij} \geq 0$. Typically, a Markov chain $(P_t)_{t \geq 0}$ can be derived from a discrete time Markov kernel $\{K(i,j)\}_{i,j \in E}$, so that $L_{ij} = K(i,j) - \delta_{ij}$.

Choosing μ as the measure or reference, we can define a density kernel $(k_t)_{t \geq 0}$ on $E \times E$ as follows: for any $t \geq 0$,

$$k_t(i,j) = \frac{\mathrm{e}^{tL}(i,j)}{p(j)} \quad \Rightarrow \quad P_t(f)(i) = \sum_j k_t(i,j)\, f(j)\, p(j)\,. \tag{2.30}$$

In this setting, the carré du champ operator takes the form

$$\Gamma(f,h)(i) = \frac{1}{2} \sum_j L_{ij} \left(f(i) - f(j)\right) \left(h(i) - h(j)\right),$$

and the Dirichlet form is defined as:

$$\mathcal{E}(f,h) = -\sum_{ij} f(i)\, L_{ij}\, h(j)\, p(i) = \frac{1}{2} \sum_{ij} L_{ij} \left(f(i) - f(j)\right) \left(h(i) - h(j)\right) p(i)\,,$$

where the second identity holds when $(P_t)_{t \geq 0}$ is reversible with respect to μ. Here, (2.28) no longer holds. However, one can show the weaker statements (see Lemma 2.7 of [Diaconis and Saloff-Coste, 1996a]):[1]

Lemma 2.4.1. *For any Markov chain $(P_t)_{t \geq 0}$ with invariant measure μ of corresponding probability mass function p, and any $f \geq 0$:*

$$\mathcal{E}(\ln f, f) \geq 2\,\mathcal{E}(\sqrt{f}, \sqrt{f})\,. \tag{2.31}$$

Furthermore, if the chain is reversible with respect to μ,

$$\mathcal{E}(\ln f, f) \geq 4\,\mathcal{E}(\sqrt{f}, \sqrt{f})\,. \tag{2.32}$$

In fact, this last inequality can be seen as the limit as $p \to 1$ of the the second family of inequalities listed below (see Lemma 2.6 of [Diaconis and Saloff-Coste, 1996a]):

Lemma 2.4.2. *With the notations of Lemma 2.4.1, for any $f \geq 0$, and $p \geq 2$:*

$$\mathcal{E}(f, f^{p-1}) \geq \frac{2}{p}\,\mathcal{E}(f^{p/2}, f^{p/2})\,. \qquad \text{(Weak } \mathbb{L}_p\text{-regularity)}$$

---

[1]In fact, Lemmas 2.4.1 and 2.4.2 can be extended to the general setting of Section 2.2. For symmetric semigroups, this is done in Proposition 3.1 of [Bakry, 1994].

If $(P_t)_{t\geq0}$ *is symmetric with respect to* μ, *then for any* $p > 1$

$$\mathcal{E}(f, f^{p-1}) \geq \frac{4(p-1)}{p^2} \mathcal{E}(f^{p/2}, f^{p/2}). \qquad \text{(Strong } \mathbb{L}_p\text{-regularity)}$$

The above inequalities play an important role in establishing the hypercontractivity of a given Markov semigroup (see Chapter 3). They were recently partially extended to the quantum setting in [Bardet, 2017, Beigi et al., 2018] (see Section 5.4).

2.5. One-parameter convolution semigroups on $\mathrm{Aut}(\mathcal{B}(\mathcal{H}))$

We end this chapter by introducing a particular type of Markov semigroup when the underlying space E is a group. These will play a major role in Chapter 9. Let G be a locally compact separable group, and let $\mathcal{M}(G)$ be the space of all probability measures on $(G, \mathcal{B}(G))$, where $\mathcal{B}(G)$ denotes the associated Borel σ-algebra. The space $\mathcal{M}(G)$ can be given the structure of a semigroup by defining the *convolution* $\mu * \nu$ of two elements $\mu, \nu \in \mathcal{M}(G)$ as follows: for any Borel set $A \in \mathcal{B}(G)$:

$$\mu * \nu(A) = \int_G \mu(Ag^{-1}) \nu(dg) = \int_G \mu(g^{-1}A) \nu(dg), \quad A \in \mathcal{B}(G),$$

where, for any $g \in G$ and $A \in \mathcal{B}(G)$, $gA = \{g\,h,\ h \in G\}$ and $Ag = \{h\,g,\ h \in G\}$. The unit element is given by the measure μ_e concentrated at the identity element e. Next, a *one-parameter convolution semigroup* of a Markov process on G is a family $(\mu_t)_{t\geq0}$ of measures from $\mathcal{M}(G)$ such that μ_t converges weakly to μ_e as $t \to 0$, and for all $t, s \geq 0$,

$$\mu_{t+s} = \mu_t * \mu_s.$$

The following theorem due to [Hunt, 1956] provides a characterization of the generator of one-parameter convolution semigroups when $G = \mathrm{Aut}(\mathcal{B}(\mathcal{H}))$ stands for the group of automorphisms on the algebra $\mathcal{B}(\mathcal{H})$, where $\mathcal{H}$ is finite dimensional:

Theorem 2.5.1. *Let* $(\mu_t)_{t\geq0}$ *be a weak* continuous semigroup of probability measures on* $\mathrm{Aut}(\mathcal{B}(\mathcal{H}))$. *Then there exist real numbers* c_i, $i \in \{1, ..., n^2 - 1\}$, *a positive semidefinite symmetric* $(n^2 - 1) \times (n^2 - 1)$ *matrix* $B = b_{ij}$ *and a Lévy measure* ν *such that for all* $f\ C_2(\mathrm{Aut}(\mathcal{B}(\mathcal{H})))$, *the derivative* $\frac{d}{dt}\mu_t(f)|_{t=0}$ *exists and is given by*

$$\frac{d}{dt}\mu_t(f)\Big|_{t=0} = \sum_{i=1}^{n^2-1} c_i\, D_i f(\mathrm{id}) + \sum_{ij=1}^{n^2-1} b_{ij}\, D_i D_j\, f(\mathrm{id})$$

$$+ \int_{\mathrm{Aut}(\mathcal{B}(\mathcal{H}))} \left(f(\alpha) - f(\mathrm{id}) - \sum_{i=1}^{n^2-1} g_i(\alpha)\, D_i f(\mathrm{id}) \right) d\nu(\alpha). \qquad (2.33)$$

Conversely, if $\{c_i\}$, $\{b_{ij}\}$ *and* ν *satisfy the above conditions, then Equation (2.33) determines exactly one convolution semigroup* $(\mu_t)_{t\geq0}$ *on* $\mathrm{Aut}(\mathcal{B}(\mathcal{H}))$.

Chapter 3.

Functional inequalities

In Chapter 2, we saw that a reversible Markov semigroup $(P_t)_{t\geq 0}$ on a probability space $(E,\mathcal{G},\mu)$ whose space of fixed points $\mathcal{F}(P)$ equals $\mathbb{C}\,\mathbb{1}_E$ converges in $\mathbb{L}_2(\mu)$: for any $f \in \mathbb{L}_2(\mu)$,

$$P_t(f) \underset{t\to\infty}{\to} \mathbb{E}_\mu[f]. \tag{3.1}$$

More generally, we say that an $\mathbb{L}_2(\mu)$-Markov semigroup $(P_t)_{t\geq 0}$ with respect to an invariant probability measure μ is *ergodic* whenever (3.1) holds. We are now interested in finding the speed at which the latter convergence occurs. In order to tackle this problem, a toolbox of related functional inequalities was discovered over the past few decades. The purpose of this section is to introduce the main functional inequalities and shed light on the links existing between them. This chapter can be seen as the classical counterpart of Part IV, where quantum versions of these inequalities will be introduced and studied. Spending some time on these classical functional inequalities will also prove useful in Chapter 9 where a transference method from the commutative to the non-commutative framework is devised in order to estimate the speed of decoherence of various QMS.

Throughout this chapter, we fix an $\mathbb{L}_2(\mu)$-Markov semigroup $(P_t)_{t\geq 0}$ on a probability space $(E,\mathcal{G},\mu)$, with generator $(L_2, \mathrm{dom}(L_2))$, and $\mathcal{A}$ an algebra satisfying Condition 2.2.1. We are interested in finding bounds on the *mixing time* of the process, which is defined as

$$\tau(\varepsilon) := \inf\{\, t \geq 0;\ \|P_{t*}(\nu) - \mu\|_{\mathrm{TV}} \leq \varepsilon \quad \forall \nu << \mu \,\}, \tag{3.2}$$

where given two probability measures μ and ν on $\mathcal{G}$, $\|\mu - \nu\|_{\mathrm{TV}} := \sup_{A\in\mathcal{G}} |\mu(A) - \nu(A)|$ denotes the *total variation distance* between μ and ν.

Layout of the chapter: In Sections 3.1 to 3.4, we briefly summarize the main tools, namely Poincaré, (modified) logarithmic and Nash inequalities, used to find estimates on the mixing time $\tau(\varepsilon)$. After a quick reminder of some basic concepts of Riemannian geometry in Section 3.5, we expose the Bakry-Émery criterion that provides an elegant derivation of these inequalities in Section 3.6. We end this chapter with some examples for semigroups acting on finite and compact Lie groups in Section 3.7.

3.1. The spectral gap method

Some symmetric ergodic Markov semigroups are better behaved than others: these are the ones for which there exists a *spectral gap* in the spectrum of L_2. This means that there exists a constant $\lambda > 0$ such that the spectrum of L_2 is contained in $\{0\} \cup [\lambda, \infty)$. In this case the convergence towards the

mean is exponential:

$$\forall f \in \mathbb{L}_2(\mu), \quad \| P_t(f) - \mathbb{E}_\mu[f(X)] \|_{\mathbb{L}_2(\mu)} \le e^{-\lambda t} \| f \|_{\mathbb{L}_2(\mu)}. \tag{3.3}$$

The inequality (3.3) is actually equivalent to the spectral gap property. The largest λ satisfying this inequality is called the *spectral gap* and denoted by $\lambda(L)$. The inequality (3.3) can also be satisfied when μ is simply invariant. The following result holds in this more general case [Diaconis and Stroock, 1991, Lawler and Sokal, 1988, Bakry, 1994]:

Proposition 3.1.1. *Let $(P_t)_{t \ge 0}$ be an $\mathbb{L}_2(\mu)$-Markov semigroup for which μ is invariant. Then, (3.3) is equivalent to the Poincaré inequality:*

$$\forall f \in \mathcal{A}, \qquad \mathrm{Var}_\mu(f) := \mathbb{E}_\mu \left[f^2(X) \right] - \mathbb{E}_\mu[f(X)]^2 \le -\frac{1}{\lambda} \langle f, L_2 f \rangle_{\mathbb{L}_2(\mu)}. \tag{PI(λ)}$$

From (3.3), one can easily derive the following mixing time: if $\nu_t = P_{t*}(\nu)$,

$$\| \nu_t - \mu \|_{\mathrm{TV}} = \frac{1}{2} \int_E \left| \frac{d\nu_t}{d\mu}(x) - 1 \right| d\mu(x) \le \frac{1}{2} \left(\int_E \left| \frac{d\nu_t}{d\mu}(x) - 1 \right|^2 d\mu(x) \right)^{\frac{1}{2}} = \frac{1}{2} \left\| \frac{d\nu_t}{d\mu} - 1 \right\|_{\mathbb{L}_2(\mu)},$$

where the second line above follows from Jensen's inequality. Now, for any $f \in \mathbb{L}_2(\mu)$

$$\begin{aligned}
\langle f, \frac{d\nu_t}{d\mu} \rangle_{\mathbb{L}_2(\mu)} &= \int_E f(x) \frac{d\nu_t}{d\mu}(x) \, d\mu(x) \\
&= \int_E f(x) \, d\nu_t(x) \\
&= \int_E P_t(f)(x) \, d\nu(x) \\
&= \langle P_t(f), \frac{d\nu}{d\mu}(x) \rangle_{\mathbb{L}_2(\mu)} \qquad \Rightarrow \qquad \frac{d\nu_t}{d\mu} = \hat{P}_t \left(\frac{d\nu}{d\mu} \right),
\end{aligned}$$

where, for each $t \ge 0$, $\hat{P}_t$ denotes the adjoint of P_t with respect to inner product $\mathbb{L}_2(\mu)$. From Equation (2.21), PI(λ) is also satisfied for its generator $\hat{L}_2$, and therefore

$$\| P_{*t}(\nu) - \mu \|_{\mathrm{TV}} \le \frac{1}{2} \left\| \hat{P}_t \left(\frac{d\nu}{d\mu} \right) - \mathbb{1}_E \right\|_{\mathbb{L}_2(\mu)} \le \frac{1}{2} \, e^{-\lambda(L) t} \left\| \frac{d\nu}{d\mu} \right\|_{\mathbb{L}_2(\mu)}. \tag{3.4}$$

Estimates on the quantity on the right strongly depend on the problem. In the case of a Markov chain on a finite sample space E,

$$\left\| \frac{d\nu}{d\mu} \right\|_{\mathbb{L}_2(\mu)}^2 = \sum_{x \in E} \left(\frac{\nu(\{x\})}{\mu(\{x\})} \right)^2 \mu(\{x\}) \le \left[\min_{x \in E} \mu(\{x\}) \right]^{-1}. \tag{3.5}$$

In this particular case,

$$\tau(\varepsilon) \le -\frac{1}{\lambda(L)} \ln \left(2 \varepsilon \left[\min_{x \in E} \mu(\{x\}) \right]^{\frac{1}{2}} \right). \tag{3.6}$$

3.2. Logarithmic Sobolev inequality and hypercontractivity

Bounds derived from the Poincaré method described above are usually not tight. This is due to the bad behavior of the right hand side of Equation (3.4) with respect to the dimension of the system. In the case of Markov chains for instance, the upper bound found in Equation (3.5) usually scales

linearly with the dimension of the system. This is not good enough at short times where mixing can occur much faster. One idea that greatly improves this estimate is to use more refined notions of contraction than the $\mathbb{L}_2 \to \mathbb{L}_2$ contraction provided by the Poincaré inequality. For instance:

$$
\begin{aligned}
\|\nu_{t+s} - \mu\|_{\mathrm{TV}} &\leq \frac{1}{2} \left\| \hat{P}_{t+s}\left(\frac{d\nu}{d\mu}\right) - \mathbb{1}_E \right\|_{\mathbb{L}_2(\mu)} \\
&\leq \frac{1}{2} \left\| \frac{d\nu}{d\mu} \right\|_{\mathbb{L}_1(\mu)} \left\| \hat{P}_s : \mathbb{L}_1(\mu) \to \mathbb{L}_2(\mu) \right\| \left\| \hat{P}_t \circ (\mathrm{id} - \mathbb{E}_\mu[.]\,\mathbb{1}_E) : \mathbb{L}_2(\mu) \to \mathbb{L}_2(\mu) \right\| \\
&\leq \frac{1}{2} \left\| P_s : \mathbb{L}_2(\mu) \to \mathbb{L}_\infty(\mu) \right\| \mathrm{e}^{-\lambda(L)\,t},
\end{aligned} \tag{3.7}
$$

where we used that $\|d\nu/d\mu\|_{\mathbb{L}_1(\mu)} = 1$. Since P_s is contractive for any $\mathbb{L}_p$ norm with $p \geq 1$ (cf. (2.13)), and in view of the limit in (3.1), one can hope to find a time $s > 0$ such that $\|P_s : \mathbb{L}_2(\mu) \to \mathbb{L}_\infty(\mu)\| \leq 2$ (2 here is of course arbitrary). However, even for classical Markov chains, this quantity is in practice difficult to estimate (see Section 3.4). Hypercontractivity provides a tool to interpolate between this norm and the $\mathbb{L}_2 \to \mathbb{L}_2$ norm given by the spectral gap method, by providing an estimation of the $\mathbb{L}_2 \to \mathbb{L}_p$ norm for $p > 2$ instead. In this case, the factor $\|d\nu/d\mu\|_{\mathbb{L}_{\hat{p}}(\mu)}$ appears, which indeed interpolates between the previous two methods. The great discovery of Gross was that finding a time $t \geq 0$ for which P_t becomes a contractive operator from $\mathbb{L}_2$ to $\mathbb{L}_p$ is a problem equivalent to the one of optimizing the so-called logarithmic Sobolev inequality. Exploiting this equivalence, Diaconis and Saloff-Coste were able to find optimal or near to optimal upper bounds of the mixing time [Diaconis and Saloff-Coste, 1996a, Diaconis and Saloff-Coste, 1993, Diaconis et al., 1993]: in the context of Markov chains, the bound derived in (3.6) via the Poincaré method was of order $\ln\left[\min_x \mu(\{x\})\right]^{-1}$, whereas hypercontractivity leads to an upper bound of order $\ln\ln[\min_x \mu(\{x\})]^{-1}$. Thus hypercontractivity provides an exponential improvement over the Poincaré inequality.

First introduced by Nelson in the context of quantum field theory [Nelson, 1973b, Nelson, 1966] (see also the bibliographic review by [Davies et al., 1992]), *hypercontractivity* denotes the following property for a semigroup $(P_t)_{t \geq 0}$: for any $q \leq p \leq 1 + (q-1)\,\mathrm{e}^{4t/c}$,

$$
\|P_t : \mathbb{L}_q(\mu) \to \mathbb{L}_p(\mu)\| \leq \exp\left\{ 2d\left(\frac{1}{q} - \frac{1}{p}\right) \right\}, \tag{HC(c, d)}
$$

for positive constants $c > 0$ and $d \geq 0$.

Hypercontractivity is intimately related to the so-called logarithmic Sobolev inequality, first introduced by [Gross, 1975a, Gross, 1975b, Gross, 1993] for the Ornstein Uhlenbeck semigroup: denote by $\mathcal{A}^+$ the subset of functions in $\mathcal{A}$ which are bounded below by a positive constant. Then, the entropy of $f \in \mathcal{A}^+$ is defined as:

$$
\mathrm{Ent}\,(f) := \int_E f \ln f \, d\mu - \left(\int_E f \, d\mu \right) \ln \left(\int_E f \, d\mu \right). \tag{3.8}
$$

In the case when $\|f\|_{\mathbb{L}_1(\mu)} = 1$, we easily get $\mathrm{Ent}\,(f) = D(\nu\|\mu)$, where $d\nu/d\mu = f$. The function $x \mapsto x \log x$ being convex, the above quantity is always positive, and is equal to 0 if and only if f is constant μ-almost surely. Then, given $p > 1$, $(P_t)_{t \geq 0}$ is said to satisfy a *logarithmic Sobolev inequality* of order p if for any $f \in \mathcal{A}^+$:

$$
\mathrm{Ent}(f^p) \leq c\frac{p\hat{p}}{4}\mathcal{E}(f^{p-1}, f) + 2d\|f\|_{\mathbb{L}_p(\mu)}^p. \tag{LSI$_p$(c, d)}
$$

Remark 3.2.1. The positivity of $\mathcal{E}(f^{p-1}, f)$ for $p \geq 1$ can be easily proved as follows: since $x \mapsto x^p$ is

convex, $(P_t(f))^p \le P_t(f^p)$. Therefore, the invariance of μ implies that for any $f \in \mathcal{A}^+$

$$\mathbb{E}_\mu[P_t(f)^p] \le \mathbb{E}_\mu[f^p].$$

This implies that $\frac{d}{dt}\mathbb{E}_\mu[P_t(f)^p]\big|_{t=0} \le 0$. We can conclude after noticing that the derivative on the left hand side is exactly equal to $-p\mathcal{E}(f^{p-1}, f)$.

The following equivalence between $\text{LSI}_p(c, d)$ and $\text{HC}(c, d)$ was first observed by [Gross, 1975a] in the case of the Ornstein-Uhlenbeck semigroup: This is due to the differentiation of the $\mathbb{L}_p$ norm along the semigroup together with a use of the regularity of the Dirichlet forms (cf. Equation (2.29) and lemma 2.4.2) (see e.g. Proposition 3.4 of [Bakry, 1994], Theorem 3.5 of [Diaconis and Saloff-Coste, 1996a]).

Theorem 3.2.2. *Assume that $(P_t)_{t\ge0}$ is either reversible, or a diffusion semigroup. Then, given two constants $c > 0$, $d \ge 0$, $\text{HC}(c, d)$ is equivalent to $\text{LSI}_2(c, d)$. In the general case, $\text{HC}(2c, d)$ is still equivalent to $\text{LSI}_2(c, d)$.*

In words, the logarithmic Sobolev inequality can be thought of as an infinitesimal version of the hypercontractivity property. The advantage of the former lies in that for most of the relevant physical evolutions modeled by a Markovian dynamics, the semigroup is known only through its generator. Hence, a logarithmic Sobolev inequality provides contractivity properties of a semigroup $(P_t)_{t\ge0}$ without the necessity of knowing its exact expression.

In the case of a reversible Markov semigroup, interpolation techniques ensure that LSI_2 holds as long as, for some $t > 0$ and two real numbers $1 < p < q \le \infty$, $\|P_t : \mathbb{L}_p(\mu) \to \mathbb{L}_q(\mu)\| < \infty$. This fact is due to [Gross, 1975b], yet the idea can be traced back to [Simon and Høegh-Krohn, 1972] (see Theorem 3.6 of [Bakry, 1994]).

Strong LSI and link to Poincaré   Choosing $f = 1$ in $\text{LSI}_p(c, d)$, one easily observes that d has to be nonnegative for the inequality to hold. In fact, in the case of the Ornstein Uhlenbeck semigroup $\text{LSI}_2(c, d)$ holds with $d = 0$. More generally, we say that a semigroup $(P_t)_{t\ge0}$ of generator L satisfies a *strong log-Sobolev inequality* of order p when it satisfies $\text{LSI}_p(c, 0)$. In this case, it is standard to denote the inverse of the best constant c achieving $\text{LSI}_p(c, 0)$ by $\alpha_p(P)$ or $\alpha_p(L)$, and call it the *p-logarithmic Sobolev constant* of $(P_t)_{t\ge0}$:

$$\alpha_2(L)\,\text{Ent}(f^2) \le \mathcal{E}(f, f). \tag{3.9}$$

Remark 3.2.3. The strong logarithmic Sobolev inequality can be given an operational interpretation as follows: the left hand side of the inequality is the entropy function, which for $f = (d\nu/d\mu)^{1/2}$, $\nu \ll \mu$, reduces to the relative entropy $D(\nu\|\mu)$. By Sanov's theorem, this is known to be the large deviation functional corresponding to the task of inferring the measure μ from the empirical measure constructed from the observation of i.i.d. instances of it. On the other hand, Donsker and Varadan proved that the Dirichlet form $\mathcal{E}(f, f) := -\langle\sqrt{d\nu/d\mu}, L_2(\sqrt{d\nu/d\mu})\rangle_{\mathbb{L}_2(\mu)}$ is the large deviation functional corresponding to the task of inferring μ from the occupation time measure associated to the underlying process $(X_t)_{t\ge0}$. In this context, the strong logarithmic Sobolev inequality justifies that, up to the multiplicative constant $\alpha_2(L)$, inferring the measure μ from i.i.d. observations is asymptotically less efficient than from counting the occupation times of a process converging to it (see [den Hollander, 2008]). While the operational interpretation of the relative entropy also applies in the quantum setting [van Horssen and Guţă, 2015, Ogata, 2010], obtaining an interpretation of the

quantum Dirichlet form seems more complicated. This is largely due to the lack of an appropriate notion of quantum trajectory needed to define occupation times.

Example 3.2.4. The first logarithmic Sobolev inequality was found by Gross [Gross, 1975b] for the (classical) Ornstein Uhlenbeck semigroup $(P_t^{\mathrm{cOU}})_{t\geq 0}$ on $\mathbb{R}^n$ of generator

$$L^{\mathrm{cOU}}(f)(x) := \Delta f(x) - x.\nabla f(x).$$

This process, which can be interpreted as a heat diffusion with friction term, has the standard Gaussian measure μ_G as its invariant. Moreover, for any f for which each side of the inequality below is well-defined:

$$\mathrm{Ent}(f^2) \leq 2 \int_{\mathbb{R}} |\nabla f|^2 d\mu_G\,, \tag{Gross-LSI}$$

where the entropy is defined with respect to μ_G. In other words, $\alpha_2(L^{\mathrm{cOU}}) = \frac{1}{2}$.

The strong logarithmic Sobolev inequality is closely related to the Poincaré inequality introduced in Section 3.1. In fact, the following result was originally proved in [Rothaus, 1981]. The proof is now standard and consists in a perturbation of $\mathrm{LSI}_2(c,0)$ applied to $1 + \varepsilon f$, where $\varepsilon > 0$ is the perturbation parameter (see e.g. Proposition 3.7 of [Bakry, 1994]).

Proposition 3.2.5. *Given an $\mathbb{L}_2(\mu)$-Markov semigroup $(P_t)_{t\geq 0}$ with invariant probability measure μ, $\mathrm{LSI}_2(c,0)$ implies $\mathrm{PI}(\lambda)$ with $\lambda \geq 2/c$.*

The converse of last proposition is known to be false in general (see e.g. example 21.19 of [Villani, 2008]). However, combining the spectral gap together with a logarithmic Sobolev inequality results in a tightening of the latter (see Proposition 3.9 of [Bakry, 1994]):

Proposition 3.2.6. *Given a $\mathbb{L}_2(\mu)$-Markov semigroup with invariant probability measure μ, assume $\mathrm{PI}(\lambda)$ holds for some $\lambda > 0$. Then $\mathrm{LSI}_2(c,d)$ implies $\mathrm{LSI}_2(c',0)$ with*

$$c' = c + \frac{2(d+1)}{\lambda}.$$

In the context of Markov chains on finite sample spaces, the above inequality, a combination of interpolation techniques with Proposition 3.2.6 and the following estimate

$$\| \mathrm{id} : \mathbb{L}_2(\mu) \to \mathbb{L}_4(\mu)\| \leq \left[\min_{x \in E} \mu(\{x\}) \right]^{-\frac{1}{4}} \tag{3.10}$$

yields the following bounds on the constant c: $\mathrm{LSI}_2(c,0)$ always holds, with

$$c \leq \frac{2 - \ln\left[\min_{x \in E} \mu(\{x\}) \right]}{\lambda(L)}. \tag{3.11}$$

Back to our original problem of the estimation of the mixing time in the absence of an estimate for the $\mathbb{L}_2 \to \mathbb{L}_\infty$ norm of the semigroup, the steps followed in (3.7) can be modified. Assume that $(P_t)_{t\geq 0}$

satisfies $\mathrm{LSI}_2(\alpha^{-1}, 0)$. Then, given $p \equiv p(s) = 1 + \mathrm{e}^{2\alpha s}$:

$$
\begin{aligned}
\|\nu_{t+s} - \mu\|_{\mathrm{TV}} &\leq \frac{1}{2} \left\| \hat{P}_{t+s} \left(\frac{d\nu}{d\mu} \right) - \mathbb{1}_E \right\|_{\mathbb{L}_2(\mu)} \\
&\leq \frac{1}{2} \left\| \frac{d\nu}{d\mu} \right\|_{\mathbb{L}_{\hat{p}}(\mu)} \left\| \hat{P}_s : \mathbb{L}_{\hat{p}}(\mu) \to \mathbb{L}_2(\mu) \right\| \left\| \hat{P}_t \circ (\mathrm{id} - \mathbb{E}_\mu[.] \, \mathbb{1}_E) : \mathbb{L}_2(\mu) \to \mathbb{L}_2(\mu) \right\| \\
&\leq \frac{1}{2} \left\| \frac{d\nu}{d\mu} \right\|_{\mathbb{L}_{\hat{p}}(\mu)} \mathrm{e}^{-\lambda(L)\,t},
\end{aligned}
\tag{3.12}
$$

In the context of a Markov chain on a finite sample space, $\left\| \frac{d\nu}{d\mu} \right\|_{\mathbb{L}_{\hat{p}}(\mu)} \leq \left[\min_{x \in E} \mu(\{x\}) \right]^{-\frac{1}{p}}$. Choosing $p(s) = 1 - \ln\left[\min_{x \in E} \mu(\{x\}) \right]$, we find that for $t' = t + s \geq s = \frac{1}{2\alpha} \ln\ln \left[\min_{x \in E} \mu(\{x\}) \right]^{-1}$ (with the condition that $\min_{x \in E} \mu(\{x\}) \leq 1/\mathrm{e}$),

$$
\|\nu_{t+s} - \mu\|_{\mathrm{TV}} \leq \frac{1}{2} \, \mathrm{e}^{-\lambda(L)t} \ .
$$

In other words,

$$
\tau(\varepsilon) \leq \frac{1 - \ln(2\varepsilon)}{\lambda(L)} + \frac{1}{2\alpha} \ln\ln \left[\min_{x \in E} \mu(\{x\}) \right]^{-1},
\tag{3.13}
$$

which constitutes a logarithmic improvement over the bound found in (3.6) from the unique use of the spectral gap method. In the case of a reversible semigroup, we can even replace α by 2α.

3.3. Entropy decay and the modified logarithmic Sobolev inequality

We saw in Section 3.1 that the spectral gap inequality $\mathrm{PI}(\lambda)$ is linked, via a Lyapunov type argument, to the exponential convergence in $\mathbb{L}_2(\mu)$ of $P_t(f)$ towards $\mathbb{E}_\mu[f]$ as $t \to \infty$. Since we saw in Proposition 3.2.5 that $\mathrm{PI}(\lambda)$ can be understood as a linearization of $\mathrm{LSI}_2(2\lambda^{-1}, 0)$, we expect the convergence to be stronger in general in the case when the latter is satisfied. In order to show this, we introduce the analogue for $p = 1$ of the logarithmic Sobolev inequalities defined above. More precisely, an $\mathbb{L}_2(\mu)$-Markov semigroup is said to satisfy $\mathrm{LSI}_1(c, 0)$ for some $c > 0$ if, for any $f \in \mathcal{A}^+$:

$$
\mathrm{Ent}(f) \leq \frac{c}{4} \mathcal{E}(\ln f, f). \tag{$\mathrm{LSI}_1(c, 0)$}
$$

$\mathrm{LSI}_1(c, 0)$ can be interpreted as the limit when $p \to 1$ of $\mathrm{LSI}_p(c, 0)$. From Lemma 2.4.1, one readily sees that, if μ is invariant, then $\mathrm{LSI}_2(c, 0)$ implies $\mathrm{LSI}_1(2c, 0)$. The conditions of reversibility or diffusivity improve this result by a factor of two: $\mathrm{LSI}_2(c, 0)$ implies $\mathrm{LSI}_1(c, 0)$. In fact, in the case of diffusions, these two inequalities are equivalent, due to Equation (2.28). Moreover, by an argument similar to the one leading to Proposition 3.2.5, one can show that $\mathrm{LSI}_1(c, 0)$ implies $\mathrm{PI}(2/c)$ whenever μ is invariant [Rothaus, 1981].

The following result was first established in [Bakry and Émery, 1984] in the case of diffusion semigroups. The proof consists in a Lyapunov argument similar to the one leading to the decay of the variance from $\mathrm{PI}(\lambda)$ (see also [Bakry, 1994, Diaconis and Saloff-Coste, 1996a]).

Proposition 3.3.1. *Let $(P_t)_{t \geq 0}$ be a $\mathbb{L}_2(\mu)$-Markov semigroup with invariant measure μ. Then*

$(P_t)_{t\geq 0}$ *satisfies* $\mathrm{LSI}_1(c,0)$ *if and only if for every positive function* f,

$$\mathrm{Ent}(P_t(f)) \leq \mathrm{e}^{-\frac{4}{c}t}\,\mathrm{Ent}(f)\,. \tag{3.14}$$

Comparing the last result to Proposition 3.1.1 we see that the spectral gap inequality provides an exponential convergence of $P_t(f)$ towards $\mathbb{E}_\mu[f]$ in $\mathbb{L}_2(\mu)$, whereas $\mathrm{LSI}_1(c,0)$ gives it in terms of entropy. The fact that this convergence is stronger than the convergence in $\mathbb{L}_2$ provided by $\mathrm{PI}(\lambda)$ is a consequence of Pinsker's inequality [Pinsker, 1960, Csiszár, 1967, Kullback, 1967]:

$$\| P_{t*}(\nu) - \mu \|_{\mathrm{TV}}^2 \leq \frac{1}{2}D(P_{t*}(\nu)\|\mu) = \frac{1}{2}\,\mathrm{Ent}\left(\hat{P}_t\left(\frac{d\nu}{d\mu}\right)\right)\,. \tag{3.15}$$

Now, assuming that $\mathrm{LSI}_2(c,0)$ holds for $(P_t)_{t\geq 0}$, it also holds for $(\hat{P}_t)_{t\geq 0}$, by Equation (2.21). Therefore, if μ is invariant, $\mathrm{LSI}_1(2c,0)$ holds and by Proposition 3.3.1,

$$\| P_{t*}(\nu) - \mu \|_{\mathrm{TV}} \leq \mathrm{e}^{-\frac{t}{c}}\,\sqrt{D(\nu\|\mu)/2}\,.$$

In the case of reversibility or diffusivity, the rate of convergence on the right hand side above can be strengthened into $2c^{-1}$, since in this case $\mathrm{LSI}_1(c,0)$ holds.

The above discussion motivates the introduction of the following *modified logarithmic Sobolev inequality*: $(P_t)_{t\geq 0}$ satisfies the modified logaritmic Sobolev inequality of parameter $\alpha > 0$ if for any measure $\nu << \mu$:

$$4\alpha\,D(\nu\|\mu) \leq -\int \ln\left(\frac{d\nu}{d\mu}\right)\hat{L}_2\left(\frac{d\nu}{d\mu}\right)d\mu \equiv \mathrm{EP}_{L_2}(\nu)\,, \tag{MLSI(α)}$$

where $\mathrm{EP}_{L_2}(\nu)$ is called the *entropy production* of the semigroup $(P_t)_{t\geq 0}$. This inequality coincides with $\mathrm{LSI}_1(\alpha^{-1},0)$ in the reversible situation. $\mathrm{MLSI}(\alpha)$ was first introduced by [Bobkov and Tetali, 2003, Bobkov and Tetali, 2006]. The reason for introducing this last inequality is that it leads to the convergence of the relative entropy without the necessity to go through LSI_2:

$$\| P_{t*}(\nu) - \mu \|_{\mathrm{TV}} \leq \mathrm{e}^{-2\alpha t}\,\sqrt{D(\nu\|\mu)/2}\,.$$

The best constant achieving $\mathrm{MLSI}(\alpha)$ is referred to as the MLSI constant, and denoted by $\alpha_1(L)$. Similarly to the Poincaré and hypercontractivity methods, one is still left with the problem of controlling the relative entropy $D(\nu\|\mu)$. Again, this will depend on the geometry of the sample space E. In the case of a Markov chain on a finite sample space,

$$D(\nu\|\mu) \leq \ln\left(\left[\min_{x\in E}\mu(\{x\})\right]^{-1}\right)\,, \tag{3.16}$$

so that,

$$\| P_{t*}(\nu) - \mu \|_{\mathrm{TV}} \leq \mathrm{e}^{-2\alpha_1(L)t}\,\sqrt{\frac{1}{2}\ln\left(\left[\min_{x\in E}\mu(\{x\})\right]^{-1}\right)}\,.$$

This leads to the following estimation of mixing times:

$$\tau(\varepsilon) \leq \frac{1}{2\alpha_1(L)}\ln\left(\varepsilon^{-1}\sqrt{\frac{1}{2}\ln\left(\left[\min_{x\in E}\mu(\{x\})\right]^{-1}\right)}\right)\,, \tag{3.17}$$

hence the promised exponential improvement over 3.6 as long as $\alpha_1(L) \sim \lambda(L)$.

3.4. Nash and Sobolev inequalities

In the last section, we saw that hypercontractivity provides estimates on the $\mathbb{L}_2 \to \mathbb{L}_p$ norms of the maps P_t for $2 < p < \infty$, which lead to better estimates on the mixing time of the semigroup. As discussed at the beginning of that section, having an estimate on the $\mathbb{L}_2 \to \mathbb{L}_\infty$ norms of P_t would provide even more precise estimates. Finding these estimates is typically only possible at short times. A semigroup satisfying such a contraction property is called *ultracontractive*. In a similar way as for hypercontractivity and the logarithmic Sobolev inequality, the notion of ultracontractivity possesses a infinitesimal versions called Nash inequality, or the Sobolev inequality. These inequalities have found many applications in precise estimations of the density of a semigroup. (see [Bakry, 1994, Ledoux, 1992, Varopoulos, 1985, Carlen, 1987, Davies and Simon, 1984, Bakry, 1990], and the book [Davies, 1989] for an exhaustive bibliography on the subject). However, they generally don't hold in infinite dimensions, which is one of the main motivations behind the study of log-Sobolev inequalities (see chapter Chapter 10 for more details).

The classical Nash inequality was introduced in [Nash, 1958] to obtain regularity properties on the solutions of parabolic partial differential equations. In the Euclidean space $\mathbb{R}^n$, it can be stated as follows: there exists a constant $c_n > 0$ (depending only on n), such that for any real-valued, smooth function f vanishing at infinity,

$$\|f\|_{\mathbb{L}^2(\mathbb{R}^n)}^{1+n/2} \leq c_n \|f\|_{\mathbb{L}^1(\mathbb{R}^n)} \|\nabla f\|_{\mathbb{L}^2(\mathbb{R}^n)}^{n/2}. \tag{Nash}$$

The optimal constant c_n was later evaluated by Carlen and Loss [Carlen and Loss, 1993]. Nash inequality implies *ultracontractivity* of the heat convolution semigroup P_t^{Heat}. This means that it maps $\mathbb{L}_1(\mathbb{R}^n)$ to $\mathbb{L}_\infty(\mathbb{R}^n)$ with

$$\|P_t^{\mathrm{Heat}} : \mathbb{L}_1(\mathbb{R}^n) \to \mathbb{L}_\infty(\mathbb{R}^n)\| \leq \left(\frac{1}{\pi e t}\right)^{n/2}, \tag{UC}$$

where $P_t^{\mathrm{Heat}}(f) = f * g_{2t}$, and g_{2t} denotes the probability distribution function of a centered Gaussian random variable on $\mathbb{R}^n$ of variance $2t$.

More generally, given a probability space $(E, \mathcal{G}, \mu)$, and an $\mathbb{L}_2(\mu)$-Markov semigroup $(P_t)_{t \geq 0}$, $(P_t)_{t \geq 0}$ is said to satisfy a *Sobolev inequality* if there exist constants $C > 0$, $n > 2$ and $T \geq 0$ such that for any $f \in \mathcal{A}$:

$$\|f\|_{\mathbb{L}_{\frac{2n}{n-2}}(\mu)}^2 \leq C \left(\mathcal{E}(f, f) + \frac{1}{T}\|f\|_{\mathbb{L}_2(\mu)}^2\right). \tag{SI(C, n, T)}$$

Diffusion semigroups on a Riemannian manifold satisfy SI(C, n, T) with n that coincides with the dimension of the manifold. This justifies calling the coefficient n the *dimension of the semigroup*. Moreover, for $n > 2$, Nash inequality can be derived from SI(C, n, T) via Hölder's inequality (see [Zegarlinski, 1992]):

$$\|f\|_{\mathbb{L}_2(\mu)}^{2(1+2/n)} \leq C \left(\mathcal{E}(f, f) + \frac{1}{T}\|f\|_{\mathbb{L}_2(\mu)}^2\right) \|f\|_{\mathbb{L}_1(\mu)}^{4/n}, \tag{Nash(C, n, T)}$$

As mentioned above, in the case when the semigroup is reversible, Nash inequality is equivalent to the following *ultracontractivity* property of the semigroup (see e.g. [Saloff-Coste, 1997, Carlen et al., 1987]):

Theorem 3.4.1. *Assume that* $\mathrm{Nash}(C, n, T)$ *holds, then for any* $t \leq T$:

$$\|P_t : \mathbb{L}_1(\mu) \to \mathbb{L}_2(\mu)\|, \; \|P_t : \mathbb{L}_2(\mu) \to \mathbb{L}_\infty(\mu)\| \leq \mathrm{e} \left(\frac{nC}{4t} \right)^{n/4}. \tag{3.18}$$

On the other hand, assume that $(P_t)_{t \geq 0}$ *is reversible with respect to* μ. *Then* $\|P_t : \mathbb{L}_1(\mu) \to \mathbb{L}_2(\mu)\| \leq \left(\frac{C_0}{t} \right)^{n/4}$ *for any* $t \leq T$ *implies* $\mathrm{Nash}(C, n, T)$ *with* $C \equiv 2C_0(1 + 2/n)(1 + n/2)^{2/n}$.

We already explained the importance of such a bound on the $\mathbb{L}_2(\mu) \to \mathbb{L}_\infty(\mu)$ norm of P_t at the beginning of Section 3.2: in particular, when combined with the spectral gap $\lambda(L)$, it leads to the following estimate: for $t \geq T$:

$$\|P_{t*}(\nu) - \mu\|_{\mathrm{TV}} \leq \frac{1}{2} \left(\frac{nC}{4T} \right)^{n/4} \mathrm{e}^{-\lambda(L)(t-T)}.$$

This leads to the following very strong bound on the mixing time of $(P_t)_{t \geq 0}$:

$$\tau(\varepsilon) \leq T - \frac{1}{\lambda(L)} \ln \left(2\varepsilon \left(\frac{4T}{nC} \right)^{n/4} \right).$$

We argued the estimate $\mathbb{L}_2(\mu) \to \mathbb{L}_\infty(\mu)$ should typically provide better bounds on the mixing times. This is also suggested by the following theorem (see Theorem 2.3.6 of [Saloff-Coste, 1997]):

Proposition 3.4.2. *Assume that the semigroup* $(P_t)_{t \geq 0}$ *is reversible with respect to* μ. *Assume* $\mathrm{Nash}(C, n, T)$, *then the logarithmic Sobolev constant* $\alpha_2(L)$ *is lower bounded by*

$$\alpha_2(L) \geq \frac{\lambda(L)}{2 \left(1 + \lambda(L) t_0 + \frac{n}{4} \ln \left(\frac{nC}{4t_0} \right) \right)},$$

for any $0 < t_0 \leq T$.

In fact, one can combine the different techniques described up to now in order to get an even better estimation of the mixing time by dividing time t into three parts: given $t = s + \eta + u$, and for $p = 1 + \mathrm{e}^{2\alpha_2(L)\eta}$,

$$
\begin{aligned}
\|\nu_t - \mu\|_{\mathrm{TV}} &\leq \frac{1}{2} \left\| \hat{P}_t \left(\frac{d\nu}{d\mu} \right) - \mathbb{1}_E \right\|_{\mathbb{L}_2(\mu)} \\
&\leq \frac{1}{2} \left\| \hat{P}_s \left(\frac{d\nu}{d\mu} \right) \right\|_{\mathbb{L}_{\hat{p}}(\mu)} \left\| \hat{P}_\eta : \mathbb{L}_{\hat{p}}(\mu) \to \mathbb{L}_2(\mu) \right\| \left\| \hat{P}_u \circ (\mathrm{id} - \mathbb{E}_\mu[.] \, \mathbb{1}_E) : \mathbb{L}_2(\mu) \to \mathbb{L}_2(\mu) \right\| \\
&= \frac{1}{2} \left\| \hat{P}_s \left(\frac{d\nu}{d\mu} \right) \right\|_{\mathbb{L}_{\hat{p}}(\mu)} \left\| P_\eta : \mathbb{L}_2(\mu) \to \mathbb{L}_p(\mu) \right\| \left\| \hat{P}_u \circ (\mathrm{id} - \mathbb{E}_\mu[.] \, \mathbb{1}_E) : \mathbb{L}_2(\mu) \to \mathbb{L}_2(\mu) \right\| \\
&\leq \frac{1}{2} \left\| \hat{P}_s \left(\frac{d\nu}{d\mu} \right) \right\|_{\mathbb{L}_2(\mu)}^{\frac{2}{p}} \mathrm{e}^{-\lambda(L)u} \\
&\leq \frac{1}{2} \left\| P_s : \mathbb{L}_2(\mu) \to \mathbb{L}_\infty(\mu) \right\|^{\frac{2}{p}} \mathrm{e}^{-\lambda(L)u}, \tag{3.19}
\end{aligned}
$$

where the fourth line arises by interpolation between $\mathbb{L}_1$ and $\mathbb{L}_2$. Therefore, assuming $\mathrm{Nash}(C, n, T)$ and $s = T$, we get the following bound

$$\|\nu_t - \mu\|_{\mathrm{TV}} \leq \frac{1}{2} \left(\frac{nC}{4T} \right)^{\frac{n}{2p}} \mathrm{e}^{-\lambda(L)u}. \tag{3.20}$$

by choosing $e^{2\alpha_2(L)\eta} = 1 + \frac{n}{2}\ln\left(\frac{nC}{4T}\right)$, we directly get

$$\tau(\varepsilon) \leq T + \frac{1}{2\alpha_2(L)}\ln\left(1 + \frac{n}{2}\ln\frac{nC}{4T}\right) + \frac{1 - \ln(2\varepsilon)}{\lambda(L)}. \tag{3.21}$$

Once again, this can be strengthened by replacing $\alpha_2(L)$ by $2\alpha_2(L)$ in the case of a reversible semigroup.

The functional inequalities that we introduced here were seen as properties satisfied by a Markov semigroup, in the original spirit of Bakry and Émery. In Chapter 4, we operate a shift and view these inequalities as a property satisfied by the measure μ, in the case of a diffusive Markov semigroup defined on a manifold $\mathcal{M}$. Indeed, one can relate the constants appearing in the various functional inequalities to the curvature and dimension of $\mathcal{M}$. In order to make this shift, we first introduce some concepts from Riemannian geometry that are needed. These will also be useful in Chapter 12.

3.5. Interlude: Riemannian geometry

3.5.1. Geodesics

Given a metric space $(\mathcal{X}, d)$, a curve $\gamma : [0,1] \to \mathcal{X}$ is called a *constant speed geodesic* provided for any $s, t \in [0,1]$:

$$d(\gamma_s, \gamma_t) = |t - s|\, d(\gamma_0, \gamma_1). \tag{3.22}$$

Now, a metric space $(\mathcal{X}, d)$ is called *geodesic* if for every $x, y \in \mathcal{X}$ there exists a constant speed geodesic connecting them, i.e. a constant speed geodesic $(\gamma_s)_{s\in[0,1]}$ such that $\gamma_0 = x$ and $\gamma_1 = y$. Such spaces include Riemannian manifolds.

We recall that, given a Riemannian manifold $(\mathcal{M}, g)$, the gradient of a continuously differentiable function $\Phi : \mathcal{M} \to \mathbb{R}$ is the vector field $\nabla\Phi$ defined by the equation

$$d_x\Phi \,.\, \vec{v} = g_x(\nabla_x\Phi, \vec{v}) \qquad \forall \vec{v} \in T_x\mathcal{M},$$

where $T_x\mathcal{M}$ denotes the tangent space at $x \in \mathcal{M}$, and where $d_x\Phi$ denotes the differential of Φ at x. Equivalently, for any smooth path $(\Phi_s)_{-\varepsilon \leq s \leq \varepsilon}$ in $\mathcal{M}$ with $\gamma_0 = x$ and $\frac{d\gamma_s}{ds}\big|_{s=0} = \vec{v}$,

$$\frac{d}{ds}\bigg|_{s=0}\Phi(\gamma_s) = g_x(\nabla_x\Phi, \vec{v}).$$

Then, for a fixed norm of the vector $\vec{v}$, $\nabla_x\Phi$ indicates the direction in which Φ increases most rapidly.

In what follows, by Riemannian manifold we mean a smooth, complete connected finite-dimensional Riemannian manifold $(\mathcal{M}, g)$, of dimension n, distinct from a point, and equipped with a smooth metric tensor. For $L > 0$, a curve $\gamma : (-L, L) \to \mathcal{M}$ is called a geodesic if it satisfies the geodesic equation

$$\ddot{\gamma} = 0.$$

In the Eulerian description, we describe the curve γ by its derivative at each point $s \in (-L, L)$, defining the velocity field $\xi(s, \gamma(s))$ at s, so that for each $s \in (-L, L)$, $\xi(s, \gamma(s)) \in T_{\gamma(s)}\mathcal{M}$. The geodesic equation is then rewritten as

$$\frac{\partial\xi}{\partial s} + \nabla_\xi\xi = 0, \tag{3.23}$$

which is also known as the *pressureless Euler equation*. Here, ∇_ξ stands for the covariant derivation along the vector ξ. Further assuming that ξ is a gradient of the form $\nabla\Phi$, the above equation reduces to the so-called *Hamilton-Jacobi equation*

$$\partial_s\Phi + \frac{\|\nabla\Phi\|^2}{2} = 0\,,$$

where $\|\xi\|^2 := g(\xi,\xi)$ is the square of the norm associated to the metric g.

3.5.2. Ricci curvature and Bochner formula
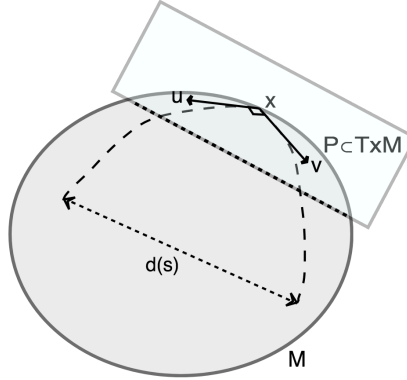
Figure 3.1.: Riemannian distance in positive curvature.

Let $(\mathcal{M},g)$ be a Riemannian manifold, $x \in \mathcal{M}$ and let $\gamma_{\vec{v}}$, $\gamma_{\vec{w}}$ be two constant speed geodesics of speed 1 starting from x, of respective tangent vectors at x denoted by $\vec{v}$ and $\vec{w}$. Then, for each $s \geq 0$, the Riemannian distance between $\gamma_{\vec{v}}(s)$ and $\gamma_{\vec{w}}(s)$ can be shown to satisfy the following Taylor expansion:

$$\delta(s) = \sqrt{2(1-\cos\theta)}\,s\left(1 - \frac{\kappa_x\cos^2(\theta/2)}{6}\,s^2 + \mathcal{O}(s^3)\right),$$

where θ is the angle between $\vec{v}$ and $\vec{w}$, and κ_x denotes the *Gauss curvature* at x of the surface $\exp_x(P)$ defined by all the geodesics intersecting at x with associated tangent vectors in $P := \operatorname{span}(\vec{v},\vec{w})$. It is independent of the vectors $\vec{v}$ and $\vec{w}$ and intrinsic. In the case when $\kappa_x = 0$, we recover the Euclidean case where $\delta(s)$ is simply linear in s, up to third order. Next, the *sectional curvature* at x with respect to the plane $P \subset T_x\mathcal{M}$, denoted by $\sigma_x(P)$, is simply the Gauss curvature of the surface $\exp_x(P)$.

We are now ready to give a definition of the *Ricci curvature*: given $x \in \mathcal{M}$ and a norm 1 tangent vector $\vec{u} \in T_x\mathcal{M}$, fix an orthonormal basis $(\vec{u} = \vec{e}_1, \vec{e}_2, ..., \vec{e}_n)$ of $T_x\mathcal{M}$ with first basis vector $\vec{u}$. Then

$$\operatorname{Ric}_x(\vec{u}) := \sum_{j=2}^{n} \sigma_x(\vec{u},\vec{e}_j)\,, \qquad\qquad \text{(Ricci curvature)}$$

where given two tangent vectors $\vec{u},\vec{v} \in T_x\mathcal{M}$, $\sigma(\vec{u},\vec{v})$ denotes the sectional curvature of the plane spanned by $\vec{u}$ and $\vec{v}$. The Ricci curvature can be interpreted as an average over all the sectional curvatures associated to the orthogonal planes $(\vec{u},\vec{e}_j)$. The Ricci curvature is known to satisfy the following *Bochner-Weitzenböck-Lichnerowicz formula* (see [Villani, 2008]): for any smooth enough

function $\Phi : \mathcal{M} \to \mathbb{R}$:

$$-\Delta \left(\frac{\|\nabla \Phi\|^2}{2} \right) + \nabla \Phi . \nabla(\Delta \Phi) + \mathrm{tr}\left(\nabla^2 \Phi \right)^2 + \mathrm{Ric}(\nabla \Phi) = 0 \,, \tag{3.24}$$

where $\vec{v} . \vec{w} \equiv g(\vec{v}, \vec{w})$, $\nabla^2 \Phi$ denotes the Hessian of Φ and Δ is the Laplace Beltrami operator associated with the manifold. This formula can be interpreted as a relation of commutation between second and third order differential operators of Φ. By a simple Cauchy Schwarz inequality, we derived the so-called *Bochner inequality*:

$$-\Delta \left(\frac{\|\nabla \Phi\|^2}{2} \right) + g(\nabla \Phi, \nabla \Delta \Phi) + \frac{(\Delta \Phi)^2}{n} + \mathrm{Ric}(\nabla \Phi) \leq 0 \,. \tag{3.25}$$

The connection with functional inequalities can be made by thinking of the Laplacian operator in the above inequality as the generator of the heat equation, whose invariant measure (assuming that $\mathcal{M}$ is compact) is the normalized Lebesgue measure on $\mathcal{M}$. More generally, given a C^2 function V on $\mathcal{M}$, let $\mu := e^{-V(x)} \mathrm{vol}(dx)$ be a probability measure on $\mathcal{M}$ that is absolutely continuous with respect to Lebesgue, and defined $L = \Delta - \nabla V . \nabla$ the generator of a Markov diffusion of invariant measure μ. Then, the so-called *generalized Bochner formula* holds:

$$-L \left(\frac{\|\nabla \Phi\|^2}{2} \right) + g\left(\nabla \Phi, \nabla L(\Phi) \right) + \frac{(L\Phi)^2}{N} + \mathrm{Ric}_{N,\mu}(\nabla \Phi) \leq 0 \,. \tag{3.26}$$

This equation is completely analogous to the Bochner inequality (3.25), where the Laplacian Δ has been replaced by the generator L for which μ is the invariant measure, and $N \geq n$. $\mathrm{Ric}_{N,\mu}$ is the so-called *generalized Ricci curvature* and is defined as follows:

$$\mathrm{Ric}_{N,\mu}(\vec{v}) := \left(\mathrm{Ric} + \nabla^2 V \right)(\vec{v}) - \frac{(\nabla V . \vec{v})^2}{N - n} \,.$$

This observation is what lead Bakry and Émery to their geometrical interpretation of all the functional inequalities that we reviewed above. Assuming that the generalized Ricci curvature is bounded below by $\kappa \|\nabla \Phi\|^2$, we arrive at the following *curvature dimension inequality*[1]

$$L \left(\frac{\|\nabla \Phi\|^2}{2} \right) - \nabla \Phi . \nabla L(\Phi) \geq \frac{(L\Phi)^2}{N} + \kappa \|\nabla \Phi\|^2. \tag{CD(κ, N)}$$

3.6. The Bakry-Émery condition

The curvature dimension inequality introduced at the end of last section can be extended in two ways. In this section, we follow the approach of [Bakry and Émery, 1985] (see also [Bakry, 1994]) where it is generalized to an inequality satisfied by the generator of a Markov process on an abstract probability space $(E, \mathcal{G}, \mu)$. In Chapter 4, we see a more geometric approach in terms of optimal transport of mass on metric spaces.

Let $(P_t)_{t \geq 0}$ be an $\mathbb{L}_2(\mu)$ Markov process, and denote by $(L_2, \mathrm{dom}(L_2))$ its corresponding generator. Here, we still assume the existence of a densely defined subalgebra $\mathcal{A}$ that satisfies Condition 2.2.1. We recall the definition of its corresponding carré du champ (2.17): for any

---

[1]One can actually show that the two bounds are equivalent [Villani, 2008].

$f, h \in \text{dom}(L_2)$:

$$\Gamma(f, h) := \frac{1}{2} \left(L_2(fh) - f L_2(h) - h L_2(f) \right).$$

In some sense, the operator Γ quantifies the deviation of L_2 from being a derivation operator. Next, we define the *carré du champ itéré* as follows: for any $f, h \in \mathcal{A}$:

$$\Gamma_2(f, h) = \frac{1}{2} \left(L \Gamma(f, h) - \Gamma(f, Lh) - \Gamma(h, Lf) \right). \tag{3.27}$$

Then, the semigroup $(P_t)_{t \geq 0}$ satisfies the *Bakry-Emery condition* if there exist $\kappa \in \mathbb{R}$ and $N > 1$ such that the following holds: for any $f \in \mathcal{A}$,

$$\Gamma_2(f, f) \geq \frac{(Lf)^2}{N} + \kappa \Gamma(f, f). \tag{BE(κ, N)}$$

The diffusive case : Assume that $E = \mathcal{M}$ is a Riemannian manifold, and consider the diffusion semigroup $(P_t)_{t \geq 0}$ of generator $L = \Delta - \nabla V . \nabla$. In this case, we saw in Equation (2.27) that the carré du champ reduced to the following: for $f, h \in \mathcal{A}$,

$$\Gamma(f, h) = \nabla f . \nabla h,$$

Moreover, a direct calculation provides the following expression for the carré du champ itéré: for all $f \in \mathcal{A}$:

$$\Gamma_2(f) = L \left(\frac{\|\nabla f\|^2}{2} \right) - \nabla f . \nabla (Lf).$$

Therefore, BE(κ, N) reduces to

$$L \left(\frac{\|\nabla f\|^2}{2} \right) - \nabla f . \nabla (Lf) \geq \frac{(Lf)^2}{N} + \kappa \|\nabla f\|^2.$$

This is noting but CD(κ, N).

The power of Bakry-Emery criterion is that it implies all the inequalities that we introduced in the previous section (see Proposition 6.1 of [Saloff-Coste, 1994]):

Proposition 3.6.1. *Let $(P_t)_{t \geq 0}$ be an $\mathbb{L}_2(\mu)$-Markov semigroup on a probability space $(E, \mathcal{G}, \mu)$, where μ is invariant, and assume that BE(κ, N) holds for some $\kappa > 0$ and $N > 1$. Then,*

- *Bochner-Licnérowicz-Weitzenböck formula: If $(P_t)_{t \geq 0}$ is reversible and ergodic, then $\lambda(L) \geq \frac{\kappa N}{N-1}$;*

- *Bakry-Emery theorem [Bakry and Émery, 1985]: If $(P_t)_{t \geq 0}$ is a reversible diffusion, then $\alpha_2(L) \geq \frac{\kappa N}{N-1}$;*

- *Sobolev inequality [Ledoux, 1992]: Under the same assumption, and if $N > 2$, then for any $f \in \text{dom}(L_2)$,*

$$\|f\|_{\mathbb{L}_{\frac{2N}{N-2}}(\mu)}^2 \leq \frac{4(N-1)}{(N-2)\kappa N} \left(\mathcal{E}(f, f) + \frac{(N-2)N\kappa}{4(N-1)} \|f\|_{\mathbb{L}_2(\mu)}^2 \right). \tag{3.28}$$

Even though we defined the Bakry-Émery condition in the case of a Markov semigroup defined on an abstract probability space, it has mostly been found for diffusions on smooth Riamannian

manifolds. Recently, an extension of the theory to discrete Markov chains was found by [Johnson, 2017].

3.7. Markov semigroups acting on groups

We end this chapter with a list of important examples of Markov semigroups acting on groups.

3.7.1. Finite groups

Given a finite group and a discrete time Markov chain of kernel $K(x, y) = k(x - y)$ and uniform invariant measure $\mu(\{x\}) = |G|^{-1}$, consider the kernel of the associated continuous time chain $(P_t)_{t \geq 0}$ defined by

$$k_t(x, y) = |G| \exp(t \, (\, K - \mathrm{id}_{\mathbb{L}_\infty(G)} \,))(x, y) \,. \tag{3.29}$$

By construction, this kernel is right-invariant.

The hypercube: Let $G = \mathbb{Z}_2^n$ and define the following classical Markov chain: for $i = 1, ..., n$, let e_i be the vector in $\mathbb{Z}_2^n$ with all coordinates 0 but the ith coordinate, which is set to be 1. Next, define a probability mass function k on $\mathbb{Z}_2^n$ by setting $k(0) = k(e_i) = 1/(n + 1)$ for $i = 1, .., n$, and $k(x) = 0$ otherwise. In words, at each time, the discrete-time chain jumps from one vertex to a neighboring one with probability $1/(n + 1)$, and stays where it was with same probability.

The strong logarithmic Sobolev constant and the spectral gap for this chain are known [Diaconis and Saloff-Coste, 1996a]:

$$2 \, \alpha_2(L^{\mathrm{Hyp}}) = \lambda(L^{\mathrm{Hyp}}) = \frac{2}{n + 1} \,.$$

This, together with the estimate (3.13), provides the essentially sharp bound

$$\tau(\varepsilon) \leq \frac{1 - \ln(2\varepsilon)}{\lambda(L)} + \frac{n + 1}{4} \ln \ln \left[\min_{x \in E} \mu(\{x\}) \right]^{-1} \,.$$

The finite circle: We now consider the simple random walk on $G = \mathbb{Z}_m$ with $m \geq 4$, of associated kernel $K(x, x \pm 1) = 1/2$ and uniform stationary measure. The spectral gap of the corresponding continuous time Markov chain $(P_t^{\mathrm{Cir}})_{t \geq 0}$ is given by the formula [Diaconis and Saloff-Coste, 1996a]

$$\lambda(L^{\mathrm{Cir}}) = 1 - \cos \frac{2\pi}{m} \,.$$

It was shown in [Diaconis and Saloff-Coste, 1996a] that $(P_t^{\mathrm{Cir}})_{t \geq 0}$ satisfies the following bound:

$$\|P_t^{\mathrm{Cir}} - \mathbb{E}_{\mu_G} : \mathbb{L}_2(\mathbb{Z}_m) \to \mathbb{L}_\infty(\mathbb{Z}_m)\|^2 \leq 2 \left(1 + \frac{\sqrt{5}m}{8\sqrt{\pi t}} \right) \exp \left(-\frac{16\pi^2 t}{5m^2} \right) + \frac{m + 1}{2} \, \mathrm{e}^{-2t} \,.$$

In particular, in the case $m \geq 5$, the above expression yields the following simpler bound for $t_\infty = 5m^2/16\pi^2$:

$$\|P_{t_\infty}^{\mathrm{Cir}} - \mathbb{E}_{\mu_G} : \mathbb{L}_2(\mathbb{Z}_m) \to \mathbb{L}_\infty(\mathbb{Z}_m)\|^2 \leq \mathrm{e} \,.$$

The logarithmic Sobolev constant $\alpha_2(L^{\mathrm{Circ}})$ satisfies:

$$\frac{2\pi^2}{m^2} \geq \frac{\lambda(L^{\mathrm{Circ}})}{2} \geq \alpha_2(L^{\mathrm{Circ}}) \geq \frac{8\pi^2}{25m^2} \geq \frac{2\lambda(L^{\mathrm{Circ}})}{25}\,.$$

The exact value of $\alpha_2(L^{\mathrm{Circ}})$ is not know for $m \geq 4$. For $m = 3$, $\alpha_2(L^{\mathrm{Circ}}) = (2\ln 2)^{-1}$.

Now, choose the uniform random walk of kernel $K(x,y) = 1/m$ for any $x,y \in \mathbb{Z}_m$. This is a special case of the Markov chain studied in Theorem A.1 of [Diaconis and Saloff-Coste, 1996a], for which the strong log-Sobolev constant α_2 was shown to be equal to

$$\alpha_2 = \frac{1 - 2/m}{\ln(m-1)}, \qquad \lambda = 1 - \frac{1}{m}. \tag{3.30}$$

Random transpositions: Here, let $G = S_n$ be the symmetric group on n symbols. Consider the discrete time chain kernel $K(\theta, \sigma) = 2/[n(n-1)]$ if $\theta^{-1}\sigma$ is a transposition, and 0 otherwise. In [Diaconis, 1988], it was shown that the spectral gap of the corresponding continuous time semigroup $(P_t^{\mathrm{Tra}})_{t \geq 0}$ is

$$\lambda(L^{\mathrm{Tra}}) = \frac{2}{n-1}\,.$$

Moreover, the following bound can be found in the same paper:

$$\|P_{t_\infty}^{\mathrm{Tra}} - \mathbb{E}_{\mu_G} : \mathbb{L}_2(\mathbb{Z}_m) \to \mathbb{L}_\infty(\mathbb{Z}_m)\|^2 \leq 1$$

for $t_\infty = n \ln n$, which implies by interpolation results that

$$\frac{1}{3n \ln n} \leq \alpha_2(L^{\mathrm{Tra}}) \leq \frac{1}{n-1}\,.$$

3.7.2. Compact Lie groups

Now, we recall some well-known estimates for the heat semigroup defined on various compact Lie groups, of associated right-invariant transition density $(k_t)_{t \geq 0}$ with respect to the Haar measure μ_G:

$$P_t(f)(g) = \int_G k_t(gh^{-1})\, f(h)\, d\mu_G(h)\,. \tag{3.31}$$

We recall that, in this case, $\|h \mapsto k_t(gh^{-1}) - 1\|_{\mathbb{L}_2(\mu_G)}$ does not depend on $g \in G$.

1-dimensional torus The heat semigroup on the circle has spectral gap $\lambda(L^{\mathrm{Heat}}) = 1$. Moreover, the following estimate was derived in example 1 of Section 3 of [Saloff-Coste, 1994]:

$$\|h \mapsto k_t(h) - 1\|_{\mathbb{L}_2(\mu_G)} \leq \sqrt{2 + \sqrt{\pi/2t}}\; \mathrm{e}^{-t}\,.$$

n-dimensional torus $(n > 1)$: The logarithmic Sobolev constant associated to the heat semigroup on the n dimensional torus $\mathbb{T}^n$ is known to achieve the bound $2\alpha_2(L^{\mathrm{Heat}}) = \lambda(L^{\mathrm{Heat}}) = 1$. In Theorem 5.3 of [Saloff-Coste, 1994], the following upper bound on its kernel (and in fact on the kernel of any uniformly elliptic generator) was found:

$$\|h \mapsto k_t(h) - 1\|_{\mathbb{L}_2(\mu_G)} \leq \exp\left(-t + \frac{1}{2}\ln\left(\frac{1}{2}n\ln n\right) + 6\right). \tag{3.32}$$

Matrix Lie groups: In [Saloff-Coste, 1994], precise estimates on the kernel of diffusion semigroups on various Riemannian manifolds were obtained starting from a curvature dimension inequality

$CD(\kappa, N)$. These estimates simply follow from Proposition 3.6.1 and Equation (3.20). Applying this to the curvature dimension inequalities satisfied by semi-simple Lie groups [Rothaus, 1986], Saloff-Coste derived the following straightforward corollary (stated here as a theorem for sake of completeness):

Theorem 3.7.1. *Let $(G, \mathfrak{g})$ be a real connected semi-simple compact Lie group of dimension n endowed with the Riemannian metric induced by its Killing form. Then, the heat diffusion satisfies*

$$\|h \mapsto k_t(h) - 1\|_{\mathbb{L}_2(\mu_G)} \leq \exp\left(1 + \lambda(\Delta)\left[-t + \frac{16}{n} + 2\ln\left(1 + \frac{1}{2}n\ln\frac{n}{4}\right)\right]\right), \tag{3.33}$$

where $\lambda(\Delta)$ is the spectral gap of $(G, \mathfrak{g})$. In particular, for $\varepsilon > 0$ and $t_n = 2(1 + \varepsilon)\ln n$, the above bound converges to 0. Moreover, the following bounds the logarithmic Sobolev and Poincaré constants hold:

$$\lambda(\Delta) \geq \frac{n}{8(n-1)}, \tag{3.34}$$

$$\alpha_2(\Delta) \geq \frac{n}{4(n-1)}. \tag{3.35}$$

Finally, the semigroup satisfies $\mathrm{Nash}(C, n, T)$ *with* $C = T = \frac{16(n-1)}{(n-2)n}$.

3.7.3. Hörmander systems:

More generally, given a Riemannian manifold $\mathcal{M}$, a Hörmander system on $\mathcal{M}$ is a set of vector fields $V = \{V_1, ..., V_m\}$ such that, at each point $p \in \mathcal{M}$, there exists an integer K such that the iterated commutators $[V_{i_1}, [V_{i_2}, ..., V_{i_k}]...]$, $k = 2, ..., K$, generate the tangent space $T_p\mathcal{M}$. Specializing to the case of a Lie group G, a Hörmander system $V = \{V_1, ..., V_m\}$ can more simply be defined as a set of vectors in the Lie algebra, i.e. the tangent space at the neutral element e, such that for some $K \in \mathbb{N}$ the iterated commutators of order at most K span the whole tangent space. For fixed $j \in [K]$ we find a geodesic $g_j(t)$ with $g_j(0) = e$ such that for any $f \in C^1(G)$

$$V_j(f)(h) = \frac{d}{dt}f(g_j(t)h)|_{t=0}.$$

This leads to the corresponding left invariant classical generator

$$L_V = \sum_j V_j^2. \tag{3.36}$$

The generator L_V generates a semigroup $P_t = e^{tL_V}$ on $\mathbb{L}_\infty(G)$. Since the semigroup commutes with the right action of the group it is implemented by a convolution kernel as in Equation (9.1):

$$P_t(f)(g) = \int k_t(gh^{-1}) f(h) \, d\mu_G(h).$$

and is reversible with respect to the Haar measure. For Hörmander systems, the following kernel estimates go back to the seminal work of Stein and Rothshield [Rothschild and Stein, 1976], see also [Lugiewicz and Zegarlinski, 2007].

Theorem 3.7.2. *Let $V = \{V_1, ..., V_m\}$ be a Hörmander system such that K iterated commutators span a Lie algebra of dimension d. Then L_V has a spectral gap and there exists $C_V > 0$ such that, for all $0 < t \leq 1$:*

$$\sup_{g \in G} k_t(g) \leq C_V \, t^{-Kd/2}.$$

The above kernel estimate can be related to estimates of the form of (3.33): indeed, since the

semigroup is reversible, $k_t(g) = k_t(g^{-1})$ for any $g \in G$, so that:

$$\|h \mapsto k_t(h)\|_{\mathbb{L}_2(\mu_G)}^2 = \int_G |k_t(h)|^2 \, d\mu_G(h) = \int_G k_t(eh^{-1}) \, k_t(he) d\mu_G(h) = |k_{2t}(e)| \le \sup_{g \in G} k_{2t}(g) \, .$$

where we used the semigroup property in the third identity above. On the other hand, it is simple to see that

$$\|h \mapsto k_t(h)\|_{\mathbb{L}_2(\mu_G)} = \|P_t : \mathbb{L}_2(\mu_G) \to \mathbb{L}_\infty(G)\| \, .$$

This observation combined with (3.7) provides bounds similar to those of Theorem 3.7.1 on the mixing time for such evolutions. One can also use the following more direct bound in order to get such estimates:

$$\|P_t : \mathbb{L}_1(\mu_G) \to \mathbb{L}_\infty(G)\| = \sup_{g \in G} k_t(g) \le C_V \, t^{-Kd/2} \, .$$

Indeed, for any measure $\nu \ll \mu_G$ of corresponding density f evolving according to $(P_t)_{t \ge 0}$, the following bound holds at short times:

$$\begin{aligned}
\|P_{t*}(\nu) - \mu_G\|_{\mathrm{TV}} &= \frac{1}{2} \|P_t(f) - 1\|_{\mathbb{L}_1(\mu_G)} \\
&\le \frac{1}{2} \|P_t(f) - 1\|_\infty \\
&\le \frac{1}{2} \|P_t : \mathbb{L}_1(\mu_G) \to \mathbb{L}_\infty(G)\| \, \|f - 1\|_{\mathbb{L}_1(\mu_G)} \\
&\le C_V \, t^{-Kd/2} \, .
\end{aligned}$$

Chapter 4.

Optimal transport

In Section 3.6, we introduced the Bakry-Émery criterion $BE(\kappa, N)$ as an generalization of the curvature-dimension inequality $CD(\kappa, N)$ to abstract probability spaces. However in practice, this condition almost exclusively applies to diffusion semigroups on smooth Riemannian manifolds. The impossibility to extend the proof of Bakry-Émery theorem (cf. Proposition 3.6.1) to nonsmooth situations, due to the lack of a chain rule in these settings, witnesses that fact. This in particular means that Bakry-Émery's approach completely breaks down in the discrete setting of Markov chains [Johnson, 2017]. The situation is even worse in the quantum setting, due to the apparent lack of an underlying sample space.

Fortunately, the theory reached another milestone with the appearance of articles establishing a connection between Bakry-Émery's criterion and the well established theory of optimal transport [McCann, 1997, Jordan et al., 1998, Otto and Villani, 2000a]. In these articles, proofs of already known functional inequalities purely based on optimal transport arguments became available. In the case when $E = \mathcal{M}$ is a Riemannian manifold, the idea is that the geometry of $\mathcal{M}$ should be related to the one of the set $\mathcal{P}(\mathcal{M})$ of probability measures on $\mathcal{M}$. More precisely, curvature-dimension inequalities should be understood in terms of the properties of convexity of some particular functionals defined on $\mathcal{M}$ along optimal paths. The mesoscopic notion of transportation of *masses* in $\mathcal{P}(\mathcal{M})$ being more robust than the local notion of transportation of *points* in $\mathcal{M}$, this new approach lead to a particularly prolific generalization of the CD criterion to non-smooth metric-measure length spaces. Since then, the original framework of Bakry and Émery was further extended to the case of Markov chains over discrete sets [Maas, 2011, Erbar and Maas, 2012] and, more recently, to finite dimensional quantum Markov semigroups [Carlen and Maas, 2014, Carlen and Maas, 2017]. Here, we only give the main notions of the theory and refer to [Villani, 2008] for further details.

The premises of optimal transport theory can be traced back to 1781's French pre-revolution area with the early work of [Monge, 1781]. In this essay, the mathematician was interested in the cheapest way of transporting soil from its extraction point to the place where it would be incorporated in a construction. The problem here is to find where each unit of soil should go in such a way as to minimize the total transportation cost. Studying the problem in three dimensions and for a continuous distribution of mass, Monge showed that transportation should follow straight lines that are orthogonal to a family of surfaces called lines of curvature. A few centuries later, Monge's problem was rediscovered in 1942 by the Russian Nobel laureate Leonid Vitaliyevich Kantorovich who developed the theory of linear programming arising from various optimization problems. More precisely, Kantorovich defined a particularly flexible notion of distance between probability measures, nowadays known as the *Kantorovich-Rubinstein distance* (also known as *Wasserstein distance*).

Layout of the chapter: In Section 4.1, we mention the Monge-Kantorovich problem and recall the definition of the commutative Wasserstein distance. The link between Bakry Émery's curvature-dimension inequality and displacement convexity of the entropy is drawn in Section 4.2. A more direct link between the latter and the functional inequalities of Chapter 3 is provided in Section 4.3. In Sections 4.4 and 4.5, we briefly explain the relationship to other geometric and information theoretic inequalities. We end this chapter with a discussion on the concentration of measure phenomenon in Section 4.6.

4.1. The Monge-Kantorovich problem

Monge's original problem can be reformulated in modern terms as follows: given two probability spaces $(\mathcal{X}, \mu)$, and $(\mathcal{Y}, \nu)$. A *coupling* between μ and ν is a random vector (X, Y) on some probability space $(\Omega, \mathbb{P})$ such that $X \sim \mu$, resp. $Y \sim \nu$. The couple (X, Y) (or its associated law) is then called a *coupling* of (μ, ν). A coupling is said to be *deterministic* if there exists a measurable function $T : \mathcal{X} \to \mathcal{Y}$ such that $Y = T(X)$.

Among the many possible ways of defining a coupling, we are interested in the *optimal coupling* (also called *optimal transport*). For this, we introduce a *cost function* $c : \mathcal{X} \times \mathcal{Y} \ni (x, y) \mapsto c(x, y)$ that models the work needed to move one *unit of mass* from location $x \in \mathcal{X}$ to location $y \in \mathcal{Y}$. Monge's original framework, which corresponds to looking at the subclass of deterministic couplings, can then be stated as the minimization of the following quantity:

$$\int_{\mathcal{X}} c\left(x, T(x)\right) \mu_0(dx),\qquad \text{(M)}$$

over mappings $T : \mathcal{X} \to \mathcal{Y}$ transporting a fixed mass μ_0 to another fixed mass $\mu_1 = T_* \mu_0$, where $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ is a given pointwise cost function. Here, T_* is the image measure, predual of T acting on probability measures, defined by

$$(T_*)\, \mu(A) = \mu(T^{-1}(A)).$$

More generally, the *Monge-Kantorovich* minimization problem is the one of the minimization of the quantity:

$$\mathbb{E}_\pi\left[c(X, Y)\right] \equiv \int_{\mathcal{X} \times \mathcal{Y}} c(x, y)\, d\pi(x, y),\qquad \text{(M-K)}$$

where the pair $(X, Y) \sim \pi$ runs over all possible couplings of (μ, ν).

The infimum in (M-K) is called the *optimal transport cost* between the probability measures μ and ν. Existence of an optimal coupling has been proved in great generality under some very mild assumptions on the spaces $(\mathcal{X}, \mathcal{Y})$ as well as on the cost function c (see Theorem 4.1 of [Villani, 2008]).

4.1.1. Kantorovich duality

Whereas the central object of the Monge-Kantorovich problem is cost, the central object in the dual formulation derived by Kantorovich is price: the *dual Kantorovich problem* is the one of the maximization of:

$$\int_{\mathcal{Y}} \phi(y)\, d\nu(y) - \int_{\mathcal{X}} \psi(x)\, d\mu(x),\qquad \text{(K)}$$

over functions $\psi \in \mathbb{L}^1(\mathcal{X}, \mu)$, $\phi \in \mathbb{L}^1(\mathcal{Y}, \nu)$ such that for any x, y, $\phi(y) - \psi(x) \le c(x, y)$. Here, the pairs (ψ, ϕ) should be interpreted as price functions: assume a company is offering to transport the shape μ into ν by buying a unit of mass at point x at a price of $\psi(x)$, and selling it at point y at a price of $\phi(y)$. Then the condition $\phi(y) - \psi(x) \le c(x, y)$ can be interpreted as the fact that the company's margin $\phi(y) - \psi(x)$ made by transporting a unit of mass from x to y should not locally exceed the cost $c(x, y)$. Kantorovich's duality theorem establishing the equivalence between (M-K) and (K) is shown under very weak regularity conditions in Theorem 5.10 of [Villani, 2008].

4.1.2. Wasserstein distances

It is natural to assume that the cost function depends on the distance between the extraction site and the construction site. In this case, assume that $\mathcal{X} = \mathcal{Y}$ is a Polish space and let $p \in [1, \infty)$. Then, for any two probability measures μ, ν on $\mathcal{X}$, the *Wasserstein distance of order p* between μ and ν is defined by the formula

$$W_p(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \left(\int_{\mathcal{X}} d(x, y)^p d\pi(x, y) \right)^{1/p},$$

where $\Pi(\mu, \nu)$ denotes the set of couplings between μ and ν. On can show that W_p does satisfies the axioms of a distance, up to the fact that it might take the value $+\infty$, on the set of all probability measures. In the case $p = 1$, W_1 is also known as *Kantorovich-Rubinstein* distance:

$$W_1(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{X}} d(x, y) \, d\pi(x, y) = \sup_{\|\psi\|_{\mathrm{Lip}} \le 1} \int_{\mathcal{X}} \psi \, d\mu - \int_{\mathcal{X}} \psi \, d\nu,$$

where the equality follows from Kantorovich duality. Perhaps the most known example of a W_1 distance is the one when $d(x, y) = \mathbb{1}_{x \ne y}$. The associated Wasserstein distance is the *total variation*:

$$\|\mu - \nu\|_{\mathrm{TV}} = 2 \inf_{(X, Y)} \mathbb{P}(X \ne Y) = \sup_{A \in \mathcal{B}(\mathcal{X})} |\mu(A) - \nu(A)|,$$

where the infimum is over all couplings (X, Y) of μ and ν, and $\mathcal{B}(\mathcal{X})$ denotes the Borel σ-algebra associated to $\mathcal{X}$.

The *Wasserstein space of order p* is the space $\mathcal{P}_p(\mathcal{X})$ of probability measures such that for some $x_0 \in \mathcal{X}$ (and therefore any $x_0 \in \mathcal{X}$):

$$\int_{\mathcal{X}} d(x_0, x)^p d\mu(x) < +\infty.$$

It is easy to verify that W_p is finite on $\mathcal{P}_p(\mathcal{X})$ and turns it into a Polish space (Theorem 6.18 of [Villani, 2008]). Moreover, a use of Hölder's inequality directly reveals the following hierarchy of Wasserstein distances:

$$1 \le p \le q \quad \Rightarrow \quad W_p \le W_q.$$

Hence, W_1 is the weakest distance. In the next sections, we will mostly focus on the Wasserstein distance of order 2. In particular, we will see that it can be reinterpreted as a Riemannian metric, hence endowing the space $\mathcal{P}_2(\mathcal{X})$ with a formal Riemannian structure.

4.2. From curvature-dimension inequality to displacement convexity

In this section, we clarify the role played by optimal transport in the study of the curvature-dimension inequality $\mathrm{CD}(\kappa, N)$. In a nutshell, the qualitative properties of the optimal transport occurring on a manifold are influenced by its curvature, and conversely. This fundamental equivalence between the notion of displacement convexity, that we introduce below, and Ricci curvature lower bound is of paramount importance. Many of the properties of a Riemannian manifold arising from the curvature-dimension inequality (and in particular the functional inequalities that we mainly care about) can be re-derived in the framework of optimal transport. We will see that Markovian evolutions naturally arise as certain *gradient flows* in the Wasserstein space of order 2. This shift of paradigm is what allows the definition of analogous objects in the classical discrete, and quantum cases where the lack of an underlying Riemmanian manifold structure marks the absence of a curvature-dimension inequality.

4.2.1. Benamou-Brenier dynamical formulation

In Section 4.1, we introduced two formulations of the Monge-Kantorovich problem, namely, the Monge-Kantorovich problem (M-K), and the dual Kantorovich problem (K). As mentioned, these two formulations are equivalent under some regularity conditions on the underlying cost function. These conditions hold in particular in the case of the Wasserstein distance: let $\mathcal{M}$ be a smooth complete Riemannian manifold without boundary. We recall the Monge-Kantorovich formulation of the Wasserstein distance of order 2: for any $\mu, \nu \in \mathcal{P}_2(\mathcal{M})$,

$$W_2(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \left(\int_{\mathcal{M} \times \mathcal{M}} d(x, y)^2 d\pi(x, y) \right)^{1/2}.$$

The Kantorovich dual problem writes

$$W_2(\mu, \nu)^2 = \sup \ \int_{\mathcal{M}} f \, d\mu - \int_{\mathcal{M}} g \, d\nu \,,$$

where the supremum is taken over functions $f \in \mathbb{L}^1(\mathcal{M}, \mu)$ and $g \in \mathbb{L}^1(\mathcal{M}, \nu)$ such that fur any $x, y \in \mathcal{M}$, $f(x) - g(y) \le d(x, y)^2$.

There exists yet another approach to the optimal transport problem, that is traditionally referred to as the *dynamical formulation*, also known as the *Benamou-Brenier formula*. It was formally stated by [Benamou and Brenier, 2000] in the case of $\mathcal{M} = \mathbb{R}^d$, and since then proved with increasing degree of rigor and generality. Let $(\mathcal{M}, g)$ be a smooth compact Riemannian manifold without boundary. Then $(\mathcal{P}_2(\mathcal{M}), W_2)$ is geodesic (Theorem 3.10 of [Ambrosio and Gigli, 2013])[1]. We now endow $(\mathcal{P}_2(\mathcal{M}), W_2)$ with a *weak Riemannian structure*. For this, we need to introduce the notion of an *absolutely continuous curve*: a curve $(\mu_t)_{t \in [0,1]}$ is said to be absolutely continuous if there exists an integrable function f such that

$$\forall t < s \in [0, 1], \quad d(\mu_t, \mu_s) \le \int_t^s f(r) dr \,, \tag{4.1}$$

where d denotes the Riemannian distance associated to the metric g. Given such a geodesic curve, its

---

[1] In fact, this still holds for Polish geodesic spaces

derivative $|\dot{\mu}_t|$ is almost everywhere defined, and is given by

$$|\dot{\mu}_t| = \lim_{h \to 0} \frac{d(\mu_{t+h}, \mu_t)}{h}, \tag{4.2}$$

so that $t \mapsto |\dot{\mu}_t|$ is integrable on $[0,1]$ and is the smallest function f satisfying (4.1). Then (see Theorem 2.29 and Proposition 2.30 of [Ambrosio and Gigli, 2013]):

Theorem 4.2.1. *Let $\mathcal{M}$ be a smooth complete Riemannian manifold without boundary. Then for every absolutely continuous curve $(\mu_t)_{t \in [0,1]}$ on $\mathcal{P}_2(\mathcal{M})$, there exists a Borel family of vector fields $(v_t)_{t \in [0,1]}$ on $\mathcal{M}$ such that*

$$\|v_t\|_{\mathbb{L}_2(\mu_t, T_{\mu_t}\mathcal{P}_2(\mathcal{M}))} := \sqrt{\int_{\mathcal{M}} |v_t|^2 \, d\mu_t} \leq |\dot{\mu}_t| \quad a.e.$$

and the following continuity equation *holds in the sense of distributions:*

$$\frac{d}{dt}\mu_t + \nabla.(v_t \, \mu_t) = 0. \qquad \text{(continuity equation)}$$

Now, for any $\mu, \nu \in \mathcal{P}_2(\mathcal{M})$,

$$W_2(\mu, \nu) = \inf \left(\int_0^1 \|v_t\|^2_{\mathbb{L}_2(\mu_t, T_{\mu_t}\mathcal{P}_2(\mathcal{M}))} \, dt \right)^{\frac{1}{2}}, \qquad \text{(Benamou-Brenier)}$$

where the infimum is taken over weakly continuous distributional solutions (μ_t, v_t) of the continuity equation.

By Theorem 4.2.1, W_2 induces a *weak Riemannian structure* on $\mathcal{P}_2(\mathcal{M})$, where the scalar product in $\mathbb{L}_2(\mu, T_\mu \mathcal{P}_2(\mathcal{M}))$ should be thought of as the associated metric tensor at the point $\mu \in \mathcal{P}_2(\mathcal{M})$. This intuition was first formulated by [Otto, 2001], and the associated formal calculus that we briefly sketch in the next subsection is therefore often referred to as *Otto calculus*. One problem is that given a curve $(\mu_t)_{t \in [0,1]}$, the associated vector field $(v_t)_{t \in [0,1]}$ is not uniquely defined through the continuity equation. Indeed, for any vector field $(w_t)_{t \in [0,1]}$ with $\nabla.(w_t \mu_t) = 0$, the continuity equation is still satisfied by $(v_t + w_t)_{t \in [0,1]}$. Hence, we choose $(v_t)_{t \in [0,1]}$ to be the a.e. unique such family with minimal norm for a.e. $t \in [0,1]$. This is equivalent to asking that almost everywhere, v_t belongs to the set

$$\overline{\{\nabla \varphi : \; \varphi \in C_c^\infty(\mathcal{M})\}}^{\mathbb{L}^2(\mu_t, T_{\mu_t}\mathcal{P}_2(\mathcal{M}))}$$
$$= \left\{ v \in \mathbb{L}_2(\mu_t, T_{\mu_t}\mathcal{P}_2(\mathcal{M})) : \; \int \langle v, w \rangle \, d\mu_t = 0, \; \forall w \in \mathbb{L}^2(\mu_t, T_{\mu_t}\mathcal{P}_2(\mathcal{M})) \text{ s.t. } \nabla.(w\mu_t) = 0 \right\}, \tag{4.3}$$

the equality coming from the usual reasoning that $A^{\perp\perp} = \overline{\operatorname{span}(A)}$ for any subset A of a Hilbert space $\mathcal{H}$. By analogy with Riemannian geometry, the space defined by (4.3) is referred to as the *tangent space* $T_\mu \mathcal{P}_2(\mathcal{M})$ at $\mu \in \mathcal{P}_2(\mathcal{M})$.

4.2.2. Otto calculus

The advantage of the Benamou-Brenier formulation lies in the possibility to carry out formal calculations on the weak Riemannian structure that it defines. This idea was first investigated by [Jordan et al., 1998, Otto, 2001]. In words, the *Otto calculus* [Otto, 2001] that emerged from [Jordan et al., 1998] mainly served as a formal, yet very powerful, heuristic tool that served to see the connection between Markov processes and the notion of *gradient flows*, as well as the equivalence between the Ricci-curvature

lower bound $\mathrm{CD}(\kappa, N)$ and the so-called *geodesic convexity* of certain entropic functionals. Here, we follow the treatment of [Villani, 2008].

Given a Riemmanian manifold $(\mathcal{M}', g')$, the *gradient flow* associated to a smooth function $f : \mathcal{M}' \to \mathbb{R}$ is the flow induced by the differential equation

$$\frac{d\gamma_t}{dt} = -\operatorname{grad}_{\gamma_t} f \,,$$

where we refer to Section 3.5 for the definition of the gradient $\operatorname{grad}_x f$ of f at the point $x \in \mathcal{M}'$. [Jordan et al., 1998] discovered that many important partial differential equations on a manifold $(\mathcal{M}, g)$ can be formulated as gradient flows in the weak Riemannian manifold $(\mathcal{M}', g') = (\mathcal{P}_2(\mathcal{M}), W_2)$. Perhaps the easiest example of this nature is the heat equation

$$\frac{\partial \rho}{\partial t} = \Delta \rho$$

that is a flow in $(\mathcal{P}_2(\mathcal{M}), W_2)$ with respect to the entropic functional $f : \rho \mapsto \int \rho \log \rho$.

More precisely, Otto calculus consists of rules for formally differentiating functionals on $(\mathcal{P}_2(\mathcal{M}), W_2)$. Let vol be the volume measure on $\mathcal{M}$, and let $\mu(dx) := \exp(-V(x)) \operatorname{vol}(dx)$ denote a measure on $\mathcal{M}$ which is absolutely continuous with respect to vol. The density $\exp(-V)$ of μ with respect to vol is such that $V : \mathcal{M} \to \mathbb{R}$ is regular enough for the following expressions to make sense. As we already saw in Section 3.5 the measure μ is invariant under the semigroup generated by the second order differential operator $L := \Delta - \nabla V . \nabla$. Then, for any twice differentiable function $\phi : (0, \infty) \to \mathbb{R}$, define the following functional on the set of probability measures $\nu \ll \mu$, $\nu = \rho \mu$:

$$\phi_\mu(\nu) := \int_{\mathcal{M}} \phi(\rho(x)) \, d\mu(x) \,.$$

The function ϕ can be thought of as a "local energy" for a fluid of mass density given by ρ. Hence, it is natural to assume that $\phi(0) = 0$. On the other hand, ϕ_μ should be understood as a total amount of energy contained in the fluid. Pursuing the analogy to fluid dynamics, we further define the local *pressure* and *iterated pressure* as

$$p(\rho) := \rho \phi'(\rho) - \phi(\rho) \,, \qquad p_2(\rho) = \rho p'(\rho) - p(\rho) \,.$$

Then, for any such functional ϕ_μ, one can derive at least formally the following two formulas:

(i) *Gradient formula in Wasserstein space*:

$$\operatorname{grad}_\nu \phi_\mu = -\nabla . (\nu \nabla \phi'(\rho)) = -L(p(\rho)) \mu \,;$$

(ii) *Hessian formula in Wasserstein space*: for any tangent vector $\dot\nu = -\nabla . (\nu \nabla \psi)$ at ν,

$$\operatorname{Hess}_\nu \phi_\mu(\dot\nu) = \int_{\mathcal{M}} \left[\|\nabla^2 \psi\|_{\mathrm{HS}}^2 + (\operatorname{Ric} + \nabla^2 V)(\nabla \psi) \right] p(\rho) \, d\nu + \int_{\mathcal{M}} (L\psi)^2 p_2(\rho) \, d\mu \,. \qquad (4.4)$$

Example 4.2.2. In the important case when $\phi(\rho) = \rho \log \rho$ and $V = 0$, we recover the celebrated *Boltzman H-functional* that will play a crucial role in the next sections:

$$H(\nu) := \int_{\mathcal{M}} \rho \log \rho \, d\operatorname{vol} \,. \qquad (4.5)$$

More generally, for a general volume $\mu = e^{-V} \operatorname{vol}$, the H-functional relative to the reference measure μ

is given by

$$H_\mu(\nu) := \int_\mathcal{M} \rho \log \rho \, d\mu \,.$$

This is nothing but the relative entropy $D(\nu\|\mu)$ between ν and μ. In this case, one can easily verify that $p(\rho) = \rho$ and $p_2(\rho) = 0$, so that the gradient formula simplifies to

$$\operatorname{grad} H_\mu = -L(\rho)\mu \,, \tag{4.6}$$

which suggests that the Markov process of generator L is the gradient flow of H_μ. The statement can be made more rigorous (see e.g. Corollary 23.23. of [Villani, 2008]).

Moreover, given $\nu << \mu$, $\rho = d\nu/d\mu$, and $\dot\nu$ of corresponding tangent vector $v = \nabla\varphi$ through the continuity equation,

$$\operatorname{Hess}_\nu H_\mu(\dot\nu) = \int_\mathcal{M} \Gamma_2(\varphi,\varphi)\,\rho\,d\mu \tag{4.7}$$

which suggests that the condition $\mathrm{CD}(\kappa, N)$ should be equivalent to the one of κ-convexity of the Boltzmann H_μ functional along geodesics. This is indeed the case, as we see in next section.

4.2.3. Displacement convexity

The intuition provided by the Otto calculus can be made rigorous: given a vector space $\mathcal{V}$, a function $f : \mathcal{V} \to \mathbb{R} \cup \{\infty\}$ is said to be convex if for all $t \in [0, 1]$ and any $x, y \in \mathcal{V}$,

$$F((1-t)\,x + t\,y) \le (1-t)F(x) + tF(y) \,.$$

This notion can be extended to the case of functions on metric spaces: Let $(\mathcal{X}, d)$ be a complete geodesic space. Then a function $F : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is said to be *geodesically convex* if for any constant-speed geodesic path $(\gamma_t)_{t\in[0,1]}$ valued in $\mathcal{X}$, and any $t \in [0, 1]$,

$$F(\gamma_t) \le (1-t)F(\gamma_0) + tF(\gamma_1) \,. \tag{4.8}$$

In the case when $\mathcal{X} \equiv \mathcal{P}_2^{\mathrm{ac}}(\mathcal{M})$ is the space of measures that are absolutely continuous with respect to the volume vol on a smooth complete Riemannian manifold $\mathcal{M}$ without boundary, any functional F satisfying Equation (4.8) is said to be *displacement convex*. In this context, F is moreover said to be κ-displacement convex for some $\kappa \in \mathbb{R}$ if, whenever $(\mu_t)_{t\in[0,1]}$ is a constant speed geodesic in $\mathcal{P}^{\mathrm{ac}}(\mathcal{M})$, and for any $t \in [0, 1]$,

$$F(\mu_t) \le (1-t)F(\mu_0) + tF(\mu_1) - \frac{\kappa\,t\,(1-t)}{2}\,W_2(\mu_0,\mu_1)^2 \,. \tag{4.9}$$

The next theorem, which can be found in Corollary 17.19 of [Villani, 2008], provides the promised equivalence between the curvature-dimension inequality and κ-convexity of Boltzmann's H functional. A more general statement can be found in [Villani, 2008], where $\mathrm{CD}(\kappa, N)$ is showed to be equivalent to the κ-convexity of a certain class of entropy-like functionals.

Theorem 4.2.3. *Let $(\mathcal{M}, g)$ be a Riemannian manifold and $\kappa \in \mathbb{R}$. Then $\mathcal{M}$ satisfies $\mathrm{CD}(\kappa, \infty)$ if and only if Boltzmann's H functional is κ-displacement convex on $\mathcal{P}_c^{\mathrm{ac}}(\mathcal{M})$.*

The advantage of the displacement convexity approach to $\mathrm{CD}(\kappa, N)$ lies in the fact that it does not rely on analytic computations (e.g. of the Ricci tensor). This fact allowed several authors

in the mid 00's to study geometric properties of non-smooth spaces. [Maas, 2011] further extended this approach to the case of continuous time Markov chains. In the next section, we clarify the link between displacement convexity and the functional inequalities of Chapter 3.

4.3. Functional inequalities retrieved

In the last section, we mentioned that Bakry-Émery's curvature-dimension inequality can be converted into the displacement convexity of a particular class of entropy-like functionals. In the case when $N = \infty$, it states that "the graph of Boltzmann's H functional lies below its chord". Extending the analogy, one can ask for an infinitesimal version of this condition, that is, "the graph of H lies above its tangent". This is the object of the HWI inequalities, that we state in the special case of $\mathrm{CD}(\kappa, \infty)$ (for a more general statement, see Theorem 20.10 of [Villani, 2008]).

Theorem 4.3.1 (HWI inequality). *Let $\mathcal{M}$ be a Riemannian manifold equipped with a reference measure $\mu = \mathrm{e}^{-V} \mathrm{vol} \in \mathcal{P}_2(\mathcal{M})$, $V \in C^2(\mathcal{M})$, satisfying the curvature-dimension bound $\mathrm{CD}(\kappa, \infty)$, $\kappa \in \mathbb{R}$. Then, for any $\nu \in \mathcal{P}_2(\mathcal{M})$:*

$$D(\nu \| \mu) \equiv H_\mu(\nu) \leq W_2(\mu, \nu) \sqrt{I_\mu(\nu)} - \kappa \frac{W_2(\mu, \nu)^2}{2}, \qquad (\mathrm{HWI}(\kappa))$$

where we recall that the Fisher information I_μ is defined as follows: given $\nu = \rho\mu$,

$$I_\mu(\nu) := \int_{\mathcal{M}} \frac{|\nabla \rho|^2}{\rho} d\mu \equiv \mathrm{EP}_L(\nu).$$

The $\mathrm{HWI}(\kappa)$ inequality can be interpreted as a nonlinear interpolation inequality: the Kullback Leibler divergence H is partially controlled by the Fisher information I (stronger, since it requires smoothness conditions through the gradient in its expression) and the Wasserstein distance W (weaker). It was first established in [Otto and Villani, 2000b] in the above case when $N = \infty$ [Cordero-Erausquin, 2002, Bobkov et al., 2001]. The case $N < \infty$ was first shown in [Lott and Villani, 2009, Lott and Villani, 2007]. From the $\mathrm{HWI}(\kappa)$ inequality, one can retrieve the logarithmic Sobolev inequality:

Theorem 4.3.2 (Bakry-Émery theorem, retrieved). *Let $\mathcal{M}$ be a Riemannian manifold equipped with a reference probability $\mu = \mathrm{e}^{-V} \mathrm{vol}$, $V \in C^2(\mathcal{M})$, satisfying $\mathrm{CD}(\kappa, \infty)$ for some $\kappa > 0$. Then μ satisfies the logarithmic Sobolev inequality $\mathrm{MLSI}(\kappa/2)$, i.e. for any $\nu << \mu$:*

$$2\kappa \, H_\mu(\nu) \leq I_\mu(\nu).$$

Proof. By the $\mathrm{HWI}(\kappa)$ inequality, the result follows from an application of Young's inequality:

$$xy \leq cx^2 + \frac{1}{4c}y^2, \qquad \forall x, y \in \mathbb{R}, c > 0,$$

in which we set $x = W_2(\mu, \nu)$, $y = \sqrt{I_\mu(\nu)}$, and $c = \frac{\kappa}{2}$. □

The original proof of Bakry and Émery used the second variation of information along the heat semigroup. As expected, the logarithmic Sobolev inequality does not contain any information on the dimension. Dimension-dependent inequalities like Soblev inequalities can also be retrieved (see [Villani, 2008]).

Extension to continuous time Markov chains In the last sections, we introduced the notion of displacement convexity, equivalent to $\mathrm{CD}(\kappa, N)$, from which one can directly recover functional

inequalities. The advantage of this reformulation in terms of optimal transportation metrics lies in that it can be extended to non-smooth geodesic spaces E (see Part III of [Villani, 2008] for more details). However, this framework does not apply if the Wasserstein distance W_2 over E does not contain geodesics. This is in particular the case for Markov chains defined on finite sample spaces. In fact, since the metric derivative of the heat flow in the Wasserstein metric is typically infinite in a discrete setting, the heat flow cannot be interpreted as the gradient flow of any functional with respect to W_2. This is what lead [Maas, 2011] to introduce a modification of W_2 with respect to which the whole theory can be recovered [Erbar and Maas, 2012]. Interestingly enough there are situations where the modified Wasserstein distance defined in [Maas, 2011] for Markov chains acting on grids still converges to W_2 in the limit of small mesh (see [Gigli and Maas, 2013]).

In the remaining of this chapter, we provide a quick overview of some important geometric, information theoretic and concentration inequalities that are directly related to the functional inequalities as well as the curvature-dimension inequality introduced before. We refer to [Gardner, 2002] and [Villani, 2008] for more details.

4.4. Brunn-Minkowski and isoperimetric inequalities

The discovery of the equivalence between displacement convexity of a certain class of functionals and the curvature-dimension inequality has lead to new proofs and extensions of celebrated geometric and information theoretic inequalities. Among those inequalities, perhaps the most famous one is the isoperimetric inequality.

The Euclidean case The classical *Brunn-Minkowski inequality* in $\mathbb{R}^n$, here stated in its general form [Lusternik, 1935, Hadwiger and Ohmann, 1956], asserts that for $0 < \lambda < 1$ and any two nonempty measurable sets A_0 and A_1 in $\mathbb{R}^n$ such that their Minkowski sum $\lambda A_0 + (1-\lambda) A_1 := \{\lambda x_0 + (1-\lambda) x_1;\ x_0 \in A_0,\ x_1 \in A_1\}$ is also measurable,

$$\operatorname{Leb}\left(\lambda A_0 + (1-\lambda) A_1\right)^{\frac{1}{n}} \geq \lambda \operatorname{Leb}\left(A_0\right)^{\frac{1}{n}} + (1-\lambda) \operatorname{Leb}\left(A_1\right)^{\frac{1}{n}}, \tag{B-M}$$

where Leb stands for the Lebesgue measure. The original inequality was proved by Brunn around 1887 for $n = 3$ and in the case of convex bodies, and corrected a little later after Minkowski pointed out an error in the original proof. The equality case was shown to hold if and only if A_0 and A_1 are equal up to translation and dilatation by Brunn and Minkowski.

Taking A_1 to be equal to the ball $B_0(\varepsilon)$ centered at 0 and of radius $\varepsilon > 0$ and letting $\varepsilon \to 0$, the B-M inequality implies the *Euclidean isoperimetric inequality*: for any bounded open set $\Omega \subset \mathbb{R}^n$ with Lipschitz boundary $\partial \Omega$

$$\frac{\operatorname{Leb}(\partial \Omega)}{\operatorname{Leb}(\Omega)^{\frac{n}{n-1}}} \geq \frac{\operatorname{Leb}(\partial B)}{\operatorname{Leb}(B)^{\frac{n}{n-1}}}, \tag{Isop}$$

for any ball B, with equality if and only if Ω itself is a ball. This is more commonly known in its simplest version on $\mathbb{R}^2$: $L^2 \geq 4\pi A$, where A is the area of a domain enclosed by a curve of length L. In words, it states that a domain of fixed parameter and maximal area must be a disk.

Even before Hadwiger and Ohmann's proof of the Brunn-Minkowski inequality, [Henstock and Macbeath, 1953] showed an alternative proof that lead to the following functional inequality:

Theorem 4.4.1 ([Prékopa, 1973, Leindler, 1972]). *Let $0 < \lambda < 1$, and f, g, h nonnegative integrable*

functions on $\mathbb{R}^n$ satisfying

$$\forall x, y \in \mathbb{R}^n, \quad h\big((1-\lambda)\,x + \lambda\,y\big) \geq f(x)^{1-\lambda}\,g(y)^\lambda.$$

Then

$$\|h\|_1 \geq \|f\|_1^{1-\lambda}\,\|g\|_1^\lambda. \tag{P-L}$$

The Prékopa-Leindler inequality provides a simple functional analytical proof of the B-M inequality by simply taking $f = \mathbb{1}_{A_0}$, $g = \mathbb{1}_{A_1}$ and $h = \mathbb{1}_{\lambda A_0 + (1-\lambda)A_1}$. The latter can also be derived from a Sobolev inequality (see Theorem 8.2 as well as the discussion following Theorem 7.1 of [Gardner, 2002]). The Prekopa Leibler inequality has since then been used to prove concentration inequalities [Maurey, 1991], logarithmic Sobolev inequalities as well as other functional inequalities [Bobkov and Ledoux, 2000].

Extension to curved and non-smooth geometries   [McCann, 1997] was the first to provide a proof of the B-M inequality via displacement convexity, therefore establishing a clear connection with optimal transport. The extension to curved geometries requires the correct generalization of Minkowski sums. This was first done by [Cordero-Erausquin, 1999] where a Prékopa-Leindler inequality was obtained on the sphere. The general Prékopa-Leindler inequality in curved geometry was rigorously obtained in [Cordero-Erausquin et al., 2001], and its proof was adapted to get the Brunn-Minkowski inequality for general reference measures in [Sturm, 2006a] (see also Theorem 19.16 and Theorem 19.18 of [Villani, 2008] for a proof of Prékopa-Leindler's inequality using displacement convexity):

Theorem 4.4.2 (Curved Brunn-Minkowski inequality)**.** *If $\mathcal{M}$ satisfies $\mathrm{CD}(0, N)$, then for any X, Y two compact subsets of $\mathcal{M}$,*

$$\nu\big(\mathrm{mid}(X, Y)\big)^{1/N} \leq \frac{1}{2}\big(\mathrm{vol}(X)^{1/N} + \mathrm{vol}(Y)^{1/N}\big),$$

where for each point $x \in X$ and $y \in Y$, $\mathrm{mid}(x, y)$ is the middle point of any geodesic relating x and y.

This is exactly what having a positive curvature morally means: the volume of a compact set constructed as the set of middle points between two sets can never be smaller than the geometric sum of the volumes of the end sets. For a generalization to manifolds satisfying $\mathrm{CD}(\kappa, N)$, see Theorem 18.5 of [Villani, 2008]. The curved version of the isoperimetric inequality, originally proved by Lévy in 1919 for Riemannian manifolds of strictly positive sectional curvature embedded in $\mathbb{R}^n$, and generalized/repaired by Gromov under a unique hypothesis on the Ricci curvature, was very recently generalized to the unsmooth case by optimal transport arguments [Cavalletti and Mondino, 2017]. The curvature-dimension inequality is also known to imply *Cheeger's isoperimetric inequality*: there exists $\kappa > 0$ such that for any Borel subset Ω of $\mathcal{M}$,

$$\nu(\Omega) \leq \frac{1}{2} \quad \Rightarrow \quad \nu(\partial\Omega) \geq \kappa\,\nu(\Omega). \tag{Cheeger}$$

Cheeger's inequality in turn implies $\mathrm{PI}(\lambda)$ [Cheeger, 1969, Milman, 2009a]. We end this section by mentioning the following bound on the diameter of a Riemannian manifold $\mathcal{M}$ satisfying $\mathrm{CD}(\kappa, N)$:

Theorem 4.4.3 (Bonnet-Myers inequality)**.** *If $\mathrm{CD}(\kappa, N)$ holds, for $\kappa > 0$ and $N < \infty$, then $\mathcal{M}$ is*

necessarily compact and

$$\text{diam}(\mathcal{M}) \le \pi \sqrt{\frac{N-1}{K}}\,. \tag{4.10}$$

For other isoperimetric inequalities in (possibly non-smooth) metric measure spaces, we refer to [Cavalletti and Mondino, 2017, Sturm, 2006a, Sturm, 2006b, Villani, 2008] and the references therein. Isoperimetric inequalities in the context of Markov chains on finite sample spaces with lower bounded modified Ricci curvature have also been recently investigated in [Erbar and Fathi, 2018].

4.5. Entropy power inequality and entropic isoperimetry

Prékopa-Leindler's inequality can itself be seen as a special case of a reverse sharp Young's inequality for norms of convolutions, proved independently in [Beckner, 1975] and [Brascamp and Lieb, 1976a].

Theorem 4.5.1 (Sharp Young's inequalities)**.** *Let* $1 \le p, q, r < \infty$ *be such that* $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, *and let* $f \in \mathbb{L}_p(\mathbb{R}^n)$, $g \in \mathbb{L}_q(\mathbb{R}^n)$. *The following inequality holds:*

$$\|f * g\|_{\mathbb{L}_r} \le \frac{C_p C_q}{C_r} \|f\|_{\mathbb{L}_p} \|g\|_{\mathbb{L}_q}\,, \tag{S-Y}$$

where $C_m = \left(\frac{m^{1/m}}{\hat{m}^{1/\hat{m}}} \right)^{1/2}$ *for any Hölder conjugate* $m, \hat{m}$. *In the case when* $p, q, r \le 1$, *the inequality is reversed:*

$$\|f * g\|_{\mathbb{L}_r} \ge \frac{C_p C_q}{C_r} \|f\|_{\mathbb{L}_p} \|g\|_{\mathbb{L}_q}. \tag{RS-Y}$$

Both inequalities are saturated by Gaussian densities.

The proof of Brascamp and Lieb is quite involved, as they actually prove a much more general theorem, called the *Brascamp Lieb inequality*, from which Theorem 4.5.1 appears to be a special case. For a more recent proof of Brascamp Lieb inequality from optimal transport, we refer to [Barthe, 1997].

We mention that in the context of general metric measured spaces, a Brascamp-Lieb inequality was recently shown to be equivalent to a particular subadditivity of the entropy [Carlen and Cordero-Erausquin, 2009]. This result was further extended to the non-commutative scenario in [Carlen, 2008]. Brascamp and Lieb discovered that the limiting case $r \to 0$ of the reverse Young inequality implies an improved version of the Prékopa-Leindler inequality [Brascamp and Lieb, 1976a, Brascamp and Lieb, 1976b].

Interestingly, [Lieb, 1978] showed that the limit $r \to 1$ provides yet another famous inequality coming from information theory, called the *entropy power inequality*, which is equivalent to Gross' Gaussian logarithmic Sobolev inequality (see Section 4.5): given a random variable X on $\mathbb{R}^n$, with law absolutely continuous with respect to the Lebesgue measure, define its *entropy power* as

$$N(X) := e^{2S(X)/n}\,.$$

Theorem 4.5.2 (Entropy power inequality)**.** *Let* X *and* Y *be two independent random variables in* $\mathbb{R}^n$, *of respective laws absolutely continuous with respect to the Lebesgue measure. Then the random variable* $X + Y$ *admits a density given by the multidimensional convolution of the densities of* X *and*

Y, and:

$$N(X + Y) \geq N(X) + N(Y).$$ (EPI)

Equivalently, for any $\lambda \in [0,1]$, and any independent random variables X and Y taking values in $\mathbb{R}^n$, the entropy convex combination inequality holds:

$$S(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \geq \lambda S(X) + (1-\lambda)S(Y).$$ (ECCI)

The inequality was proposed by [Shannon, 1948], and a first proof was obtained a decade later by [a.J. Stam, 1959]. Stam's proof was later simplified by [Blachman, 1965]. Since then, various proofs, refinements and generalizations of EPI have been studied [Courtade, 2017, Courtade et al., 2017, Ram and Sason, 2016, Rioul, 2017, Courtade, 2018, Costa, 1985, Cover and Thomas, 2012, Toscani, 2012, Johnson and Yu, 2010].

Following the analogy between the Brunn-Minkoswki inequality and the entropy power inequality, the latter provides an entropic version of the isoperimetric inequality:

Theorem 4.5.3 (Entropic isoperimetry). *Given a random variable X on $\mathbb{R}^n$ with associated density function f_X, the following inequality holds:*

$$I(f_X)N(X) \geq 2\pi\,e\,n,$$ (e − Isop)

where the Fisher information I is defined in Equation (1.71). The inequality is saturated if and only if X is Gaussian.

This inequality can be used to prove the uncertainty principle [a.J. Stam, 1959]. For a discussion about the link to Brunn-Minkowski's inequality, we refer to [Costa and Cover, 1984].

The general meaning of the entropy power inequality is that entropy increases, and therefore information decreases, when adding up random signals. The addition of noise results in the convolution of the corresponding density. In particular, assuming that signal Y is a Gaussian random variable Z_t of variance t, the sum $X + Z_t$ is the Brownian motion whose density is the solution to the heat equation

$$\frac{\partial_t f_{X+Z_t}}{\partial_t} = \frac{1}{2}\Delta f_{X+Z_t}.$$

In this case, Theorem 4.5.2 implies the following *concavity of the entropy power* along the heat semigroup:

$$\frac{d^2}{dt^2}N(X + Z_t) \leq 0.$$ (4.11)

This was first proved by Costa in [Costa, 1985] by means of direct computations. Later, Dembo [Dembo et al., 1991, Dembo, 1989] simplified the proof, by an argument based on the so-called Blachman-Stam inequality [Blachman, 1965], which is at the core of his proof of Theorem 4.5.2. More recently, Villani [Villani, 2000] gave a direct proof of the same inequality.

Lemma 4.5.4 (Fisher information inequality). *Let X and Y be two independent random variables in $\mathbb{R}$, with respective densities f_X, f_Y. Then for any $\alpha, \beta > 0$,*

$$(\alpha + \beta)^2\,I(f_{X+Y}) \leq \alpha^2\,I(f_X) + \beta^2\,I(f_Y).$$ (4.12)

The advantage of the concavity of the entropy power is that it extends to any geometry where there is a notion of heat. In the next section, we see how these are related to the logarithmic Sobolev inequality for the Ornstein Uhlenbeck semigroup.

Equivalence between Gross-LSI **and** $e-\mathrm{Isop}$ In Example 3.2.4 we mentioned the first logarithmic Sobolev inequality was derived by Gross for the Gaussian measure. In this section, we show that this inequality is nothing but a reformulation of $e-\mathrm{Isop}$. The main difference between the two inequalities lies in the choice of the reference measure in the integrals involved: Shannon's differential entropy is expressed in terms of the Lebesgue measure whereas the entropy lying on the left hand side of Gross' LSI is expressed in terms of the Gaussian measure. The equivalence between the two inequalities was observed in the nineties [Carlen, 1991] (see also [Chafai, 2005] and references therein).

More precisely, the logarithm of $e-\mathrm{Isop}$ can be written as follows:

$$\int_{\mathbb{R}^n} f(x) \ln f(x) \, dx \le \frac{n}{2} \ln\left(\frac{1}{2\pi \, e \, n} \int_{\mathbb{R}^n} \frac{|\nabla f(x)|^2}{f(x)} \, dx \right).$$

Taking $f = g^2$, we see that the entropic isoperimetric inequality is equivalent to the following: for any $g : \mathbb{R}^n \to \mathbb{R}$ such that $\int g^2 = 1$ and g converges to 0 fast enough at ∞:

$$\int_{\mathbb{R}^n} g(x)^2 \ln g(x)^2 \, dx \le \frac{n}{2} \ln\left(\frac{2}{\pi n \, e} \int_{\mathbb{R}^n} |\nabla g(x)|^2 \, dx \right). \tag{4.13}$$

The isoperimetric inequality written in this form is usually referred to as the *Stam-Carlen inequality*. Taking $g = h\sqrt{f_Z}$ in (4.13), where f_Z is the density function of the standard Gaussian on $\mathbb{R}^n$, gives for any $A > 0$:

$$\int_{\mathbb{R}^n} h(x)^2 \ln\left(h(x)^2 f_Z(x) \right) f_z(x) \, dx \le \frac{n}{2} \ln\left(\frac{2}{\pi n \, e} \int_{\mathbb{R}^n} |\nabla h(x) \sqrt{f_Z}(x)|^2 \, dx \right)$$
$$= \frac{n}{2} \ln\left(A \int_{\mathbb{R}^n} |\nabla h \sqrt{f_Z}|^2 \, dx \right) + \frac{n}{2} \ln\left(\frac{2}{A \pi n \, e} \right).$$

A simple calculation involving integration by parts and the fact that $\int_{\mathbb{R}^n} h^2 f_Z(x) dx = 1$ leads to the following

$$\int_{\mathbb{R}^n} h^2 \ln h^2 \, f_Z(x) \, dx - \frac{n}{2} \ln(2\pi) + \left(\frac{An}{8} - \frac{1}{2} \right) \int_{\mathbb{R}^n} h^2 |x|^2 f_Z(x) dx$$
$$\le \frac{n}{2} A \mathcal{E}(h, h) + \frac{An^2}{2} - \frac{n}{2} - n + \frac{n}{2} \ln \frac{2}{A \pi n}.$$

Gross-LSI follows after choosing $A = \frac{4}{n}$. Moreover, the argument is reversible (see Corollary 6.9 in [Chafai, 2005]). Hence, we proved the following:

Theorem 4.5.5. *The $e-\mathrm{Isop}$ inequality is equivalent to* Gross-LSI.

More recently, links between Gross' LSI and the entropy power inequality have been further studied [Courtade, 2016].

4.6. The concentration of measure phenomenon

Concentration of measure is the phenomenon according to which almost all the points of a set are close to a subset of positive measure. More precisely let (E, d) be a metric space, and μ a probability measure on the Borel sets $\mathcal{B}(E)$. Then given a set $A \in \mathcal{B}(E)$ such that $\mu(A) \ge 1/2$, the complement

$(A^r)^c$ of its r-enlargements $A^r := \{x : d(x, A) \le r\}$ should rapidly decay with r. [Marton, 1986] showed that the concentration of measure phenomenon holds as long as the following inequality is true for any probability measure $\nu << \mu$:

$$W_2(\mu, \nu) \le C \sqrt{H_\mu(\nu)}. \tag{4.14}$$

The above is usually referred to as a transportation cost inequality, and morally quantifies the difficulty of transporting a measure ν to μ in terms of their divergence. Now, for $\nu_A(.) = \frac{\mu(\,.\,\cap A)}{\mu(A)}$:

$$r \le W_2(\nu_A, \nu_{(A^r)^c}) \le W_2(\nu_A, \mu) + W_2(\nu_{(A^r)^c}, \mu)$$

$$\le C \left(\sqrt{H_\mu(\nu_A)} + \sqrt{H_\mu(\nu_{(A^r)^c})} \right)$$

$$= C \left(\sqrt{\log \frac{1}{\mu(A)}} + \sqrt{\log \frac{1}{\mu((A^r)^c)}} \right),$$

where the first inequality can be interpreted as the fact that it takes at least a cost r to transport each particle from A to $(A^r)^c$. Inverting the above inequality, we get Gaussian concentration:

$$\mu(A^r) \ge 1 - \exp\left\{ -\left(\frac{r}{C} - \sqrt{\log \frac{1}{\mu(A)}} \right)^2 \right\}. \tag{4.15}$$

Transportation cost inequalities represent an efficient way of encoding a whole profile of Gaussian concentrations. More generally, the measure μ is said to satisfy a *transportation cost inequality* of order p and constant $c > 0$ if for any $\nu \in \mathcal{P}_p(E)$:

$$W_p(\mu, \nu) \le \sqrt{2\, c\, H_\mu(\nu)}, \tag{$\text{TC}_p(c)$}$$

where $\mathcal{P}_p(E)$ denotes the Wasserstein space of order p defined in Section 4.1.2. Since $W_p \le W_q$ for any $1 \le p \le q$, TC_q implies TC_p. Perhaps the simplest example of a transportation cost inequality is *Pinsker's inequality*:

$$\|\mu - \nu\|_{\text{TV}} \le \sqrt{\frac{H_\mu(\nu)}{2}}.$$

As discussed above, the motivation for introducing these inequalities is because they imply the concentration of measure phenomenon for μ (see Theorem 22.10 of [Villani, 2008]):

Theorem 4.6.1. *Let $(E.d)$ be a Polish space, equipped with a reference probability measure μ. Then the following properties are all equivalent:*

(i) *μ lies in $\mathcal{P}_1(E)$ and satisfies a TC_1 inequality*

(ii) *There is $\lambda > 0$ such that for any $\varphi \in C_b(E)$, and any $t \ge 0$:*

$$\int e^{t \inf_{y \in \mathcal{X}} [\varphi(y) + d(x,y)]} \mu(dx) \le e^{\frac{t^2}{2\lambda}} e^{t \int \varphi d\mu}$$

(iii) *There is a constant $C > 0$ such that for any Borel set $A \subset E$, $\mu(A) \ge \frac{1}{2}$ implies that for any $r > 0$:*

$$\mu(A^r) \ge 1 - e^{-C r^2}.$$

(iv) There is a constant $C > 0$ such that for all $f \in \mathbb{L}^1(\mu) \cap \mathrm{Lip}(E)$, $\forall \varepsilon > 0$,

$$\mathbb{P}_\mu(f - \mathbb{E}_\mu[f] \geq \varepsilon) \leq \exp\left(-C \, \frac{\varepsilon^2}{\|f\|_{\mathrm{lip}}^2}\right),$$

and the same inequality holds by replacing $\mathbb{E}_\mu[f(X)]$ by the median of f.

(v) There is a constant $C > 0$ such that for all random variable $f \in \mathbb{L}^1(\mu) \cap \mathrm{Lip}(E)$, $\forall \varepsilon > 0$, $\forall N \in \mathbb{N}$:

$$\mathbb{P}_\mu^{\otimes n}\left(\frac{1}{N}\sum_{i=1}^{N} f(X_i) - \mathbb{E}_\mu[f(X)] \geq \varepsilon\right) \leq \exp\left(-C \, \frac{N\,\varepsilon^2}{\|X\|_{\mathrm{lip}}^2}\right).$$

The implication (i)$\Rightarrow$(iii) of the above theorem is due to [Marton, 1996b] (see also [Bobkov and Goetze, 1999, Djellout et al., 2004, Bolley and Villani, 2005]). In the case of TC$_2$, there exists stronger, dimension-free Gaussian concentration, due to the tensorization property of W_2 (see [Gozlan, 2009]).

Link to other inequalities The transportation cost inequality of order 2 can be derived from CD(κ, ∞) when $\kappa > 0$. This was first realized by [Talagrand, 1996] when μ is the Gaussian measure on $\mathbb{R}^n$. The result was extended to other measures by [Blower, 2003, Cordero-Erausquin, 2002]. Other approaches can also be found in the literature [Otto and Villani, 2000a, Bobkov and Ledoux, 2000]. We also mention that stronger kinds of transportation cost inequalities (and hence concentrations) were obtained in [Lott and Villani, 2007, Gentil, 2002] by imposing a bound $N < \infty$ on the dimension. Moreover, TC$_2$ can be related to log Sobolev and Poincaré inequalities [Otto and Villani, 2000a].

Part III.

Quantum Markovian evolutions

Chapter 5.

Quantum Markov semigroups

In Chapter 0, we briefly introduced the standard mathematical model for the description of the continuous time evolution of open quantum systems. After taking a suitable limit, we showed that it can be described in terms of a one-parameter semigroup of unital, completely positive maps acting on the von Neumann algebra generated by the observables of the system.

The theory of Markov semigroups acting on abstract operator algebras was initiated by [Albeverio and Høegh-Krohn, 1977] where the bijective correspondence between symmetric Markov semigroups and Dirichlet forms has been established. Their work was followed by [Davies and Lindsay, 1992] which treated the case of semigroups acting on a finite von Neumann algebra that are symmetric with respect to its trace. The general von Neumann algebraic symmetric case was later treated by [Goldstein and Lindsay, 1995].

A large part of the thesis deals with evolutions occurring on matrix algebras. In this chapter, we however decided to introduce the objects with a fair amount of generality, namely weak*-continuous semigroups on the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on a separable Hilbert space $\mathcal{H}$, which describes dissipative evolutions occurring in quantum optical information processing devices and *continuous variables* (CV) quantum information theory. First of all, this is the amount of generality needed to introduce the quantum heat and Ornstein Uhlenbeck semigroups studied in Chapter 11. Moreover, we believe that the concepts introduced in the next chapters can be extended to this setting, after taking care of the issues arising from the manipulation of unbounded operators that constitute the daily worries of CV information theorists. Quantum optical systems are known to be prominent candidates for implementation of quantum information processing devices. Hence, this can be seen as a series of first steps towards the author's goal to extend the present results to the case of Markovian noise occurring in quantum optical systems. In particular, after an introduction to the non-commutative differential calculus in Section 5.3, we prove a non-commutative, infinite dimensional *Stroock-Varopoulos inequality*. This result is the main ingredient of the proof of the equivalence between the logarithmic Sobolev inequality and hypercontractivity of reversible quantum Markov semigroups.

Layout of the chapter: For sake of comparison with the classical setting, this chapter is organized similarly to Chapter 2: After a brief recapitulation of the theory of semigroups of operators on Banach algebras in Section 5.1. We introduce the notions of invariant states, $\mathbb{L}_p$ quantum Markov semigroups, reversibility and Dirichlet forms in Section 5.2. The notion of a non-commutative differential calculus associated to a QMS is presented in Section 5.3. In Section 5.4, this notion is then used to prove a quantum extension of the $\mathbb{L}_p$ regularity of Dirichlet forms seen in Section 2.4. We end this chapter by introducing examples that will reappear in the sequel.

5.1. Semigroups of operators on Banach algebras

We recall that a *one-parameter semigroup of operators* on a Banach space $\mathcal{X}$ is a map $\mathcal{P} : [0, \infty) \to \mathcal{B}(\mathcal{X})$ (or equivalently the family $(\mathcal{P}_t)_{t \geq 0}$ of maps on $\mathcal{X}$) such that

(i) $\mathcal{P}_0 = \mathrm{id}_{\mathcal{X}}$

(ii) *semigroup property*: for all $s, t \geq 0$,

$$\mathcal{P}_{s+t} = \mathcal{P}_s \circ \mathcal{P}_t . \tag{5.1}$$

5.1.1. Different notions of continuity:

Such a semigroup is called *uniformly continuous* if the following holds:

$$\| \mathcal{P}_t - \mathrm{id} : \mathcal{X} \to \mathcal{X} \| \underset{t \to 0}{\to} 0 .$$

The semigroup $(\mathcal{P}_t)_{t \geq 0}$ is called *strongly continuous* if for all $X \in \mathcal{X}$:

$$\| \mathcal{P}_t (X) - X \| \underset{t \to 0}{\to} 0 . \tag{5.2}$$

In the following chapters, we also consider weaker notions of continuity: the semigroup $(\mathcal{P}_t)_{t \geq 0}$ is *weak* continuous* if:

(i) For all $X \in \mathcal{X}$ and all $\eta \in \mathcal{X}_*$, the map $t \mapsto \eta (\mathcal{P}_t(X))$ is continuous on $\mathbb{R}_+$;

(ii) For all $t \geq 0$ and any $\eta \in \mathcal{X}_*$, $\eta \circ \mathcal{P}_t \in \mathcal{X}_*$.

Similarly, the semigroup $(\mathcal{P}_t)_{t \geq 0}$ is said to be *weakly continuous* if the above two conditions hold after replacement of the predual $\mathcal{X}_*$ by the dual $\mathcal{X}^*$ of $\mathcal{X}$. In fact, any weakly continuous semigroup is strongly continuous (see Corollary 3.1.8 of [Bratteli and Robinson, 1979]). Therefore, a simple ordering of the topologies implies the following inclusion of the classes introduced:

uniform continuous $\subset$ strongly continuous = weakly continuous $\subset$ weak* continuous .

5.1.2. Generators:

Let $(\mathcal{P}_t)_{t \geq 0}$ be a weak* continuous semigroup of operators on the Banach space $\mathcal{X}$. The *weak* generator* $(\mathcal{L}, \mathrm{dom}(\mathcal{L}))$ of $(\mathcal{P}_t)_{t \geq 0}$ is defined as the linear operator $\mathcal{L}$ on $\mathcal{X}$ whose domain is composed of elements $X \in \mathcal{X}$ such that there exists $Y \in \mathcal{X}$ with the property that for all $\eta \in \mathcal{X}_*$, $\eta(Y) = \lim_{t \to 0} t^{-1} \eta((\mathcal{P}_t - \mathrm{id})(X))$. In this case

$$\mathcal{L}(X) := \mathrm{weak}^* - \lim_{h \to 0} \frac{\mathcal{P}_h(X) - X}{h} .$$

In the case when $(\mathcal{P}_t)_{t \geq 0}$ is strongly continuous, its *generator* is the linear operator $(\mathcal{L}, \mathrm{dom}(\mathcal{L}))$, where $\mathrm{dom}(\mathcal{L})$ is the subspace of vectors $X \in \mathcal{X}$ for which there exists $Y \in \mathcal{X}$ such that $\lim_{h \to 0} h^{-1} \| \mathcal{P}_h(X) - X - hY \| \to 0$. In this case, we set $Y = \mathcal{L}(X)$. In particular, this generator coincides with the weak generator, defined similarly to the weak* generator by simply requiring the limit to exists for any $\eta \in \mathcal{X}^*$.

By the semigroup property (5.1), the domain of $\mathcal{L}$ is then the set of vectors $X \in \mathcal{X}$ for which the map $t \mapsto \mathcal{P}_t(X)$ is continuously differentiable of derivative denoted by $\frac{d}{dt} \mathcal{P}_t(X)$. Moreover, for any

such X,

$$\frac{d}{dt}\mathcal{P}_t(X) = \mathcal{L} \circ \mathcal{P}_t(X) = \mathcal{P}_t \circ \mathcal{L}(X).$$

A strongly continuous contraction semigroup $(\mathcal{P}_t)_{t\geq 0}$ with generator $(\mathcal{L}, \mathrm{dom}(\mathcal{L}))$ satisfies the following equation (see (10.3.7) of [Hille and Phillips, 1996]): for any $X \in \mathrm{dom}(\mathcal{L})$ and $t \geq 0$,

$$\mathcal{P}_t(X) - X = \int_0^t \mathcal{P}_s \circ \mathcal{L}(X)\,ds, \tag{5.3}$$

where the integral is defined in the Bochner sense.

The generator $\mathcal{L}$ is bounded, which equivalently means that $\mathrm{dom}(\mathcal{L}) = \mathcal{X}$, if and only if $(\mathcal{P}_t)_{t\geq 0}$ is uniformly continuous. A linear operator on $\mathcal{X}$ is the infinitesimal generator of at most one strongly continuous semigroup. Conversely, the Hille-Yosida-Phillips theorem provides necessary and sufficient conditions under which a linear operator $\mathcal{L}$ is the generator of a semigroup (see [Hille and Phillips, 1996]).

A strongly continuous semigroup $(\mathcal{P}_t)_{t\geq 0}$ is called a *contraction semigroup* if it satisfies $\|\mathcal{P}_t : \mathcal{X} \to \mathcal{X}\| \leq 1$ for any $t \geq 0$. One can show [Bratteli and Robinson, 1979] that in this case, $\mathcal{L}$ is weak* densely defined. The generator associated to a strongly continuous contraction semigroup is *dissipative*: for all $X \in \mathcal{X}$ there exists $X^* \in \mathcal{X}^*$, the dual of $\mathcal{X}$, such that $\|X^*\|^2 = \|X\|^2 = X^*(X)$, and $\mathrm{Re}\left(X^*(\mathcal{L}(X))\right) \leq 0$. More precisely:

Theorem 5.1.1 (Lumer-Phillips theorem for contraction semigroups). *Let $\mathcal{X}$ be a complex Banach space, $\mathcal{L} : \mathrm{dom}(\mathcal{L}) \to \mathcal{X}$ a densely defined complex linear operator. The following are equivalent:*

(i) $\mathcal{L}$ is the infinitesimal generator of a strongly continuous contraction semigroup.

(ii) For all $\lambda > 0$, $\lambda\,\mathrm{id} - \mathcal{L} : \mathrm{dom}(\mathcal{L}) \to \mathcal{X}$ is bijective and

$$\|(\lambda\,\mathrm{id} - \mathcal{L})^{-1}\| \leq \lambda^{-1}.$$

(iii) For all $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) > 0$, $\lambda\,\mathrm{id} - \mathcal{L} : \mathrm{dom}(\mathcal{L}) \to X$ is bijective and

$$\|(\lambda\,\mathrm{id} - \mathcal{L})^{-1}\| \leq \mathrm{Re}(\lambda)^{-1}.$$

(iv) $\mathcal{L}$ is dissipative and there exists $\lambda > 0$ such that $\lambda\,\mathrm{id} - \mathcal{L} : \mathrm{dom}(\mathcal{L}) \to \mathcal{X}$ has a dense image.

5.1.3. Quantum Markov semigroups:

As mentioned in Section 0.1.5, we are exclusively interested in the study of weak* continuous, completely positive, unital contraction semigroups $(\mathcal{P}_t)_{t\geq 0}$ acting on a von Neumann algebra of operators $\mathcal{X} \coloneqq \mathcal{M}$, and refer to them as *quantum Markov semigroup* (QMS). In fact, any completely positive, unital Markov semigroup is necessarily contractive: this is due to the Russo-Dye Theorem (Corollary 2.9 of [Paulsen, 2002]) according to which, for all $t \geq 0$:

$$\|\mathcal{P}_t : \mathcal{M} \to \mathcal{M}\| = \|\mathcal{P}_t(\mathbb{1}_{\mathcal{M}})\|_\infty = 1.$$

Quantum Markov semigroups can be viewed as dual semigroups of strongly continuous semigroups on the predual Banach space $\mathcal{M}_*$, defined by:

$$\mathcal{P}_{*t}(\omega)(X) = \omega\left(\mathcal{P}_t(X)\right),$$

for every $X \in \mathcal{M}$ and $\omega \in \mathcal{M}_*$. Indeed, since $t \mapsto \mathcal{P}_t$ is weak* continuous on $\mathcal{M}$, $t \mapsto \mathcal{P}_{*t}$ is weakly continuous on $\mathcal{M}_*$. From our previous discussion, this actually implies that $t \mapsto \mathcal{P}_{*t}$ is strongly continuous on $\mathcal{M}_*$.

Uniformly continuous QMS: In the case of a uniformly continuous completely positive semigroup the following important result was first proved by [Lindblad, 1976, Gorini et al., 1976] for matrix algebras, and extended by [Christensen and Evans, 1979] to the general C*-algebraic setting:

Theorem 5.1.2 (Standard representation of a norm continuous QMS). *Let $(\mathcal{P}_t)_{t \geq 0}$ be a uniformly continuous semigroup of completely positive normal unital maps on a von Neumann algebra $\mathcal{M}$ of operators acting on a Hilbert space $\mathcal{H}$. Then its generator $\mathcal{L}$ may be represented as*

$$\forall X \in \mathcal{M}, \quad \mathcal{L}(X) = i\,[H, X] - \{\Psi(\mathbb{1}_{\mathcal{A}}), X\} + \Psi(X) = \Psi(X) + K^* X + XK\,, \tag{5.4}$$

where $K, H = H^* \in \mathcal{M}$ and $\Psi : \mathcal{M} \to \mathcal{M}$ is a completely positive map.

In this case we obtain the following *Lindblad form* of $\mathcal{L}$ from the Kraus decomposition (0.23) of Ψ,

$$\mathcal{L}(X) = i\,[H, X] + \frac{1}{2} \sum_{j \in \mathcal{J}} 2\, L_j^* X L_j - X L_j^* L_j - L_j^* L_j X\,, \tag{5.5}$$

where $\{L_j\}_{j \in \mathcal{J}}$ is a family of bounded operators such that $\sum_{j \in \mathcal{J}} L_j^* L_j$ is bounded and $\sum_{j \in \mathcal{J}} L_j^* X L_j \in \mathcal{M}$ for any $X \in \mathcal{M}$. In this case, $\mathcal{L}$ is usually called a *Lindblad generator*.

A physical interpretation for the Lindblad form (5.5) can be given from the Dyson expansion of the solution of the forward Markov master equation

$$\mathcal{P}_t = \hat{\mathcal{P}}_t + \sum_{n=1}^{\infty} \int \cdots \int_{0 \leq t_1 \leq \ldots \leq t_n \leq t} \hat{\mathcal{P}}_{t_1} \circ \Psi \circ \hat{\mathcal{P}}_{t_2 - t_1} \circ \cdots \circ \Psi \circ \hat{\mathcal{P}}_{t - t_n}\, dt_1 \ldots dt_n\,, \tag{5.6}$$

where $\hat{\mathcal{P}}_t(X) := \mathrm{e}^{-K^* t}\, X\, \mathrm{e}^{-Kt}$. In words, the above equation says that $(\mathcal{P}_t)_{t \geq 0}$ can be described by a sequence of "spontaneous jumps" modeled by the map Ψ occurring at times $t_1 \leq \ldots \leq t_n$ on the background of a nonunitary reversible evolution given by the semigroup $(\hat{\mathcal{P}}_t)_{t \geq 0}$.

Weak* continuous QMS on $\mathcal{B}(\mathcal{H})$: The non-norm continuous case, while being interesting from both physical and mathematical reasons, is more difficult to handle. The generator of such a QMS may be unbounded, and its domain is not necessarily a *-algebra. Moreover the correspondence between the master equation and the semigroup is not one-to-one, since the solution for the former need not be unique. In a series of articles, Holevo partially introduced the so-called *form generator* associated to a QMS. In words, it corresponds to a differential equation without boundary terms, which makes issues arising from the unboundedness of the weak* generator easier. On the other hand, this formulation introduces other problems such as the one of existence, uniqueness and unitality of the solution.

More precisely, let $(\mathcal{P}_t)_{t \geq 0}$ be a QMS on a von Neumann algebra $\mathcal{M}$ acting on a separable Hilbert space $\mathcal{H}$, and assume there exists a dense domain $\mathcal{V} \subseteq \mathcal{H}$ such that the derivative

$$\frac{d}{dt} \langle \varphi, \mathcal{P}_t(X)\psi \rangle \Big|_{t=0} =: \mathcal{L}(X)[\varphi, \psi] \tag{5.7}$$

exists for all $X \in \mathcal{M}$, and any $\varphi, \psi \in \mathcal{V}$. By the semigroup property of $(\mathcal{P}_t)_{t \geq 0}$, this implies that the

following so-called *master equation* holds for any $t \geq 0$, any $X \in \mathcal{M}$ and any $\varphi, \psi \in \mathcal{V}$:

$$\frac{d}{dt} \langle \varphi, \mathcal{P}_t(X)\psi \rangle = \mathcal{L}(\mathcal{P}_t(X))[\varphi, \psi].$$

Moreover, the form $(\mathcal{M}, \mathcal{V}, \mathcal{V}) \ni (X, \varphi, \psi) \mapsto \mathcal{L}(X)[\varphi, \psi]$ satisfies the following properties, due to the definition of the QMS $(\mathcal{P}_t)_{t \geq 0}$:

- *Adjoint preservation property*: $(X, \varphi, \psi) \mapsto \mathcal{L}(X)[\varphi, \psi]$ is linear in X and ψ, antilinear in φ and

$$\mathcal{L}(X^*)[\varphi, \psi] = \overline{\mathcal{L}(X)[\psi, \varphi]};$$

- *Complete positivity*: Given any finite subsets $\{\varphi_i\} \subset \mathcal{V}$ and $\{X_i\} \subset \mathcal{M}$ such that $\sum_i X_i \varphi_i = 0$,

$$\sum_{i,j} \mathcal{L}(X_i^* X_j)[\varphi_i, \varphi_j] \geq 0;$$

- *Unitality*: for any $\varphi, \psi \in \mathcal{V}$,

$$\mathcal{L}(\mathbb{1})[\varphi, \psi] = 0; \tag{5.8}$$

- *Weak* continuity*: for any $\psi, \varphi \in \mathcal{V}$, the functional $X \mapsto \mathcal{L}(X)[\varphi, \psi]$ is weak* continuous on $\mathcal{M}$.

Any map $(X, \varphi, \psi) \mapsto \mathcal{L}(X)[\varphi, \psi]$ satisfying the above properties is called a *form generator*. The following theorem, proved in [Holevo, 1995a] provides a characterization of form generators in the case when $\mathcal{M} = \mathcal{B}(\mathcal{H})$:

Theorem 5.1.3. *Let $(X, \varphi, \psi) \mapsto \mathcal{L}(X)[\varphi, \psi]$ be a form generator on $\mathcal{B}(\mathcal{H})$. Then there exists an operator $K : \mathcal{V} \to \mathcal{H}$ and a countable family $\{L_i : \mathcal{V} \to \mathcal{H}\}_{i \in \mathcal{J}}$ of operators satisfying $\sum_i \|L_i \varphi\|^2 \leq -2 \operatorname{Re}\langle \varphi, K\varphi \rangle$ for any $\varphi \in \mathcal{V}$, and such that*

$$\mathcal{L}(X)[\varphi, \psi] = \sum_{i \in \mathcal{J}} \langle L_i \varphi, X L_i \psi \rangle + \langle K\varphi, X\psi \rangle + \langle \varphi, X K\psi \rangle. \tag{5.9}$$

The operators L_i and K can always be chosen so that the following normalizations hold: for any state $\rho := \sum_j \lambda_j |\psi_j\rangle\langle\psi_j| \in \mathcal{D}(\mathcal{H})$ whose eigenvectors are supposed to be in $\mathcal{V}$, one can without loss of generality assume that for any $i \in \mathcal{J}$:

$$\operatorname{Tr}(\rho L_i) := \sum_j \lambda_j \langle \psi_j, L_i \psi_j \rangle = 0, \quad \operatorname{Im} \operatorname{Tr}(\rho K) := \operatorname{Im} \sum_j \lambda_j \langle \psi_j, K\psi_j \rangle = 0,$$

as long as the associated series converge absolutely: $\sum_j \lambda_j |\langle \psi_j, L_i \psi_j \rangle| < \infty$, $\sum_j \lambda_j |\langle \psi_j, K\psi_j \rangle| < \infty$. Moreover, for any other operators $\tilde{L}_i$ and $\tilde{K}$ satisfying Equation (5.9) with corresponding state $\rho' = V\rho V^*$, $V \in \mathcal{U}(\mathcal{H})$, there exists a unitary operator $U = \{U_{ij}\}_{(i,j) \in \mathcal{J}^2}$, a vector $\{\alpha_i\}_{i \in \mathcal{J}}$ of complex numbers and a real number β such that

$$\tilde{L}_i = \sum_j L_j U_{ij} + \alpha_i \mathbb{1} \tag{5.10}$$

$$\tilde{K} = K - \sum_{i,j \in \mathcal{J}} \overline{U}_{ji} \alpha_j L_i - \frac{1}{2}\left(\sum_{i \in \mathcal{J}} |\alpha_i|^2 - i\beta\right)\mathbb{1}, \tag{5.11}$$

where $\alpha_i \equiv \operatorname{Tr}(\rho \tilde{L}_i)$ and $\beta \equiv \operatorname{Im} \operatorname{Tr}(\rho \tilde{K})$.

The above theorem is important for two main reasons: first, it provides a justification for the search of candidate generators of quantum Markov semigroups in a *generalized* GKLS *form*, which was already assumed in previous work [Davies, 1977, Chebotarev and Fagnola, 1998, Fagnola, 1993]. Moreover, the equivalence class of representations $(K, \{L_i\}_{i \in \mathcal{J}})$ provided in Equations (5.10) and (5.11) is at the heart of Holevo's characterization of covariant quantum Markov semigroups which lead to a noncommutative extension of the Lévy-Khinchin formula for generators of classical shift-covariant Markov semigroups [Holevo, 1993a, Holevo, 1995a, Holevo, 1993b, Holevo, 1995b, Holevo, 1996, Holevo, 1998]. In Section 5.2.2, we will provide a useful expression for the generator of symmetric quantum Markov semigroups as yet another straightforward consequence of Theorem 5.1.3.

On the other hand, given a form-generator $\mathcal{L}$ in the generalized GKLS form of Equation (5.9), the question of the existence, uniqueness and unitality of a QMS satisfying the Master equation (5.7) has been partially solved under some further conditions on the operators K and L_i, $i \in \mathcal{J}$. In what follows, we provide a brief summary of the main results in that direction. Given a countable set $\mathcal{J}$, let K and L_k, $k \in \mathcal{J}$ be (possibly unbounded) operators on $\mathcal{H}$, and assume the following holds:

Condition 5.1.4 (H-min)**.** The operator K is the generator of a strongly continuous semigroup $(P_t)_{t \geq 0}$ on $\mathcal{H}$, the domain $\mathrm{dom}(L_k)$ of each operator L_k is contained in $\mathrm{dom}(K)$, and for all $\psi, \varphi \in \mathrm{dom}(K)$:

$$\langle K\psi, \varphi \rangle + \sum_{k \in \mathcal{J}} \langle L_k \psi, L_k \varphi \rangle + \langle \psi, K\varphi \rangle = 0\,.$$

For any $X \in \mathcal{B}(\mathcal{H})$, define the quadratic form $(\psi, \varphi) \mapsto \mathcal{L}(X)[\psi, \varphi]$ as in Equation (5.9) with domain $\mathrm{dom}(K) \times \mathrm{dom}(K)$. Under Condition 5.1.4, one can construct the so-called *minimal semigroup* $(\mathcal{P}_t^{\min})_{t \geq 0}$ on $\mathcal{B}(\mathcal{H})$ associated with the operators K and L_k, $k \in \mathcal{J}$ [Davies, 1977, Chebotarev and Fagnola, 1998, Fagnola, 1999]. This is a weak* continuous semigroup $(\mathcal{P}_t^{\min})_{t \geq 0}$ of completely positive maps satisfying

$$\forall \psi, \varphi \in \mathrm{dom}(K), \quad \langle \psi, \mathcal{P}_t^{\min}(X)\varphi \rangle = \langle \psi, X\varphi \rangle + \int_0^t \mathcal{L}\left(\mathcal{P}_s^{\min}(X)\right)[\psi, \varphi]\, ds\,. \tag{5.12}$$

Moreover, $(\mathcal{P}_t^{\min})_{t \geq 0}$ satisfies the following two properties (Theorem 3.22 of [Fagnola, 1999]):

- $\mathcal{P}_t^{\min}(\mathbb{1}) \leq \mathbb{1}$ for every $t \geq 0$;

- for every other weak* continuous semigroup of positive maps $(\mathcal{P}_t)_{t \geq 0}$ satisfying Equation (5.12), every positive operator $X \in \mathcal{B}(\mathcal{H})$ and all $t \geq 0$:

$$\mathcal{P}_t^{\min}(X) \leq \mathcal{P}_t(X)\,.$$

However, Condition 5.1.4, reminiscent from Equation (5.8), does not insure that the QMS is unital (see e.g. example 3.3 of [Davies, 1977] or example 3.4 of [Fagnola, 1999]). If we further impose the minimal semigroup to be unital, then it becomes the unique weak* continuous semigroup of positive maps on $\mathcal{B}(\mathcal{H})$ that satisfies Equation (5.12). A sufficient condition for unitality was provided in [Chebotarev and Fagnola, 1998] (see also [Holevo, 1995a]). In the statement of the theorems, we will however assume the following weaker condition (see e.g. Proposition 2.3 of [Carbone and Fagnola, 2003]):

Condition 5.1.5 (H-Markov)**.** One of these equivalent conditions holds:

- The minimal semigroup $(\mathcal{P}_t^{\min})_{t \geq 0}$ is unital.

- $(\mathcal{P}_t^{\min})_{t \geq 0}$ is the unique weak* continuous family of positive contractive maps satisfying Equation (5.12) for all $X \in \mathcal{B}(\mathcal{H})$.

- The linear space $\mathcal{D} := \mathrm{span}\{|\psi\rangle\langle\varphi|, \psi, \varphi \in \mathrm{dom}(K)\}$ is a core for $\mathcal{L}_*^{\min}$. In that case, for any $\psi, \varphi \in D$:

$$\mathcal{L}_*^{\min}(|\psi\rangle\langle\varphi|) = |K\psi\rangle\langle\varphi| + |\psi\rangle\langle K\varphi| + \sum_{r\in\mathcal{J}} |L_r\psi\rangle\langle L_r\varphi| \tag{5.13}$$

The next result provides a characterization of the domain of $\mathcal{L}^{\min}$. It is also due to [Fagnola and Rebolledo, 1998]:

Proposition 5.1.6. *Suppose that Condition 5.1.4 and Condition 5.1.5 hold. Then the domain of the infinitesimal generator $\mathcal{L}^{\min}$ of $(\mathcal{P}_t^{\min})_{t\geq 0}$ is given by all elements $X \in \mathcal{B}(\mathcal{H})$ such that the sesquilinear form defined in Equation* (5.9) *on* $\mathrm{dom}(K) \times \mathrm{dom}(K)$:

$$(\psi, \varphi) \mapsto \mathcal{L}(X)[\psi, \varphi]$$

is norm continuous.

5.2. $\mathbb{L}_p$ quantum Markov semigroups, reversibility, Dirichlet forms

5.2.1. $\mathbb{L}_p$ quantum Markov semigroups and invariant states

Just as in Section 2.2, we now extend the domain of definition of a QMS $(\mathcal{P}_t)_{t\geq 0}$ defined on $\mathcal{B}(\mathcal{H})$, $\mathcal{H}$ separable, to the larger non-commutative $\mathbb{L}_p$ spaces. Such semigroups were studied by [Albeverio and Høegh-Krohn, 1977] (in the C*-algebraic setting) and [Davies and Lindsay, 1992] in the case of $\mathbb{L}_p$ spaces defined with respect to a tracial state. The theory was later generalized to semigroups that are symmetric with respect to any faithful normal state by [Goldstein and Lindsay, 1995]. Here, we restrict ourselves to semigroups defined on the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on a separable Hilbert space $\mathcal{H}$, unless otherwise specified. We refer to the above articles for more details. Let σ be a faithful invariant state on $\mathcal{H}$ and $1 \leq p < \infty$. A strongly continuous, contraction semigroup $(\mathcal{P}_t)_{t\geq 0}$ of positive unital operators on the Banach space $\mathbb{L}_p(\sigma)$ is called an $\mathbb{L}_p(\sigma)$-*quantum Markov semigroup*. Let us first consider the case $p = 2$. By the Hille-Yosida theorem, an $\mathbb{L}_2(\sigma)$-quantum Markov semigroup is fully determined by its generator

$$\mathcal{L}_2 : \begin{cases} \mathrm{dom}(\mathcal{L}_2) \to \mathbb{L}_2(\sigma) \\ X \mapsto \lim_{\epsilon\to 0} \dfrac{\mathcal{P}_\epsilon(X) - X}{\epsilon}, \end{cases} \tag{5.14}$$

the limit being taken in the $\|.\|_{\mathbb{L}_2(\sigma)}$ topology. The state σ is said to be an *invariant state* with respect to $(\mathcal{P}_t)_{t\geq 0}$ if for all $t \geq 0$ and all $X \in \mathbb{L}_\infty(\sigma)$,

$$\mathrm{Tr}(\sigma \mathcal{P}_t(X)) = \mathrm{Tr}(\sigma X). \tag{5.15}$$

Moreover, $(\mathcal{P}_t)_{t\geq 0}$ is called *primitive* if it possesses a unique faithful invariant state. From now one, we assume that the state σ with respect to which the $\mathbb{L}_p(\sigma)$ spaces are being defined is invariant with respect to $(\mathcal{P}_t)_{t\geq 0}$. As in the classical case, the following holds:

Proposition 5.2.1. *Let $(\mathcal{P}_t)_{t\geq 0}$ be a QMS and σ a faithful invariant state. Then $(\mathcal{P}_t)_{t\geq 0}$ can be extended to an $\mathbb{L}_p(\sigma)$-QMS for any $1 \leq p < \infty$.*

Proof. Since $(\mathcal{P}_t)_{t \geq 1}$ is unital, it satisfies $\mathcal{P}_t(\mathbb{1}) = \mathbb{1}$ for all $t \geq 0$. By positivity of the maps $\mathcal{P}_t$, Russo-Dye theorem (see Corollary 2.9 of [Paulsen, 2002]) implies that $\|\mathcal{P}_t : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})\| = \|\mathcal{P}_t(\mathbb{1})\|_\infty = \|\mathbb{1}\|_\infty = 1$. Moreover, by duality of the $\mathbb{L}_p(\sigma)$ spaces, for any $X \in \mathcal{B}(\mathcal{H})$,

$$
\begin{aligned}
\|\mathcal{P}_t(X)\|_{\mathbb{L}_1(\sigma)} &= \sup_{\|Y\|_\infty \leq 1} \langle Y, \mathcal{P}_t(X)\rangle_\sigma \\
&= \sup_{\|Y\|_\infty \leq 1} \langle \hat{\mathcal{P}}_t(Y), X\rangle_\sigma \\
&\leq \sup_{\|Y\|_\infty \leq 1} \|\hat{\mathcal{P}}_t(Y)\|_\infty \|X\|_{\mathbb{L}_1(\sigma)} \\
&\leq \|X\|_{\mathbb{L}_1(\sigma)},
\end{aligned}
$$

where $\hat{\mathcal{P}}_t$ is the dual of $\mathcal{P}_t$ with respect $\langle .,.\rangle_\sigma$. By density, it then follows that $\|\mathcal{P}_t : \mathbb{L}_1(\sigma) \to \mathbb{L}_1(\sigma)\| \leq 1$. The result for any $p \geq 1$ follows by interpolation. The strong continuity of the semigroup follows by Proposition 5.8 of [Goldstein and Lindsay, 1995] which we here adapt to our setting: let $\hat{p}$ be the Hölder conjugate of p. Then, for any $X \in \mathcal{B}(\mathcal{H})$ and $Y \in \mathbb{L}_{\hat{p}}(\sigma)$,

$$
\langle Y, (\mathcal{P}_t - \mathrm{id})(X)\rangle_\sigma = \mathrm{Tr}\left(\Gamma_\sigma(Y)^* (\mathcal{P}_t - \mathrm{id})(X)\right) \underset{t \to 0}{\to} 0.
$$

The limit above holds by weak* continuity of $(\mathcal{P}_t)_{t \geq 0}$ and since

$$
\|\Gamma_\sigma(Y)\|_1 = \|Y\|_{\mathbb{L}_1(\sigma)} \leq \|Y\|_{\mathbb{L}_{\hat{p}}(\sigma)} < \infty,
$$

so that $\Gamma_\sigma(Y) \in \mathcal{T}_1(\mathcal{H}) \simeq \mathcal{B}(\mathcal{H})_*$. Since the limit holds for any $Y \in \mathbb{L}_{\hat{p}}(\sigma) \simeq \mathbb{L}_p(\sigma)^*$, $t \mapsto \mathcal{P}_t - \mathrm{id}$ converges weakly to 0 as $t \to 0$. This is equivalent to strong convergence by Proposition 1.23 of [Davies, 1980]. $\square$

For sake of simplicity, we denote the $\mathbb{L}_p(\sigma)$ QMS as $(\mathcal{P}_t)_{t \geq 0}$, and its generator as $(\mathcal{L}_p, \mathrm{dom}(\mathcal{L}_p))$ of dense domain $\mathrm{dom}(\mathcal{L}_p)$ in $\mathbb{L}_p(\sigma)$. We also assume the following condition, similar to Condition 2.2.1, holds true:

Condition 5.2.2. There exists a $*$-subalgebra $\mathcal{A}_0$ of $\mathcal{B}(\mathcal{H})$ that is $\mathbb{L}_2(\sigma)$-dense in $\mathrm{dom}(\mathcal{L}_2)$, invariant under $\mathcal{L}_2$ and dense in all the $\mathbb{L}_p(\sigma)$ spaces for $p \in [1, \infty)$. Then, we define $\mathcal{A}_0^+ := \{X + c\,\mathbb{1},\ X \in \mathcal{A}_0,\ X \geq 0,\ c > 0\}$.

In practice, infinitesimal functional inequalities such as the logarithmic Sobolev inequality will be defined on the space $\mathcal{A}_0^+$. The density of $\mathcal{A}_0$ will allow us to derive contractivity properties of the semigroup for any initial operator (see Chapter 7). In the case when the $\mathbb{L}_2(\sigma)$-QMS $(\mathcal{P}_t)_{t \geq 0}$ is the extension of a QMS on $\mathcal{B}(\mathcal{H})$ with associated Lindblad form given in Equation (5.9), the algebra $\mathcal{A}_0$ can be chosen as follows:

Proposition 5.2.3. *Assume that $(\mathcal{P}_t)_{t \geq 0}$ is a weak* continuous QMS such that Condition 5.1.4 and Condition 5.1.5 hold, and let $\sigma = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ be a normal faithful invariant state. Assume moreover that, for all $i \in \mathbb{N}$, ψ_i is contained in $\mathrm{dom}(K^*)$, $\mathrm{dom}(L_l^*)$, $l \geq 1$, such that $\sum_l \|L_l^*\psi_i\|^2 < \infty$, and that $K\mathcal{V}, K^*\mathcal{V}, L_j\mathcal{V}, L_j^*\mathcal{V} \subset \mathcal{V}$, where $\mathcal{V} := \mathrm{span}\{\psi_i,\ i \in \mathbb{N}\}$. Then $\mathcal{A}_0 = \mathrm{span}\{|\psi_i\rangle\langle\psi_j|,\ i, j \in \mathbb{N}\}$ satisfies Condition 5.2.2.*

Proof. We first show that $\mathcal{A}_0$ belongs to $\mathrm{dom}(\mathcal{L}_2)$. By Proposition 5.1.6, we simply need to prove

that for any $i, j \in \mathbb{N}$, $(v, u) \mapsto \mathcal{L}(|\psi_i\rangle\langle\psi_j|)[v, u]$ is norm continuous on $\mathrm{dom}(K) \times \mathrm{dom}(K)$:

$$\mathcal{L}(|\psi_i\rangle\langle\psi_j|)[v, u] = \langle Kv, \psi_i\rangle\langle\psi_j, u\rangle + \sum_l \langle L_l v, \psi_i\rangle\langle\psi_j, L_l u\rangle + \langle v, \psi_i\rangle\langle\psi_j, Ku\rangle$$

$$= \langle v, K^*\psi_i\rangle\langle\psi_j, u\rangle + \sum_l \langle v, L_l^*\psi_i\rangle\langle L_l^*\psi_j, u\rangle + \langle v, \psi_i\rangle\langle K^*\psi_j, u\rangle$$

$$\leq \|u\| \|v\| \left(\|K^*\psi_i\| \|\psi_j\| + \sum_l \|L_l^*\psi_i\| \|L_l^*\psi_j\| + \|\psi_i\| \|K^*\psi_j\| \right),$$

and the claim follows by a use of Cauchy Schwartz inequality. This directly implies that $\mathcal{A}_0 \subset \mathrm{dom}(\mathcal{L}_2)$. By definition, $\mathcal{A}_0$ is dense in $\mathcal{T}_p(\mathcal{H})$ for any $p \geq 1$. This implies that for any $p \geq 1$, and any $X \in \mathbb{L}_p(\sigma)$, there exists a sequence $\{X_n\}_{n \in \mathbb{N}}$ of elements in $\mathcal{A}_0$ such that $\|\Gamma_\sigma^{-\frac{1}{p}}(X_n) - X\|_{\mathbb{L}_p(\sigma)} = \|X_n - \Gamma_\sigma^{\frac{1}{p}}(X)\|_p \to 0$. This also proves the density of $\mathcal{A}_0$ in $\mathcal{D}(\mathcal{L}_2) \subset \mathbb{L}_2(\sigma)$. The invariance of $\mathcal{A}_0$ under the action of $\mathcal{L}_2$ arises from the condition of invariance of $\mathcal{V}$ under the action of the operators $K, K^*, L_j, L_j^*, j \in \mathcal{J}$. $\quad\square$

5.2.2. KMS symmetry and detailed balance condition

There are various ways of extending the notion of reversibility seen in Section 2.2 to the quantum setting. One possible way is by use of the modular automorphism group associated to the corresponding invariant state (cf. Section 0.1.5).

Definition 5.2.4 (KMS-symmetry)**.** Let $\beta \in \mathbb{R}$, $(\alpha_t)_{t \in \mathbb{R}}$ be a weak* continuous group of *-automorphisms of a von Neumann algebra $\mathcal{M}$ and ω a fixed (α, β)-KMS state. A bounded map $\Phi: \mathcal{M} \to \mathcal{M}$ is said to be (α, β)-KMS *symmetric* with respect to ω if

$$\omega\left(B\Phi(A)\right) = \omega\left(\alpha_{-i\beta/2}(A)\,\Phi(\alpha_{i\beta/2}(B))\right), \tag{KMS-symmetry}$$

holds for any A, B in a weak* dense $(\alpha_t)_{t \geq 0}$-invariant $*$-subalgebra of the set $\mathcal{M}_\alpha$ of analytic elements of $(\alpha_t)_{t \in \mathbb{R}}$. A weak* continuous semigroup $(\mathcal{P}_t)_{t \geq 0}$ on $\mathcal{M}$ is said to be (α, β)-KMS *symmetric* with respect to ω if $\mathcal{P}_t$ is (α, β)-KMS symmetric with respect to ω for all $t \geq 0$.

In the case when $\beta = 0$ (ω tracial state), the condition of KMS-symmetry simplifies to

$$\langle\Phi(A), B\rangle_{1,\omega} = \langle A, \Phi(B)\rangle_{1,\omega}, \qquad A, B \in \mathcal{M}, \tag{ω-DBC}$$

where $\langle X, Y\rangle_{1,\omega} := \omega(X^*Y)$ is usually referred to as the GNS-*inner product*. This so-called *quantum detailed balance condition* (also known as GNS-*symmetry*) was introduced by [Kossakowski et al., 1977] and will play an important role in Chapter 12. For the moment, we make the following well-known observation, a proof of which can be found in [Cipriani, 1997]: a bounded linear map $\Phi: \mathcal{M} \to \mathcal{M}$ is said to commute with a weak* continuous group of automorphisms $(\alpha_s)_{s \in \mathbb{R}}$ if, for any element $X \in \mathcal{M}_\alpha$ and any $s \in \mathbb{R}$:

$$\Phi \circ \alpha_s(X) = \alpha_s \circ \Phi(X).$$

A QMS $(\mathcal{P}_t)_{t \geq 0}$ is said to commute with $(\alpha_s)_{s \in \mathbb{R}}$ if $\mathcal{P}_t$ commutes with $(\alpha_s)_{s \in \mathbb{R}}$ for any $t \geq 0$.

Lemma 5.2.5. *Let $\mathcal{M}$ be a von Neumann algebra, $\beta \in \mathbb{R}$, $(\alpha_t)_{t \in \mathbb{R}}$ a weak* continuous group of automorphisms, ω an (α, β)-KMS state on $\mathcal{M}$ and $\Phi: \mathcal{M} \to \mathcal{M}$ a bounded map. Then the following holds:*

(i) If Φ commutes with $(\alpha_t)_{t \in \mathbb{R}}$, then KMS-symmetry is equivalent to ω-DBC.

(ii) If $(\alpha_t)_{t\in\mathbb{R}} = (\alpha_t^\omega)_{t\in\mathbb{R}}$ is the modular automorphism group (so that $\beta = -1$), and if Φ satisfies ω-DBC, then Φ commutes with $(\alpha_t^\omega)_{t\in\mathbb{R}}$ and satisfies KMS-symmetry.

Proof. (i) Assume that Φ commutes with $(\alpha_t)_{t\geq 0}$ and that KMS-symmetry holds. Then for all A, B in a norm dense, α-invariant $*$-subalgebra of $\mathcal{M}_\alpha$,

$$\omega(B\Phi(A)) = \omega(\alpha_{-i\beta/2}(A)\Phi(\alpha_{i\beta/2}(B)))$$
$$= \omega(\alpha_{-i\beta/2}(A)\alpha_{i\beta/2}(\Phi(B)))$$
$$= \omega(\Phi(B)A),$$

where the first line follows from KMS-symmetry, the second one by commutation of Φ and $(\alpha_t)_{t\in\mathbb{R}}$ and the fact that $B \in \mathcal{M}_\alpha$, and the last line follows from Equation (0.13). Then ω-DBC follows by weak* density. Since all the steps above are reversible, the equivalence is proved.

(ii) This is done in Proposition 2.1 of [Kossakowski et al., 1977] (see also Lemma 2 of [Bratteli and Robinson, 1976]). $\qquad\square$

In the case when $\mathcal{M} = \mathcal{B}(\mathcal{H})$, with $\mathcal{H}$ separable, and ω is normal faithful with associated density operator σ and cyclic vector $\Omega_\omega := \sigma^{\frac{1}{2}}$, a semigroup $(\mathcal{P}_t)_{t\geq 0}$ is $((\alpha_s^\sigma)_{s\in\mathbb{R}}, -1)$-KMS symmetric with respect to σ, where $(\alpha_s^\sigma)_{s\in\mathbb{R}}$ is the modular automorphism group defined in Section 0.1.5, if and only if for any $t \geq 0$ and all A, B in a weak*-dense $(\alpha_s^\sigma)_{s\in\mathbb{R}}$-invariant subspace of the algebra $\mathcal{B}(\mathcal{H})_\sigma$ of entire analytic elements of $\mathcal{B}(\mathcal{H})$ for the modular group $(\alpha_s^\sigma)_{s\in\mathbb{R}}$:

$$\mathrm{Tr}\left(\sigma B \mathcal{P}_t(A)\right) = \mathrm{Tr}\left(\sigma^{\frac{1}{2}} A \sigma^{\frac{1}{2}} \mathcal{P}_t(\sigma^{\frac{1}{2}} B \sigma^{-\frac{1}{2}})\right). \tag{5.16}$$

However, given an eigenvector decomposition of σ as $\sigma = \sum_{i=1}^\infty \lambda_i |e_i\rangle\langle e_i|$, $\mathcal{A}_0 := \mathrm{span}\left\{|e_i\rangle\langle e_j|, i, j \in \mathbb{N}\right\}$ is a weak* dense subalgebra of $\mathcal{B}(\mathcal{H})$ whose norm closure is the algebra $\mathcal{K}$ of compact operators. Moreover, it is easy to see that it is included in the algebra $\mathcal{B}(\mathcal{H})_\sigma$. Therefore, for $A, B \in \mathcal{A}_0$, upon replacement of $\sigma^{\frac{1}{2}} B \sigma^{-\frac{1}{2}}$ by B and A by A^* in Equation (5.16), we get $\langle B, \mathcal{P}_t(A)\rangle_\sigma = \langle A, \mathcal{P}_t(B)\rangle_\sigma$. Since $\mathcal{P}_t$ is continuous and $\mathcal{A}_0$ is dense in $\mathbb{L}_2(\sigma)$, this also holds for any $A, B \in \mathbb{L}_2(\sigma)$. We then proved that in the case of a QMS $(\mathcal{P}_t)_{t\geq 0}$ on $\mathcal{B}(\mathcal{H})$, $((\alpha_s^\sigma)_{t\geq 0}, -1)$-KMS-symmetry with respect to the faithful normal state σ is equivalent to the following: for any $A, B \in \mathbb{L}_2(\sigma)$,

$$\langle B, \mathcal{P}_t(A)\rangle_\sigma = \langle A, \mathcal{P}_t(B)\rangle_\sigma. \tag{KMS-symmetry}$$

That is, the QMS is symmetric with respect to the $\mathbb{L}_2(\sigma)$ inner product $\langle .,.\rangle_\sigma$. This condition then naturally extends the classical detailed balance condition (μ-DBC). Since we will exclusively be working in this setting, from now on, we take the above equation as our definition of KMS-symmetry.

In the case of an $\mathbb{L}_2(\sigma)$-QMS that satisfies ω-DBC, Theorem 5.1.3 provides an expression for the generator $\mathcal{L}_2$ that will prove very useful in Section 5.4.

Theorem 5.2.6. *Assume that $(\mathcal{P}_t)_{t\geq 0}$ is the minimal, unital QMS associated to a form generator in the generalized GKLS form (5.9), with operators L_j, L_j^*, K and K^* densely defined on $\mathcal{V} := \mathrm{span}\{\psi_i, i \in \mathbb{N}\}$, where $\sigma = \sum_i \lambda_i|\psi_i\rangle\langle\psi_i|$ is a normal, faithful state. We further assume that $\sum_j \lambda_j |\langle\psi_j, L_i\psi_j\rangle| < \infty$, $\sum_j \lambda_j |\langle\psi_j, K\psi_j\rangle| < \infty$ and $|\mathcal{J}| < \infty$. Then,*

(a) *If $(\mathcal{P}_t)_{t\geq 0}$ commutes with the modular automorphism group $(\alpha_s^\sigma)_{s\in\mathbb{R}}$, the operators L_j can be chosen without loss of generality so that the following algebraic conditions hold: for any $j \in \mathcal{J}$, there exist $k, l \in \mathbb{N}$ and constants $\omega_j = \ln\lambda_k - \ln\lambda_l$ such that, for all $z \in \mathbb{C}$, and all $\psi \in \mathcal{V}$:*

(i) $\sigma^{iz} L_j \, \sigma^{-iz} \, \psi = \mathrm{e}^{i\,z\omega_j} \, L_j \, \psi$

(ii) $\sigma K \, \psi = K\sigma \, \psi$

(iii) $\mathrm{Tr}(\sigma L_i) = \mathrm{Tr}(\sigma L_i^*) = \mathrm{Im}\,\mathrm{Tr}(\sigma K) = \mathrm{Im}\,\mathrm{Tr}(\sigma K^*) = 0$.

(b) *Moreover, if the semigroup satisfies ω-DBC with respect to the state σ, then the operator K can be assumed to satisfy $K = K^*$ on $\mathcal{V}$, and one can choose $\{\mathrm{e}^{-\frac{\omega_r}{2}} L_r\}_{r\in\mathcal{J}} = \{L_r^*\}_{r\in\mathcal{J}}$.*

Proof. (a) By Theorem 5.1.3, we can assume without loss of generality that $\mathrm{Tr}(\sigma L_r)$, $\mathrm{Tr}(\sigma K) = 0$. Moreover, since $(\mathcal{P}_t)_{t\geq 0}$ commutes with $(\alpha_s^\sigma)_{s\in\mathbb{R}}$, it is covariant with respect to the representation $s \mapsto V(s) := \sigma^{is}$ of $\mathbb{R}$ on $\mathcal{H}$. Adapting the proof of Theorem 2 of [Holevo, 1995a], since $V(s)\mathcal{V} \subset \mathcal{V}$, and since $V(s)\sigma V(s)^* = \sigma$, the operators L_r and K can, without loss of generality, be replaced by $L_r'(s)$ and $K'(s)$ that satisfy Equations (5.10) and (5.11) for any $s \in \mathbb{R}$ and unitary matrices $U(s) \in \mathcal{U}(\mathbb{C}^{\mathcal{J}})$. Moreover, by the covariance property, the operators $L_r'(s)$ and $K'(s)$ can be chosen as follows:

$$L_r'(s) = \sigma^{is} L_r \sigma^{-is} = \sum_j U_{rj}(s) L_j \tag{5.17}$$

$$K'(s) = \sigma^{is} K \sigma^{-is} = K \tag{5.18}$$

(the vectors $(\alpha_i(s))_{i\in\mathcal{J}}$ and numbers $\beta(s)$ appearing in Theorem 5.1.3 can be taken to be null because of the invariance of the normalization condition under the action of the modular group). One can easily verify from Equation (5.17) that the unitaries $U(s)$ form a group $U(s)U(t) = U(s+t)$, $s, t \in \mathbb{R}$. Therefore, it admits a self-adjoint generator H. Without loss of generality, one can then assume that $(L_j)_{j\in\mathcal{J}}$ are the coefficients of a $|\mathcal{J}|$-dimensional vector $\mathbf{L}$ written in the eigenbasis of H, so that $(U(s)\mathbf{L})_r = \mathrm{e}^{\omega_r is} L_r$ for all $s \in \mathbb{R}$. Therefore $\alpha_s^\sigma(L_r) = \mathrm{e}^{\omega_r is} L_r$ on $\mathcal{V}$ for any $r \in \mathcal{J}$. In particular, for any $j, k \in \mathbb{N}$ and $r \in \mathcal{J}$:

$$\left(\frac{\lambda_j}{\lambda_k}\right)^{is} \langle \psi_j, L_r\psi_k \rangle = \langle \psi_j, \sigma^{is} L_r \sigma^{-is}\psi_k \rangle = \mathrm{e}^{i\omega_r s}\langle \psi_j, L_r\psi_k \rangle.$$

This implies that, unless $\langle \psi_j, L_r\psi_k \rangle = 0$, for any $r \in \mathcal{J}$, there exist $j, k \in \mathbb{N}$ such that $\omega_r = \ln(\lambda_j) - \ln(\lambda_k)$. By density of $\mathcal{V}$ in $\mathcal{H}$, for any $z \in \mathbb{C}$, and any $\psi \in \mathcal{V}$:

$$\sigma^{iz} L_r \sigma^{-iz}\psi = \mathrm{e}^{i\omega_r z} \, L_r\psi.$$

(b) This is inspired by [Fagnola and Umanitá, 2007]: we assume now that $(\mathcal{P}_t)_{t\geq 0}$ satisfies ω-DBC with respect to σ. From Equation (5.13), we know that for any $k \in \mathbb{N}$ and $\psi \in \mathcal{V}$:

$$K \, \psi = \mathcal{L}_*(|\psi\rangle\langle\psi_k|) \, \psi_k - \sum_r \langle L_r\psi_k, \psi_k \rangle \, L_r\psi - \langle K\psi_k, \psi_k \rangle \, \psi.$$

By the adjoint of Equation (5.13), we find similarly the following:

$$K^* \psi = \mathcal{L}_*(|\psi\rangle\langle\psi_k|) \, \psi_k - \langle \psi_k, K\psi_k \rangle \, \psi - \sum_r \langle L_r^* \psi_k, \psi_k \rangle \, L_r^*\psi.$$

Therefore,

$$(K - K^*)\psi = (\langle \psi_k, K\psi_k \rangle - \langle \psi_k, K^*\psi_k \rangle) \, \psi + \sum_r \langle (L_r^* - L_r)\psi_k, \psi_k \rangle \, L_r\psi,$$

After multiplication by λ_k and summation over index k, we get

$$(K - K^*)\,\psi = [\mathrm{Tr}(\sigma K) - \mathrm{Tr}(\sigma K^*)]\,\psi\,,$$

where the term involving operators L_r vanishes because of the normalization conditions $\mathrm{Tr}(\sigma L_r) = 0$. Hence, $K - K^*$ acts as an imaginary constant on $\mathcal{V}$, which is equal to 0 by the normalization condition. To prove the last claim, we start from the detailed balance condition $\mathrm{Tr}(\sigma\,\mathcal{P}_t(A)\,B) = \mathrm{Tr}(\sigma\,A\,\mathcal{P}_t(B))$, and choose $A = |\psi_i\rangle\langle\psi_k|$, $B = |\psi_j\rangle\langle\psi_l|$. After differentiation, we get

$$\mathcal{L}(|\psi_j\rangle\langle\psi_l|)[\psi_k,\,\sigma\psi_i] = \mathcal{L}(|\psi_i\rangle\langle\psi_k|)[\sigma\psi_l,\,\psi_j]\,.$$

By a direct computation, we arrive at

$$\sum_{r\in\mathcal{J}}\langle\sigma L_r^*\psi_l,\,\psi_i\rangle\langle\psi_k,\,L_r^*\psi_j\rangle = \sum_{r\in\mathcal{J}}\langle L_r\psi_k,\,\psi_j\rangle\langle\psi_l,\,L_r\sigma\psi_i\rangle$$
$$= \sum_{r\in\mathcal{J}}\langle L_r\sigma\psi_l,\,\psi_i\rangle\langle\psi_k,\,L_r\psi_j\rangle$$
$$= \sum_{r\in\mathcal{J}}\mathrm{e}^{-\omega_r}\langle\sigma L_r\psi_l,\,\psi_i\rangle\langle\psi_k,\,L_r\psi_j\rangle\,.$$

Then, by weak* density of $\mathcal{A}_0$ in $\mathcal{B}(\mathcal{H})$, we get that for all $\varphi, \psi \in \mathcal{V}$ and any $X \in \mathcal{B}(\mathcal{H})$:

$$\sum_{r\in\mathcal{J}}\langle\sigma L_r^*\psi_l,\,XL_r^*\psi_j\rangle = \sum_{r\in\mathcal{J}}\mathrm{e}^{-\omega_r}\langle\sigma L_r\psi_l,\,XL_r\psi_j\rangle$$

and therefore $\{\mathrm{e}^{-\frac{\omega_r}{2}}L_r\}_{r\in\mathcal{J}} = \{L_r^*\}_{r\in\mathcal{J}}$, without loss of generality. $\qquad\square$

The above theorem implies the existence of the following symmetric GKLS form for the form-generator of a QMS satisfying ω-DBC with respect to a normal faithful state σ as well as the domain conditions of Theorem 5.2.6: there exist operators $\{\tilde{L}_r\}_{r\in\mathcal{J}} = \{\tilde{L}_r^*\}_{r\in\mathcal{J}}$ and K defined on $\mathcal{V}$, with $K = K^*$ pointwise in $\mathcal{V}$, satisfying (i)-(ii)-(iii) of Theorem 5.2.6, and such that for any $X \in \mathcal{B}(\mathcal{H})$ and any $\psi, \varphi \in \mathcal{V}$:

$$\mathcal{L}(X)[\psi,\varphi] = \langle\psi,\,XK\varphi\rangle + \langle K\psi,\,X\varphi\rangle + \sum_{r\in\mathcal{J}}\mathrm{e}^{-\frac{\omega_r}{2}}\langle\tilde{L}_r\psi,\,X\tilde{L}_r\varphi\rangle\,. \qquad (5.19)$$

Indeed, assuming that $L_r^* = \mathrm{e}^{-\frac{\omega_r'}{2}}L_{r'}$, one simply needs to pick $\tilde{L}_{r'} := \mathrm{e}^{\frac{\omega_r'}{4}}L_r$. This form reduces to the one that was recently given by [Carlen and Maas, 2017] in the finite dimensional case: since $\mathcal{L}(\mathbb{1}) = 0$, $K = -\frac{1}{2}\sum_{r\in\mathcal{J}}\mathrm{e}^{-\frac{\omega_r}{2}}L_r^*L_r$, so that

$$\mathcal{L}(X) = \sum_{r\in\mathcal{J}}\mathrm{e}^{-\omega_r/2}\left(\tilde{L}_r^*[X,\tilde{L}_r] + [\tilde{L}_r^*,X]\tilde{L}_r\right)\,. \qquad (5.20)$$

5.2.3. Dirichlet forms

The analysis of QMS and their contractivity properties is greatly simplified by the introduction of a Dirichlet form[1]: let $(\mathcal{P}_t)_{t\geq0}$ be an $\mathbb{L}_2(\sigma)$ QMS with associated generator $(\mathcal{L}_2, \mathrm{dom}(\mathcal{L}_2))$. Then, for $X, Y \in \mathrm{dom}(\mathcal{L}_2)$, the mapping

$$\mathcal{E}_{2,\mathcal{L}}(X,Y) := -\langle X,\,\mathcal{L}_2(Y)\rangle_\sigma$$

---

[1] For a thorough study of Dirichlet forms on general von Neumann algebras, we refer to [Cipriani, 1997].

is called the *Dirichlet form* associated to $(\mathcal{P}_t)_{t\geq 0}$. In fact, symmetric, weak* continuous positive contraction semigroups are fully characterized by their Dirichlet forms. This noncommutative generalisation of the so-called Beurling–Deny characterisation of the form generators of classical symmetric Markov semigroups was studied in the general von Neumann algebraic setting in [Cipriani, 1997, Goldstein and Lindsay, 1995, Goldstein and Lindsay, 1999].

In Section 3.2 we introduced the logarithmic Sobolev inequality $\mathrm{LSI}_p(c, d)$ where the right hand side involves a classical Dirichlet form $\mathcal{E}(f^{p-1}, f)$. We will see in Part IV that, in order to define a quantum logarithmic Sobolev inequality as the infinitesimal version of the hypercontractivity property of a QMS, one needs to introduce a parametrized family of such forms: For any $p > 1$, the *Dirichlet form of order* p associated to the QMS $(\mathcal{P}_t)_{t\geq 0}$ is defined as follows for all $X \in \mathrm{dom}(\mathcal{L}_p)$:[2]

$$\mathcal{E}_{p,\mathcal{L}}(X) := -\frac{p\hat{p}}{4}\, \mathrm{Re}(\langle \mathrm{I}_{\hat{p},p}(X), \mathcal{L}_p(X)\rangle_\sigma). \qquad (5.21)$$

Since $(\mathcal{P}_t)_{t\geq 0}$ is contractive in each $\mathbb{L}_p(\sigma)$ space, the Lumer-Phillips theorem 5.1.1 directly implies the following:

Proposition 5.2.7. *For any* $p > 1$ *and any positive, semidefinite operator* $X \in \mathrm{dom}(\mathcal{L}_p)$:

$$\mathcal{E}_{p,\mathcal{L}}(X) \geq 0. \qquad (5.22)$$

Proof. Let $p > 1$. Then, by Proposition 5.2.1 the semigroup $(\mathcal{P}_t)_{t\geq 0}$ is $\mathbb{L}_p(\sigma)$ contractive. Hence, by part (iv) of Theorem 5.1.1, $\mathcal{L}_p$ is dissipative, which means that there exists $X_p^* \in \mathbb{L}_p(\sigma)^* \simeq \mathbb{L}_{\hat{p}}(\sigma)$ such that $X_p^*(X) = \|X\|_{\mathbb{L}_p(\sigma)}^2$ and $\mathrm{Re}\,(X_p^*(\mathcal{L}_p(X))) \leq 0$. However, by Proposition 1.1.4, the only operator X_p^* that satisfies $X_p^*(X) = \|X\|_{\mathbb{L}_p(\sigma)}^2$ is $X_p^* \equiv \|X\|_{\mathbb{L}_p(\sigma)}^{2-p}\, \mathrm{I}_{\hat{p},p}(X)$. Therefore:

$$\mathrm{Re}\,(X_p^*(\mathcal{L}_p(X))) = \|X\|_{\mathbb{L}_p(\sigma)}^{2-p}\, \mathrm{Re}(\langle \mathrm{I}_{\hat{p},p}(X), \mathcal{L}_p(X)\rangle_\sigma) = -\frac{4\,\|X\|_{\mathbb{L}_p(\sigma)}^{2-p}}{\hat{p}p}\, \mathcal{E}_{p,\mathcal{L}}(X) \leq 0 \quad \Rightarrow \quad \mathcal{E}_{p,\mathcal{L}}(X) \geq 0.$$

$\square$

In Section 7.4, we will need to extend the definition of $\mathcal{E}_{p,\mathcal{L}}$ to $p \in \mathbb{R}\backslash\{0,1\}$. The same definition will be of use, after replacing $\mathcal{L}_p$ by the generator $\mathcal{L}_2$ of the QMS: assume that there exists an algebra $\mathcal{A}_0$ that satisfies Condition 5.2.2. Then, for all $X \in \mathcal{A}_0^+$ and $p \in \mathbb{R}\backslash\{0,1\}$,

$$\mathcal{E}_{p,\mathcal{L}}(X) := -\frac{p\hat{p}}{4}\, \langle \mathrm{I}_{\hat{p},p}(X), \mathcal{L}(X)\rangle_\sigma. \qquad (5.23)$$

This definition coincides with the one given in Equation (5.21) for $X \in \mathcal{A}_0^+$ and $p > 1$. The verification of the following properties of the Dirichlet form is easy.

Lemma 5.2.8. *Let* $X \in \mathcal{A}_0^+$:

(i) $\mathcal{E}_{\hat{p},\mathcal{L}}(\mathrm{I}_{\hat{p},2}(X)) = \mathcal{E}_{p,\mathcal{L}}(\mathrm{I}_{p,2}(X))$ *for all* $p \in \mathbb{R}\backslash\{0,1\}$.

(ii) $\mathcal{E}_{p,\mathcal{L}}(cX) = c^p\, \mathcal{E}_{p,\mathcal{L}}(X)$ *for any* $p \in \mathbb{R}\backslash\{0,1\}$ *and* $c \geq 0$.

---

[2]The normalization chosen here differs from the one of [Kastoryano and Temme, 2013]. The advantage of our normalization resides in that the quantity stays positive even for $p < 1$, which will prove useful in Section 7.4.

5.3. Non-commutative differential calculus

5.3.1. Derivations and differential operators

In Part II, we claimed that the proof of the equivalence between hypercontractivity and the logarithmic Sobolev inequality mainly consists in the use of a differential calculus that is based on the existence of a certain chain rule. This is due to the fact that the property of $\mathbb{L}_p$-regularity of Dirichlet forms depends on it.

From this observation, the question of the existence of a noncommutative differential calculus becomes central. In this section, we recall the concept of a derivation on the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on a separable Hilbert space, and explain how in this framework, the generator of a QMS should be interpreted as a noncommutative second order differential operator. We then describe a noncommutative chain rule due to [Birman and Solomyak, 1993] that is the main ingredient in the proof of $\mathbb{L}_p$-regularity of QMS satisfying ω-DBC. This chain rule will also prove to be very useful in the development of a quantum theory of Ricci curvature (see Chapter 12).

The study of noncommutative differential calculus on abstract C*-algebras can be traced back to [Guido et al., 1996, Cipriani and Sauvageot, 2003]. Given a C*-algebra $\mathcal{A}$, an $\mathcal{A} - \mathcal{A}$ *bimodule* $\mathcal{K}$ is a Hilbert space together with commuting left and *right* representations of $\mathcal{A}$.

Definition 5.3.1 (Derivations)**.** A linear map $\nabla : \mathrm{dom}(\nabla) \to \mathcal{K}$ defined on a subspace $\mathrm{dom}(\nabla)$ of $\mathcal{A}$ is then called a *derivation* if

- $\mathrm{dom}(\nabla)$ is a subalgebra of $\mathcal{A}$;

- The *Leibniz rule* holds: for any $X, Y \in \mathrm{dom}(\nabla)$,

$$\nabla(XY) = \nabla(X)\,Y + X\,\nabla(Y),$$

where, given $X \in \mathcal{A}$ and $\mathbf{V} \in \mathcal{K}$, $X\,\mathbf{V}$, resp. $\mathbf{V}\,X$, stands for the operation of left-multiplication, resp. right-multiplication, of $\mathbf{V}$ by X.

Example 5.3.2 (Gradients on $\mathbb{R}^n$)**.** Let $\mathcal{A} = C^0(\mathbb{R}^n)$ be the algebra of continuous functions $f : \mathbb{R}^n \to \mathbb{R}$ that vanish at infinity, and let $\mathcal{S}(\mathbb{R}^n)$ be the Schwartz space of smooth, rapidly decaying functions:

$$\mathcal{S}(\mathbb{R}^n) = \left\{ f \in C^\infty(\mathbb{R}^n) : \sup_{x \in \mathbb{R}^n} |x^\alpha D^\beta f(x)| < \infty \ \forall \alpha, \beta \in \mathbb{N}^n \right\},$$

where, given $x \in \mathbb{R}^n$ and $\alpha = (\alpha_1, ..., \alpha_n)$, $\beta = (\beta_1, ..., \beta_n) \in \mathbb{N}^n$, $x^\alpha := x_1^{\alpha_1} ... x_n^{\alpha_n}$, and $D^\beta f := \partial_{x_1}^{\beta_1} ... \partial_{x_n}^{\beta_n}$. The usual gradient

$$\nabla : \mathcal{S}(\mathbb{R}^n) \to \mathbb{L}_2(\mu_{\mathrm{Leb}}(\mathbb{R}^n), \mathbb{R}^n) := \left\{ v : \mathbb{R}^n \to \mathbb{R}^n, \ \int_{\mathbb{R}^n} \|v(x)\|_{\mathbb{R}^n}^2 \, \mu_{\mathrm{Leb}}(dx) < \infty \right\}$$

is a derivation in the sense of Definition 5.3.1.

Example 5.3.3 (Discrete derivation on finite sample spaces)**.** Let $\mathcal{A} = F(I)$ the algebra of functions defined on a finite set I and let $\pi : \mathcal{M}(I \times I)$ be a positive measure on the cartesian product $I \times I$. Then, the map $\nabla : F(I) \to \mathbb{L}_2(I \times I, \pi)$ defined as

$$\forall f \in F(I), \ \forall x, y \in I, \ \ \nabla(f)(x, y) := f(x) - f(y).$$

is a derivation in the sense of Definition 5.3.1.

Example 5.3.4 (Quantum commutations)**.** In the next section, we will need to manipulate commutators $[A, B]$ between two possibly unbounded operators A and B. Let $\mathcal{A} = \mathcal{B}(\mathcal{H})$ be the algebra of bounded linear operators on a separable Hilbert space $\mathcal{H}$, and let $L : \mathrm{dom}(L) \to \mathcal{H}$ be a (possibly unbounded) closed, densely defined linear operator on $\mathcal{H}$. Let X, Y two closed operators on $\mathcal{H}$ with dense domains. Assume there exists a dense subspace $\mathcal{V}$ of $\mathcal{H}$ such that $\mathcal{V} \subset \mathrm{dom}(X), \mathrm{dom}(X^*), \mathrm{dom}(Y), \mathrm{dom}(L), \mathrm{dom}(L^*)$. Moreover, assume that the following sesquilinear form defined, for any $\psi, \varphi \in \mathcal{V}$, as

$$(\psi, \varphi) \mapsto \tilde{C}_L^{X,Y}[\psi, \varphi] := \langle L^*\psi, \, Y\varphi \rangle - \langle X^*\psi, \, L\varphi \rangle \tag{5.24}$$

is continuous on $\mathcal{H} \times \mathcal{H}$. To this form one can then associate a bounded operator, which we denote by $C_L^{X,Y}$, such that $\tilde{C}_L^{X,Y}[\psi, \varphi] = \langle \psi, C_L^{X,Y}\varphi \rangle$, $\psi, \varphi \in \mathcal{V}$. Formally, " $C_L^{X,Y} = LY - XL$ ".

In the case when $X = Y$, we will denote $C_L^{X,X}$ by $[L, X]$. If X is bounded, the commutator $[L, X]$ can be interpreted as a derivation in the sense of Definition 5.3.1: assume that $\mathcal{V} := \mathrm{span}\{\psi_i\}$ for a given orthonormal basis $\{\psi_i\}_{i\in\mathbb{N}}$ of $\mathcal{H}$, and let $\sigma := \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ be a faithful normal state in $\mathcal{D}(\mathcal{H})$. Since $\mathbb{L}_2(\sigma)$ is invariant under left and right multiplication by $\mathcal{B}(\mathcal{H})$, it defines a $\mathcal{B}(\mathcal{H}) - \mathcal{B}(\mathcal{H})$ bimodule. Then, define the operator $\nabla_L : \mathrm{dom}(\nabla_L) \to \mathbb{L}_2(\sigma)$ as follows:

$$\mathrm{dom}(\nabla_L) = \left\{ X \in \mathcal{B}(\mathcal{H}) : X\mathcal{V} \subset \mathcal{V}, \tilde{C}_L^{X,X} \text{ continuous in } \mathcal{H} \times \mathcal{H} \, , \, [L, X] \in \mathbb{L}_2(\sigma) \right\}, \quad \nabla_L(X) := [L, X] \, .$$

This set is an algebra: for any $X, Y \in \mathrm{dom}(\nabla_L)$, $XY\mathcal{V} \subset \mathcal{V}$ and for all $\psi, \varphi \in \mathcal{V}$:

$$\begin{aligned}
\tilde{C}_L^{XY,XY}[\psi, \varphi] &= \langle L^*\psi, \, XY\varphi \rangle - \langle (XY)^*\psi, \, L\varphi \rangle \\
&= \langle L^*\psi, \, X(Y\varphi) \rangle - \langle X^*\psi, \, L(Y\varphi) \rangle \\
&\quad + \langle L^*(X^*\psi), \, Y\varphi \rangle - \langle Y^*(X^*\psi), \, L\varphi \rangle \, .
\end{aligned}$$

The above also justifies the Leibnitz rule

$$\nabla_L(XY) = \nabla_L(X)Y + X\nabla_L(Y) \, ,$$

and that $\nabla_L(XY) \in \mathbb{L}_2(\sigma)$. Therefore, $\mathrm{dom}(\nabla_L)$ is an algebra.

For instance, let $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$, and let Q, P denote the position and momentum operators defined in Section 0.2. Then, for any smooth, bounded function φ on $\mathbb{R}$, $\varphi(Q) \in \mathrm{dom}(\nabla_P)$ and, on $\mathcal{S}(\mathbb{R})$,

$$\nabla_P(\varphi(Q)) = \frac{1}{i} \, \varphi'(Q) \, .$$

More generally, given a vector $\mathbf{L} := (L_j)_{j\in\mathcal{J}}$ of closed, densely defined operators L_j such that L_j and L_j^* all have common dense domain $\mathcal{V}$, we introduce the *noncommutative gradient* as follows:

$$\nabla_{\mathbf{L}} : \begin{cases} \bigcap_{j\in\mathcal{J}} \mathrm{dom}(\nabla_{L_j}) \to \bigoplus_{j\in\mathcal{J}} \mathbb{L}_2(\sigma) \\ X \mapsto (\nabla_{L_1}(X), ..., \nabla_{L_n}(X)) \end{cases} . \tag{5.25}$$

The *noncommutative divergence* is then defined in the weak sense as follows: for any $\mathbf{V} \in \bigoplus_{j\in\mathcal{J}} \mathcal{T}_2(\mathcal{H})$ and $X \in \mathcal{A}_0 \equiv \mathrm{span}\{|\psi_i\rangle\langle\psi_j|\}$,

$$\langle X, \mathrm{div}_{\mathbf{L}}(\mathbf{V}) \rangle_{\mathrm{HS}} = - \sum_{j\in\mathcal{J}} \langle \nabla_{L_j}(X), V_j \rangle_{\mathrm{HS}} \, .$$

Finally, the *noncommutative Laplacian* $\Delta_{\mathbf{L}}$ is weakly defined as follows:

$$\Delta_{\mathbf{L}} := \operatorname{div}_{\mathbf{L}} \circ \nabla_{\mathbf{L}} = -\sum_{j \in \mathcal{J}} \nabla^*_{L_j} \circ \nabla_{L_j}\,.$$

5.3.2. Double operator integrals and Noncommutative chain rule:

In order to extend some of the calculations of Part II involving differentiations to our noncommutative setting, we need the notion of a noncommutative chain rule. Such a chain rule exists and is connected to the abstract theory of double integrals [Daleckii and Krein, 1951, Daletskii and Krein, 1965, Birman and Solomyak, 1967, De Pagter et al., 2002, de Pagter and Sukochev, 2004, Ptapov and Sukochev, 2008, Potapov, D. and Sukochev, F., 2010]. Here, we follow the treatment of [Birman and Solomyak, 1993]. In spirit, given any two spectral projections $\mathcal{B}(\mathbb{R}) \ni A \mapsto E_0(A), E_1(A)$ and a function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$, one wants to make sense of the following integral:

$$T_f : X \mapsto \int_{\mathbb{R}^2} f(x,y)\, E_0(dx)\, X\, E_1(dy)\,.$$

More precisely, let A_0, A_1 be (possibly unbounded) self-adjoint operators in a separable Hilbert space $\mathcal{H}$, of respective domains $\operatorname{dom}(A_k)$ and associated spectral measures E_k, $k = 0, 1$. The measures E_k induce new spectral measures acting on $\mathcal{T}_2(\mathcal{H})$: for $T \in \mathcal{T}_2(\mathcal{H})$, set $\mathcal{E}_0(.)T = E_0(.)T$ and $\mathcal{E}_1(.)T = TE_1(.)$. The spectral measures $\mathcal{E}_k$ commute, and their product $\tilde{\mathcal{E}}$ is a spectral measure on $\mathcal{T}_2(\mathcal{H})$ on the Borel subsets of $\mathbb{R}^2$. Now, given an arbitrary $\tilde{\mathcal{E}}$-measurable and $\tilde{\mathcal{E}}$-almost everywhere finite function h defined on $\operatorname{sp}(A_0) \times \operatorname{sp}(A_1)$, we associate the integral

$$T_h = \int h\, d\tilde{\mathcal{E}}\,.$$

T_h is called the *transformer* of h. The transformer T_h defines a bounded linear operator on $\mathcal{T}_2(\mathcal{H})$, since $\|T_h : \mathcal{T}_2(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})\| = \|h\|_{\mathbb{L}_\infty(\mathbb{R}^2)}$. More generally, given $1 \leq p \leq \infty$, we are interested in the class $[T_p(\mathcal{H})]$ of transformers T_h that are bounded in $\mathcal{T}_p(\mathcal{H})$ (with $\mathcal{T}_\infty(\mathcal{H}) := \mathcal{B}(\mathcal{H})$), i.e. such that $\|T_h : \mathcal{T}_p(\mathcal{H}) \to \mathcal{T}_p(\mathcal{H})\| < \infty$. On any of these ideals, the map $f \mapsto T_f$ is

(i) *linear*: $T_{\lambda f + g} = \lambda T_f + T_g$;

(ii) *multiplicative*: $T_{fg} = T_f \circ T_g$;

(iii) *involutive*: when $p = 2$, $T_f^* = T_{\bar{f}}$, where the adjoint is taken with respect to the Hilbert-Schmidt inner product;

(iv) *positivity preserving*: $f \geq 0$ implies that T_f is a positive operator on $\mathcal{T}_2(\mathcal{H})$.

In particular, given a complex-valued Borel function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$, we are interested in functions $\tilde{f}$ of the following form

$$\tilde{f}(x,y) := \begin{cases} \dfrac{f(x) - f(y)}{x - y} & \text{if } (x,y) \in \operatorname{sp}(H_0) \times \operatorname{sp}(H_1) \\ 0 & \text{else} \end{cases} \tag{5.26}$$

Now, let X, Y and L be defined as in Example 5.3.4 on a dense subspace $\mathcal{V}$. Assume moreover that X and Y are self-adjoint, and that for any $g \in C_c^\infty(\mathbb{R})$, $g(X)\mathcal{V}, g(Y)\mathcal{V} \subset \mathcal{V}$. The following theorem from [Birman and Solomyak, 1993] provides a *noncommutative chain rule* for the derivation ∇_L:

Theorem 5.3.5. *Given $1 \le p \le \infty$, assume that the operator $C_L^{X,Y}$ defined through Equation (5.24) is in $\mathcal{T}_p(\mathcal{H})$ and let $f : \mathbb{R} \to \mathbb{C}$ be a Borel function such that the transformer $T_{\tilde{f}}^{X,Y}$ associated to the spectral measures of X and Y is in $[\mathcal{T}_p(\mathcal{H})]$. Then, the form $\tilde{C}_{L,f}^{X,Y}$ defined for any $\psi, \varphi \in \mathcal{V}$ as*

$$\tilde{C}_{L,f}^{X,Y}[\psi, \varphi] := \langle L^*\psi, f(Y)\varphi \rangle - \langle \bar{f}(X)\psi, L\varphi \rangle$$

is continuous on $\mathcal{H} \times \mathcal{H}$. Moreover, the associated bounded operator $C_{L,f}^{X,Y}$ can be represented as follows:

$$C_{L,f}^{X,Y} = T_{\tilde{f}}^{X,Y}(C_L^{X,Y}) \in \mathcal{T}_p(\mathcal{H}). \qquad \text{(n-c chain rule)}$$

Remark 5.3.6. In the case when $X = Y$ the n-c chain rule simplifies to

$$\nabla_L f(X) = T_{\tilde{f}}^{X,X}(\nabla_L X),$$

where $\nabla_L X := [L, X]$. This should be interpreted as a noncommutative extension of the usual chain rule $(f \circ g)' = f' \circ g \, g'$ for functions.

5.4. Stroock Varopoulos inequality and regularity of Dirichlet forms

In Section 2.3 and Section 2.4, we introduced regularity properties of the Dirichlet form of a classical Markov semigroup that play a key role in the proof of the equivalence between the logarithmic Sobolev inequality and hypercontractivity. In particular, we saw that if $(P_t)_{t \ge 0}$ is an $\mathbb{L}_2(\mu)$-Markov semigroup defined on a measure space $(E, \mathcal{F}, \mu)$ that is symmetric with respect to the measure μ, then, for any $p > 0$ and $f \ge 0$:

$$\mathcal{E}(f^{p-1}, f) \ge \frac{4}{p\hat{p}} \mathcal{E}(f^{p/2}, f^{p/2}). \qquad (5.27)$$

When $(P_t)_{t \ge 0}$ is a reversible Markov semigroup defined on a finite sample space, inequality (5.27) is actually a simple consequence of the so-called Stroock Varopoulos inequality (first established by [Carlen et al., 1987, Varopoulos, 1985] in the case $p = 2$ and $q \in (1, 2]$, see [Mossel et al., 2013] for the more general result provided here):

Theorem 5.4.1 (Stroock-Varopoulos inequality)**.** *Let $p, q \in (0, 2] \backslash \{1\}$ and $p > q$. Then for every function $g > 0$:*

$$q\hat{q}\,\mathcal{E}(g^{1/q}, g^{1/\hat{q}}) \ge p\hat{p}\,\mathcal{E}(g^{1/p}, g^{1/\hat{p}}). \qquad (5.28)$$

When q spans $(1, 2]$, $\hat{q}$ spans $[2, \infty)$, so that one recovers (5.27) after taking $p = 2$ and $g = f^{\hat{q}}$ in (5.28).

A fully quantum generalization of these results is currently lacking. Here we provide a partial result in this direction: let $(\mathcal{P}_t)_{t \ge 0}$ be a quantum Markov semigroup on the space $\mathcal{B}(\mathcal{H})$, $\mathcal{H}$ separable, with faithful normal invariant state $\sigma = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$. In this section, we assume that the QMS $(\mathcal{P}_t)_{t \ge 0}$ satisfies the conditions of Proposition 5.2.3, so that its form generator takes the form of Equation (5.20), and choose $\mathcal{A}_0 := \text{span}\{|\psi_i\rangle\langle\psi_j|, i, j \in \mathbb{N}\}$. By Proposition 5.2.3, the algebra $\mathcal{A}_0$ satisfies Condition 5.2.2,

and for any $X \in \mathcal{A}_0^+$:

$$\mathcal{E}_{p,\mathcal{L}}(X) = -\frac{p\hat{p}}{4} \langle \mathrm{I}_{\hat{p},p}(X), \mathcal{L}(X) \rangle_\sigma$$

$$= -\frac{p\hat{p}}{4} \sum_{r \in \mathcal{J}} \mathrm{e}^{-\frac{\omega_r}{2}} \langle \mathrm{I}_{\hat{p},p}(X), \nabla_{\tilde{L}_r^*}(X) \tilde{L}_r - \tilde{L}_r^* \nabla_{\tilde{L}_r}(X) \rangle_\sigma,$$

where $\nabla_{\tilde{L}_r} : \mathrm{dom}(\nabla_{\tilde{L}_r}) \to \mathbb{L}_2(\sigma)$, $\nabla_{\tilde{L}_r^*} : \mathrm{dom}(\nabla_{\tilde{L}_r^*}) \to \mathbb{L}_2(\sigma)$ are defined as in Example 5.3.4, with $\mathcal{V} := \mathrm{span}\{\psi_i, i \in \mathbb{N}\}$. In particular, for $p = 2$, $\mathcal{E}_{2,\mathcal{L}}$ satisfies the following *integration by parts formula* analogous to IBP: for any $X, Y \in \mathcal{A}_0$:

$$\mathcal{E}_{2,\mathcal{L}}(X,Y) = -\langle X, \mathcal{L}(Y) \rangle_\sigma = \sum_{r \in \mathcal{J}} \langle \nabla_{\tilde{L}_r} X, \nabla_{\tilde{L}_r} Y \rangle_\sigma. \qquad \text{(n-cIBP)}$$

Theorem 5.4.2 (Quantum Stroock-Varopoulos inequality)**.** *Let $(\mathcal{P}_t)_{t \geq 0}$ be a quantum Markov semigroup satisfying the conditions of Theorem 5.2.6(b). Then, for all $X \in \mathcal{A}_0^+$ and any $0 \leq p \leq q \leq 2$:*

$$\mathcal{E}_{p,\mathcal{L}}\big(\mathrm{I}_{p,2}(X)\big) \geq \mathcal{E}_{q,\mathcal{L}}\big(\mathrm{I}_{q,2}(X)\big).$$

Remark 5.4.3. We should point out that a quantum Stroock-Varopoulos inequality in the special case when σ is the completely mixed state of a QMS defined on a finite dimensional Hilbert space was proven in [Cubitt et al., 2015]. Moreover, sufficient conditions for a finite dimensional QMS to satisfy *strong $\mathbb{L}_p$-regularity* were previously provided in [Kastoryano and Temme, 2013]. The strong $\mathbb{L}_p$-regularity was later proved in [Bardet, 2017] under the condition of ω-DBC.

In the next two subsections, we list two proofs for this theorem. The first one builds on the proof of strong regularity of [Kastoryano and Temme, 2013] and only works in finite dimensions. The second one is more similar to the proof of [Bardet, 2017] and extends to the case of a separable Hilbert space. We present the two proofs since they are different in nature and we hope that the ideas used in both can be useful elsewhere.

5.4.1. First proof of Theorem 5.4.2 in the finite dimensional case

In this section, we assume that the QMS $(\mathcal{P}_t)_{t \geq 0}$ is defined on $\mathcal{B}(\mathcal{H})$, where $d_{\mathcal{H}} < \infty$, and satisfies σ-DBC.

Lemma 5.4.4. *For every $t \geq 0$ there are operators $R_k \in \mathcal{B}(\mathcal{H})$ and $\nu_k > 0$ such that $\Delta_\sigma(R_k) = \nu_k R_k$,*

$$\mathcal{P}_t(X) = \sum_k R_k X R_k^*, \qquad (5.29)$$

and $\sum_k R_k R_k^* = \mathbb{1}$.

Proof. By Lemma 5.2.5 the Lindblad generator $\mathcal{L}$ and then $\mathcal{P}_t = \mathrm{e}^{t\mathcal{L}}$ commute with Δ_σ, i.e.,

$$\mathcal{P}_t \circ \Delta_\sigma = \Delta_\sigma \circ \mathcal{P}_t. \qquad (5.30)$$

Fix an orthonormal basis $\{e_i\}_{i=1}^d$ for the underlying Hilbert space $\mathcal{H} = \mathcal{H}_A$ and define

$$\Upsilon := \sum_{i=1}^d e_i \otimes e_i \in \mathcal{H}_{AB},$$

where $\mathcal{H}_B$ is isomorphic to $\mathcal{H}_A$. It is not hard to verify that for any matrix M we have

$$(M_A \otimes \mathbb{1}_B)|\Upsilon\rangle = \mathbb{1}_A \otimes M_B^T|\Upsilon\rangle, \tag{5.31}$$

where the transpose is with respect to the basis $\{e_i\}_{i=1}^d$. The Choi-Jamiolkowski representation of $\mathcal{P}_t$ is

$$J_{AB} := (\mathcal{P}_t \otimes \mathbb{1}_B)(|\Upsilon\rangle\langle\Upsilon|).$$

Then using (5.31) it is not hard to verify that (5.30) translates to

$$(\sigma_A^{-1} \otimes \sigma_B^T)J_{AB} = J_{AB}(\sigma_A^{-1} \otimes \sigma_B^T).$$

That is, J_{AB} and $\sigma_A^{-1} \otimes \sigma_B^T$ commute. On the other hand, J_{AB} is positive semidefinite since it is the Choi-Jamiolkowski representation of a completely positive map. Therefore, J_{AB} and $\sigma_A^{-1} \otimes \sigma_B^T$ can be simultaneously diagonalized in an orthonormal basis, i.e., there exists an orthonormal basis $\{v_k\}_{k=1}^{d^2}$ of $\mathcal{H}_{AB}$ such that

$$J_{AB}v_k = \lambda_k v_k \tag{5.32}$$
$$\sigma_A^{-1} \otimes \sigma_B^T v_k = \nu_k^{-1} v_k, \tag{5.33}$$

where $\lambda_k \geq 0$, $\nu_k > 0$. Define the operator V_k by

$$(V_k \otimes \mathbb{1}_B)\Upsilon = v_k.$$

Then again using (5.31), equation (5.33) translates to

$$\sigma^{-1}V_k\sigma = \nu_k^{-1}V_k.$$

Moreover, equation (5.32) means that

$$(\mathcal{P}_t \otimes \mathbb{1}_B)(|\Upsilon\rangle\langle\Upsilon|) = J_{AB} = \sum_k \lambda_k|v_k\rangle\langle v_k| = \sum_k \lambda_k(V_k \otimes \mathbb{1}_B)|\Upsilon\rangle\langle\Upsilon|(V_k^* \otimes \mathbb{1}_B),$$

which gives

$$\mathcal{P}_t(X) := \sum_k \lambda_k V_k X V_k^*.$$

Then letting $R_k := \sqrt{\lambda_k}V_k$ we have $\sigma R_k = \nu_k R_k \sigma$ and (5.29) holds. The other equation comes from $\mathcal{P}_t(\mathbb{1}) = \mathbb{1}$.

Proof of Theorem 5.4.2 ($d_\mathcal{H} < \infty$) **:** For any $t \geq 0$ define the function $h_t : (0, \infty) \to \mathbb{R}$ by

$$h_t(s) := \left\langle I_{2/(2-s),2}(X), \mathcal{P}_t \circ I_{2/s,2}(X) \right\rangle_\sigma.$$

Since by part (ii) of Lemma 5.2.5, $\mathcal{P}_t = e^{t\mathcal{L}}$ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_\sigma$, we have $h_t(2-s) = h_t(s)$ and h_t is symmetric about $s = 1$. Therefore, all the the odd-order derivatives of h_t at $s = 1$ vanish, and we have

$$h_t(s) = h_t(1) + \sum_{j=1}^\infty \frac{c_j}{(2j)!}(s-1)^{2j}, \tag{5.34}$$

where

$$c_j = \frac{d^{2j}}{ds^{2j}} h_t(s)\Big|_{s=1}.$$

We claim that all the even-order derivatives of h_t at $s = 1$ are non-negative, i.e., $c_j \geq 0$. We use Lemma 5.4.4 to verify this. Let R_k's be operators such that

$$\sigma R_k \sigma^{-1} = \nu_k R_k, \tag{5.35}$$

with $\nu_k > 0$ and (5.29) holds. Then letting $Y := \Gamma_\sigma^{1/2}(X)$ and using (5.35) we compute

$$\begin{aligned}
h_t(s) &= \operatorname{Tr}\left[\Gamma_\sigma^{\frac{s}{2}}(Y^{2-s}) \cdot \mathcal{P}_t\left(\Gamma_\sigma^{-\frac{s}{2}}(Y^s)\right)\right] \\
&= \sum_k \operatorname{Tr}\left[Y^{2-s}\sigma^{\frac{s}{4}}R_k\sigma^{-\frac{s}{4}}Y^s\sigma^{-\frac{s}{4}}R_k^*\sigma^{\frac{s}{4}}\right] \\
&= \sum_k \nu_k^{\frac{s}{2}} \operatorname{Tr}\left[Y^{2-s}R_k Y^s R_k^*\right].
\end{aligned}$$

Now diagonalizing Y in its eigenbasis: $Y = \sum_\ell \mu_\ell |e_\ell\rangle\langle e_\ell|$, we find that

$$h_t(s) = \sum_{k,\ell,\ell'} \mu_\ell^2 |\langle e_\ell, R_k e_{\ell'}\rangle|^2 \left(\frac{\sqrt{\nu_k}\,\mu_{\ell'}}{\mu_\ell}\right)^s.$$

Therefore, $h_t(s)$ is a sum of exponential functions with positive coefficients. From this expression it is clear that c_j's are all non-negative. Let us define

$$g_t(s) := \frac{h_t(s) - h_t(0)}{(s-1)^2 - 1} = \sum_{j=1}^{\infty} \frac{c_j}{(2j)!}\left(\sum_{i=0}^{j-1}(s-1)^{2i}\right).$$

From this expression it is clear that $g_t(s)$ is non-decreasing on $[1, +\infty)$. Therefore, $\lim_{t\to 0^+} g_t(s)/t$ is non-decreasing on $[1, +\infty)$. On the other hand, we have $h_t(0) = \operatorname{Tr}(Y^2) = h_0(s)$. We thus can compute

$$\begin{aligned}
\lim_{t\to 0^+} \frac{g_t(s)}{t} &= \frac{1}{(s-1)^2 - 1} \lim_{t\to 0^+} \frac{h_t(s) - h_t(0)}{t} \\
&= \frac{1}{(s-1)^2 - 1} \lim_{t\to 0^+} \frac{h_t(s) - h_0(s)}{t} \\
&= \frac{1}{(s-1)^2 - 1} \frac{\partial}{\partial t} h_t(s)\Big|_{t=0} \\
&= \frac{1}{(s-1)^2 - 1} \left\langle I_{2/(2-s),2}(X), \mathcal{L} \circ I_{2/s,2}(X)\right\rangle_\sigma.
\end{aligned}$$

Therefore

$$s \mapsto \frac{1}{(s-1)^2 - 1}\left\langle I_{2/(2-s),2}(X), \mathcal{L} \circ I_{2/s,2}(X)\right\rangle_\sigma,$$

is non-decreasing on $[1, +\infty)$. Now the desired result follows once we identify $2/s$ with p (and $2/(2-s)$ with $\hat{p}$, its Hölder conjugate).

$\square$

5.4.2. Second proof of Theorem 5.4.2 in the infinite dimensional case

The proof of Theorem 5.4.2 presented in Section 5.4.1 has the disadvantage that it only holds in finite dimensions. This is mostly due to the fact that it relies on the introduction of the unnormalized maximally entangled state Υ that becomes ill defined when $\dim(\mathcal{H}) = \infty$. Here we provide a more

algebraic proof based on the noncommutative chain rule introduced in Theorem 5.3.5. First, for any $r, s \in \mathbb{R} \setminus \{0, 1\}$ and $X \in \mathcal{A}_0^+$, $I_{r,s}(X) \in \operatorname{dom} \nabla_{\tilde{L}_j}$ and the following holds on $\mathcal{V} = \operatorname{span}\{\psi_i, \, i \in \mathbb{N}\}$:

$$
\begin{aligned}
\nabla_{\tilde{L}_j} I_{r,s}(X) &= \tilde{L}_j \, \sigma^{-\frac{1}{2r}} \left(\sigma^{\frac{1}{2s}} X \sigma^{\frac{1}{2s}} \right)^{\frac{s}{r}} \sigma^{-\frac{1}{2r}} - \sigma^{-\frac{1}{2r}} \left(\sigma^{\frac{1}{2s}} X \sigma^{\frac{1}{2s}} \right)^{\frac{s}{r}} \sigma^{-\frac{1}{2r}} \tilde{L}_j \\
&= \mathrm{e}^{-\frac{\omega_j}{2r}} \sigma^{-\frac{1}{2r}} \tilde{L}_j \left(\sigma^{\frac{1}{2s}} X \sigma^{\frac{1}{2s}} \right)^{\frac{s}{r}} \sigma^{-\frac{1}{2r}} - \mathrm{e}^{\frac{\omega_j}{2r}} \sigma^{-\frac{1}{2r}} \left(\sigma^{\frac{1}{2s}} X \sigma^{\frac{1}{2s}} \right)^{\frac{s}{r}} \tilde{L}_j \sigma^{-\frac{1}{2r}} \quad (5.36)
\end{aligned}
$$

Next, define $Y_j \coloneqq \mathrm{e}^{-\frac{\omega_j}{4}} \, \Gamma_\sigma^{\frac{1}{2}}(X)$ and $Z_j \coloneqq \mathrm{e}^{\frac{\omega_j}{4}} \, \Gamma_\sigma^{\frac{1}{2}}(X)$. Using Equation (5.36) we compute

$$
\begin{aligned}
\mathcal{E}_{q,\mathcal{L}}\big(I_{q,2}(X)\big) &= -\frac{q\hat{q}}{4} \big\langle I_{\hat{q},q}\big(I_{q,2}(X)\big), \mathcal{L}\big(I_{q,2}(X)\big) \big\rangle_\sigma \\
&= -\frac{q\hat{q}}{4} \big\langle I_{\hat{q},2}(X), \mathcal{L}\big(I_{q,2}(X)\big) \big\rangle_\sigma \\
&= \frac{q\hat{q}}{4} \sum_j \langle \nabla_{\tilde{L}_j} I_{\hat{q},2}(X), \, \nabla_{\tilde{L}_j} I_{q,2}(X) \rangle_\sigma & (5.37) \\
&= \frac{q\hat{q}}{4} \sum_j \Big\langle \Gamma_\sigma^{-\frac{1}{\hat{q}}}\big(\tilde{L}_j Y_j^{\frac{2}{\hat{q}}} - Z_j^{\frac{2}{\hat{q}}} \tilde{L}_j\big), \, \Gamma_\sigma^{-\frac{1}{q}}\big(\tilde{L}_j Y_j^{\frac{2}{q}} - Z_j^{\frac{2}{q}} \tilde{L}_j\big) \Big\rangle_\sigma & (5.38) \\
&= \frac{q\hat{q}}{4} \sum_j \Big\langle \tilde{L}_j Y_j^{\frac{2}{\hat{q}}} - Z_j^{\frac{2}{\hat{q}}} \tilde{L}_j, \, \tilde{L}_j Y_j^{\frac{2}{q}} - Z_j^{\frac{2}{q}} \tilde{L}_j \Big\rangle_{\mathrm{HS}} & (5.39) \\
&= \frac{q\hat{q}}{4} \sum_j \Big\langle T_{\tilde{f}_{\frac{2}{\hat{q}}}}^{Z_j, Y_j}(\tilde{L}_j Y_j - Z_j \tilde{L}_j), \, T_{\tilde{f}_{\frac{2}{q}}}^{Z_j, Y_j}(\tilde{L}_j Y_j - Z_j \tilde{L}_j) \Big\rangle_{\mathrm{HS}} & (5.40) \\
&= \frac{q\hat{q}}{4} \sum_j \Big\langle \tilde{L}_j Y_j - Z_j \tilde{L}_j, \, \big(T_{\tilde{f}_{\frac{2}{\hat{q}}}}^{Z_j, Y_j}\big)^* T_{\tilde{f}_{\frac{2}{q}}}^{Z_j, Y_j}(\tilde{L}_j Y_j - Z_j \tilde{L}_j) \Big\rangle_{\mathrm{HS}} \\
&= \frac{q\hat{q}}{4} \sum_j \Big\langle \tilde{L}_j Y_j - Z_j \tilde{L}_j, \, T_{\tilde{f}_{\frac{2}{\hat{q}}} \tilde{f}_{\frac{2}{q}}}^{Z_j, Y_j}(\tilde{L}_j Y_j - Z_j \tilde{L}_j) \Big\rangle_{\mathrm{HS}}, & (5.41)
\end{aligned}
$$

where in (5.37) we used n-cIBP, in (5.38) we used (5.36), and in (5.40) we used the chain rule formula of Theorem 5.3.5 for the functions $f_\alpha : x \mapsto x^\alpha$. Finally, in (5.41) we used the fact that $f \mapsto T_f$ is multiplicative and involutive (cf. Section 5.3.2). Now, using the proofs of Theorem 2.1 and Lemma 2.4 of [Mossel et al., 2013], for any $x, y \geq 0$ and $0 \leq p \leq q \leq 2$ we have

$$
q\hat{q}\, (x^{1/\hat{q}} - y^{1/\hat{q}})(x^{1/q} - y^{1/q}) \leq p\hat{p}\, (x^{1/\hat{p}} - y^{1/\hat{p}})(x^{1/p} - y^{1/p}).
$$

This means that for all x, y we have

$$
q\hat{q}\, \tilde{f}_{\frac{2}{\hat{q}}}(x, y)\, \tilde{f}_{\frac{2}{q}}(x, y) \leq p\hat{p}\, \tilde{f}_{\frac{2}{\hat{p}}}(x, y)\, \tilde{f}_{\frac{2}{p}}(x, y).
$$

Hence, by positivity we have

$$
\begin{aligned}
\mathcal{E}_{q,\mathcal{L}}(I_{q,2}(X)) &\leq \frac{p\hat{p}}{4} \sum_j \Big\langle \tilde{L}_j Y_j - Z_j \tilde{L}_j, \, T_{\tilde{f}_{\frac{2}{\hat{p}}} \tilde{f}_{\frac{2}{p}}}^{Z_j, Y_j}(\tilde{L}_j Y_j - Z_j \tilde{L}_j) \Big\rangle_{\mathrm{HS}} \\
&= \mathcal{E}_{p,\mathcal{L}}(I_{p,2}(X)).
\end{aligned}
$$

$\square$

5.5. Examples

5.5.1. Generalized depolarizing semigroup

Perhaps the simplest example of a quantum channel is the *depolarizing channel*: given a state $\rho \in \mathcal{D}(\mathcal{H})$ and $0 \leq p \leq 1$:

$$\mathcal{P}_*^{\mathrm{depol}}(\rho) := (1-p)\,\rho + p\,\frac{\mathbb{1}}{d_\mathcal{H}}\,.$$

The depolarizing channel corresponds to the case when the information on the input state is completely lost (i.e. replaced by the completely mixed state $\mathbb{1}/d_\mathcal{H}$) with some probability p. More generally, consider the channel that replaces the input state with a full-rank fixed state σ:

$$\mathcal{P}_*^{\mathrm{depol},\sigma}(\rho) = (1-p)\,\rho + p\,\sigma\,.$$

The *generalized depolarizing semigroup* $(\mathcal{P}_t^{\mathrm{depol},\sigma})$ constitutes a simple continuous time version of the above channel. It is defined as follows: for all $t \geq 0$ and any $X \in \mathcal{B}(\mathcal{H})$, $\rho \in \mathcal{D}(\mathcal{H})$:

$$\mathcal{P}_t^{\mathrm{depol},\sigma}(X) = \mathrm{e}^{-t}\,X + (1-\mathrm{e}^{-t})\,\mathrm{Tr}(\sigma X) \quad \Leftrightarrow \quad \mathcal{P}_{t*}^{\mathrm{depol},\sigma}(\rho) = \mathrm{e}^{-t}\,\rho + (1-\mathrm{e}^{-t})\,\sigma\,.$$

The generator of $(\mathcal{P}_t^{\mathrm{depol},\sigma})_{t\geq 0}$ is then given by

$$\mathcal{L}_\sigma(X) = \mathrm{Tr}(\sigma X)\,\mathbb{1}_\mathcal{H} - X\,.$$

$(\mathcal{P}_t^{\mathrm{depol}})_{t\geq 0}$ is obviously primitive with respect to σ. It is easy to also see that it satisfies ω-DBC: for any $X, Y \in \mathcal{B}(\mathcal{H})$,

$$\mathrm{Tr}(\sigma X^* \mathcal{P}_t^{\mathrm{depol},\sigma}(Y)) = \mathrm{e}^{-t}\,\mathrm{Tr}(\sigma X^* Y) + (1-\mathrm{e}^{-t})\,\mathrm{Tr}(\sigma X^*)\,\mathrm{Tr}(\sigma Y) = \mathrm{Tr}(\sigma \mathcal{P}_t^{\mathrm{depol},\sigma}(X)^* Y)\,.$$

Therefore, $\mathcal{L}_\sigma$ takes the form of Equation (5.20): given the following eigenvalue decomposition $\sigma := \sum_{i=1}^{d_\mathcal{H}} \lambda_i |\psi_i\rangle\langle\psi_i|$, define the operators $L_{ij} := \sqrt{\lambda_i}\,|\psi_i\rangle\langle\psi_j|$. Hence for any $X \in \mathcal{B}(\mathcal{H})$, $\mathrm{Tr}(\sigma X)\,\mathbb{1}_\mathcal{H} = \sum_{i,j=1}^{d} L_{ij}^* X L_{ij}$, so that

$$\mathcal{L}_\sigma(X) = -\frac{1}{2} \sum_{i,j=1}^{d_\mathcal{H}} L_{ij}^* L_{ij} X - 2 L_{ij}^* X L_{ij} + X L_{ij}^* L_{ij}. \tag{5.42}$$

Moreover, $\Delta_\sigma(L_{ij}) = \lambda_i/\lambda_j L_{ij}$, so that $\omega_{ij} \equiv \log \lambda_j - \log \lambda_i$. Therefore, for any $X \in \mathcal{B}(\mathcal{H})$, Equation (5.42) can be rewritten in the form of Equation (5.20) by taking $\tilde{L}_{ij} = (\lambda_i \lambda_j/4)^{1/4} |\psi_i\rangle\langle\psi_j|$.

5.5.2. Quantum diffusions on phase space

Quantum Ornstein Uhlenbeck semigroups In this section, we introduce the so-called *quantum Ornstein Uhlenbeck semigroup* (also known as the *damped harmonic oscillator*) on the representation of the CCR algebra over a two-dimensional phase space (cf. Section 0.2), that is, the algebra $\mathcal{B}(\mathcal{H})$ of bounded linear operators over the Hilbert space $\mathcal{H} := \mathbb{L}_2(\mathbb{R})$. This Markovian dynamics models the evolution of a two-energy levels atom which traverses a photonic cavity. The dynamics then results from the interaction of photons, modeled as a quantized radiation field, with the incident atom. The Master equation can be obtained by different standard approximation procedures (weak coupling limit, coarse graining, see e.g. [Strunz, 2002, Fagnola et al., 1994, Weidlich and Haake, 1965]).

First, let $0 < \lambda < \mu$, $\nu := \lambda^2/\mu^2$, and defined the state:

$$\sigma_\nu := (1 - \nu) \sum_{n \in \mathbb{N}} \nu^n \, |\psi_n\rangle\langle\psi_n| = \frac{e^{\beta N}}{Z} \,, \tag{5.43}$$

where $\{\psi_n, \, n \in \mathbb{N}\}$ corresponds to the eigenbasis of the number operator N, and $\beta = \ln \nu$ can be interpreted as an inverse temperature. σ_ν can be shown to be a Gaussian state. Then, on the $*$-subalgebra $\mathcal{A}_0 := \mathrm{span}\,\{|\psi_n\rangle\langle\psi_m|, \, m, n \in \mathbb{N}\}$, the generator of the quantum Ornstein Uhlenbeck semigroup takes the following Lindblad form:

$$\mathcal{L}^{\mathrm{qOU}}(X) := -\frac{\mu^2}{2}(a^*aX - 2a^*Xa + Xa^*a) - \frac{\lambda^2}{2}(aa^*X - 2aXa^* + Xaa^*)\,.$$

In Theorem 4.2 and Corollary 4.3 of [Cipriani et al., 2000], the authors proved the existence of a weak* continuous, positive contraction semigroup $(\mathcal{P}_t^{\mathrm{qOU}})_{t \geq 0}$ on $\mathcal{B}(\mathcal{H})$ that is KMS-symmetric with respect to ρ_ν, through the careful analysis of its associated Dirichlet form (see also [Ko and Park, 2004]). They also showed that $(\mathcal{P}_t^{\mathrm{qOU}})_{t \geq 0}$ is a Feller semigroup with respect to the algebra $\mathcal{K}(\mathcal{H})$ of compact operators on $\mathcal{H}$ (see also [Carbone and Fagnola, 2003]). Moreover, they showed that $\mathcal{A}_0$ is a core for the strong generator $\mathcal{L}_{\mathcal{K}(\mathcal{H})}^{\mathrm{qOU}}$ of $(\mathcal{P}_t^{\mathrm{qOU}}|_{\mathcal{K}(\mathcal{H})})_{t \geq 0}$: for any $X \in \mathrm{dom}(\mathcal{L}_{\mathcal{K}(\mathcal{H})}^{\mathrm{qOU}})$, there exists a sequence $\{X_n\}_{n \in \mathbb{N}}$ of operators in $\mathcal{A}_0$ such that

$$\|\mathcal{L}_{\mathcal{K}(\mathcal{H})}^{\mathrm{qOU}}(X_n) - \mathcal{L}_{\mathcal{K}(\mathcal{H})}^{\mathrm{qOU}}(X)\|_\infty + \|X_n - X\|_\infty \underset{n \to \infty}{\to} 0\,.$$

$\mathcal{A}_0$ is also a weak* core for the weak* generator $\mathcal{L}^{\mathrm{qOU}}$, that is, for any $X \in \mathrm{dom}(\mathcal{L}^{\mathrm{qOU}})$, there exists a sequence $\{X_n\}_{n \in \mathbb{N}}$ of operators on $\mathcal{A}_0$ such that for any $T \in \mathcal{T}_1(\mathcal{H}) \simeq \mathcal{B}(\mathcal{H})_*$,

$$\mathrm{Tr}(T(X_n - X)) \underset{n \to \infty}{\to} 0 \qquad \text{and} \qquad \mathrm{Tr}(T(\mathcal{L}(X_n) - \mathcal{L}(X))) \underset{n \to \infty}{\to} 0\,.$$

Another approach is via the form-generator formalism introduced in Section 5.1. In particular, the form-generator associated with $L_1 = a$ and $L_2 = a^*$ satisfies Conditions 5.1.4 and 5.1.5. Hence, it generates a unique weak* continuous QMS $(\mathcal{P}_t^{\mathrm{qOU}})_{t \geq 0}$ on $\mathcal{B}(\mathcal{H})$. Finally, $(\mathcal{P}_t^{\mathrm{qOU}})_{t \geq 0}$ is primitive with respect to ρ_ν (cf. Theorem 6.2 of [Cipriani et al., 2000]). The reason behind the name of the semigroup resides in the following theorem which one can find in [Cipriani et al., 2000]:

Theorem 5.5.1. *For any polynomial p, $p(Q) \in \mathrm{dom}(\mathcal{L}^{\mathrm{qOU}})$ and*

$$\mathcal{L}^{\mathrm{qOU}}(p(Q)) = L^{\mathrm{cOU}}(p)(Q)\,,$$

where L^{cOU} is the generator of the classical Ornstein Uhlenbeck semigroup which is defined, for example on $C^2(\mathbb{R})$, as:

$$L^{\mathrm{cOU}}(f)(x) := \left(\frac{\mu^2 + \lambda^2}{4}\right) f''(x) - \left(\frac{\mu^2 - \lambda^2}{2}\right) x \, f'(x)\,.$$

On the other hand, for any polynomial q,

$$\mathcal{L}^{\mathrm{qOU}}(q(N)) = L^{\mathrm{BD}}(q)(N)\,,$$

where L^{BD} is the generator of a birth and death process with birth rates $\{\lambda^2(k+1)\}_{k \in \mathbb{N}}$ and death rates

$\{\mu^2 k\}_{k \in \mathbb{N}}$. *That is*

$$L^{\mathrm{BD}}(f)(k) = \mu^2 k \left(f(k-1) - f(k) \right) + \lambda^2 (k+1) \left(f(k+1) - f(k) \right),$$

Remark 5.5.2. The last theorem provides an example of how from a single quantum Markov semigroup, one can recover different classical Markov semigroups of very different nature, depending on the invariant commutative subalgebra on which the semigroup acts. In this case, the action of the semigroup on the algebra of functions of the position operator leads to a diffusion semigroup, namely the Ornstein Uhlenbeck semigroup, whereas its action on the algebra generated by the number operator leads to a jump semigroup, namely a birth and death process.

Quantum Brownian motion   In the infinite temperature limit, that is when $\mu = \nu \, (= 1/\sqrt{2})$, the generator of the quantum Ornstein Uhlenbeck semigroup formally converges to the one of the *quantum heat semigroup*: on the subalgebra $\mathcal{A}_0 := \mathrm{span}\{ |\varphi\rangle \langle\psi| \; : \; |\varphi\rangle, |\psi\rangle \in \cap_{j=1}^{n} \left(\mathrm{dom}(P_j^2) \cap \mathrm{dom}(Q_j^2) \right) \}$ of the CCR algebra $\mathcal{B}(\mathbb{L}_2(\mathbb{R}^n))$ over the $2n$-dimensional phase space $\mathcal{Z} = \mathbb{R}^{2n}$, consider the following generator

$$\mathcal{L}^{\mathrm{qheat}}(\cdot) := -\frac{1}{4} \sum_{j=1}^{2n} [R_j, [R_j, \cdot]], \tag{5.44}$$

with R_j being the jth element of the vector $\mathbf{R} = (Q_1, P_1, ..., Q_n, P_n)$ defined through Equation (0.30). By Section 2 of [Holevo, 1996], $\mathcal{L}^{\mathrm{qheat}}$ extends to an unbounded operator on $\mathcal{T}_1(\mathcal{H})$ such that, if an operator X has finite moments of order 2, then $X \in \mathrm{dom}(\mathcal{L}^{\mathrm{qheat}})$; this in particular is true if X is a Gaussian state or a Schwartz operator (cf. Section 0.2). The following theorem lists some important properties of the quantum heat semigroup. Some of these can already be found in [Koenig and Smith, 2014].

Proposition 5.5.3. *The generator defined in Equation* (5.44) *is the weak* generator of a unique unital quantum Markov semigroup* $(\mathcal{P}_t^{\mathrm{qheat}})_{t \geq 0}$ *on* $\mathcal{B}(\mathbb{L}_2(\mathbb{R}^n))$ *called the quantum heat semigroup. Moreover,*

1. *The quantum heat semigroup is reversible with respect to the trace:*

$$\langle A, \mathcal{P}_t^{\mathrm{qheat}}(B) \rangle_{\mathrm{HS}} = \langle \mathcal{P}_t^{\mathrm{qheat}}(A), B \rangle_{\mathrm{HS}}, \quad \forall A, B \in \mathcal{T}_2(\mathcal{H}). \tag{5.45}$$

   Denoting by $\mathcal{L}_2^{\mathrm{qheat}}$ *the* $\mathbb{L}_2(\mathrm{Tr})$ *generator of the above QMS, this implies that, for all* $A, B \in \mathrm{dom}(\mathcal{L}_2^{\mathrm{qheat}})$,

$$\langle A, \mathcal{L}_2^{\mathrm{qheat}}(B) \rangle_{\mathrm{HS}} = \langle \mathcal{L}_2^{\mathrm{qheat}}(A), B \rangle_{\mathrm{HS}}. \tag{5.46}$$

2. *Let the state* $\rho_t \equiv \mathcal{P}_t^{\mathrm{qheat}}(\rho)$ *denote the solution of the quantum diffusion equation defined through Equation* (5.44). *Then for each* $t \geq 0$, *the characteristic function of* ρ_t *is given by*

$$\mathcal{F}_{\rho_t}^{\mathrm{q}}(z) = \mathcal{F}_{\rho}^{\mathrm{q}}(z) \, e^{-|z|^2 t/4}, \quad \forall \, z \in \mathcal{Z}, \tag{5.47}$$

   and we have

$$\rho_t \equiv \rho \ast g_{t/2} := \int_{\mathcal{Z}} W_z \, \rho \, W_{-z} \, g_{t/2}(z) \, dz, \tag{5.48}$$

where $g_{t/2}$ denotes the probability density function (pdf) of a Gaussian random variable on $\mathbb{R}^{2n}$ with zero mean and variance equal to $t/2$, i.e.

$$g_{t/2}(z) = \frac{1}{(\pi t)^n} e^{-|z|^2/t}, \tag{5.49}$$

and W_z denotes the Weyl operator defined in Equation (0.28). In particular, if ρ is a Gaussian state, then so is ρ_t.

Proof. Equation (5.47) can be found in [Koenig and Smith, 2014]. In fact, Equation (5.46) can be directly verified by using Equation (5.44). The proof of Equation (5.48) is obtained as follows: Note that $e^{-|z|^2 t/4}$ is the characteristic function of a Gaussian random variable of associated pdf $g_{t/2}$. It is well-known that if f and g are two pdfs then the characteristic function of their convolution $f * g$ is equal to the product of their characteristic functions:

$$\mathcal{F}_{f*g}(z) = \mathcal{F}_f(z)\,\mathcal{F}_g(z).$$

This property also holds when one of the pdfs is replaced by a quantum state and the standard definition of convolution is replaced by the one of *quantum convolution* (given by Equation (5.48)), as shown by [Werner, 1984] (see also [Kossakowski, 1972]) in his generalization of harmonic analysis to the quantum framework. Hence, (5.47) allows us to express the state ρ_t as a quantum convolution of the initial state ρ and the pdf $g_{t/2}$, as given by Equation (5.48). This provides an easy proof ot self-duality (5.45): the operator $\mathcal{P}_t(B)$ is given by the right hand side of Equation (5.48). Hence,

$$
\begin{aligned}
\operatorname{Tr}\left(A^* \mathcal{P}_t(B)\right) &= \operatorname{Tr}\left(A^* \int_{\mathcal{Z}} W_z\, B\, W_{-z}\, g_{t/2}(z)\, dz\right) \\
&= \operatorname{Tr}\left(A^* \int_{\mathcal{Z}} W_{-z}\, B\, W_z\, g_{t/2}(z)\, dz\right) \\
&= \operatorname{Tr}\left(\int_{\mathcal{Z}} W_z\, A^*\, W_{-z}\, B\, g_{t/2}(z)\, dz\right) \\
&= \operatorname{Tr}\left(\left(\int_{\mathcal{Z}} W_z\, A\, W_{-z}\right)^* B\, g_{t/2}(z)\, dz\right) \\
&= \operatorname{Tr}\left((\mathcal{P}_t(A))^* B\right), \tag{5.50}
\end{aligned}
$$

where we have used the symmetry of the Gaussian pdf, the cyclicity of the trace, the fact that $W_z^* = W_{-z}$, and Equation (5.48).

$\square$

Remark 5.5.4. The semigroup does not possess an invariant state: if there was an invariant state ρ, then its characteristic function would satisfy $\mathcal{F}_{\rho_t}^{\mathrm{q}} = \mathcal{F}_\rho^{\mathrm{q}}$ for all $z \in \mathcal{Z}$, which is impossible by Equation (5.47). The semifinite trace Tr is, however, invariant under $(\mathcal{P}_t^{\mathrm{qheat}})_{t \geq 0}$. This is in complete analogy with the classical heat semigroup $(P_t^{\mathrm{heat}})_{t \geq 0}$ which leaves invariant the Lebesgue measure. Moreover, for any smooth function f on $\mathbb{R}^{2n}$:

$$L^{\mathrm{heat}}(f)(Q) = \mathcal{L}^{\mathrm{qheat}}(f(Q)),$$

where $L^{\mathrm{heat}} \equiv \frac{1}{4}\Delta$ is the generator of the classical heat semigroup.

5.5.3. Quantum convolution semigroups

More generally, [Kossakowski, 1972] introduced a technique to define quantum Markov semigroups on the space $\mathcal{T}_1(\mathcal{H})$, $\mathcal{H}$ separable, starting from a classical one parameter convolution semigroup (we

refer to Section 2.5 for the definition of such semigroups):

Theorem 5.5.5. *Let $(G, \mathcal{B}(G))$ be a locally compact separable group together with its Borel algebra $\mathcal{B}(G)$, and let U be a unitary representation of G on $\mathcal{H}$. Any one-parameter convolution semigroup $(\mu_t)_{t \geq 0}$ on $\mathcal{M}(G)$ induces a strongly continuous semigroup $(\mathcal{P}^{\mu}_{t*})_{t \geq 0}$ on $\mathcal{T}_1(\mathcal{H})$ defined as follows: for any $\rho \in \mathcal{T}_1(\mathcal{H})$:*

$$\mathcal{P}^{\mu}_{t*}(\rho) := \int_G U_g \, \rho \, U_{g^{-1}} \, \mu_t(dg), \qquad \rho \in \mathcal{T}_1(\mathcal{H}).$$

In finite dimensions, a characterization of the Heisenberg dual of the quantum convolution semigroups defined by Kossakowski was given by [Kümmerer and Maassen, 1987] in terms of their dilation properties: We mentioned in Section 2.1 that to any classical Markov semigroup $(P_t)_{t \geq 0}$ one can associated a unique minimal Markov process. In the noncommutative setting, the situation is more complicated and is known as the *dilation problem*. Here, by a quantum Markov process (or Markov dilation), we mean a group of *-automorphisms occurring on a larger system from which the local irreversible evolution is recovered by coarse graining (see [Hudson and Parthasarathy, 1984, Parthasarathy, 1992, Meyer, 1993, Kümmerer and Maassen, 1987, Rajarama Bhat and Parthasarathy, 1995] and references therein). In finite dimensions, [Kümmerer and Maassen, 1987] showed that the existence of a certain type of dilation characterizes those unital semigroups that can be written as a quantum convolution semigroup.

More precisely, let $(\mathcal{P}_t)_{t \geq 0}$ be a QMS defined a von Neumann algebra $\mathcal{M}$, and assume that it has a faithful normal invariant state ω. If there exists a group $(\alpha_t)_{t \in \mathbb{R}}$ of *-automorphisms on a von Neumann $\mathcal{M}'$ together with faithful normal invariant state ω' on $\mathcal{M}'$, and completely positive, unital operators $\pi : \mathcal{M} \to \mathcal{M}'$, $E_{\mathcal{M}} : \mathcal{M}' \to \mathcal{M}$ such that $\omega \circ E_{\mathcal{M}} = \omega'$ and $\omega' \circ \pi = \omega$, and such that the following diagram commutes for all $t \geq 0$:

$$
\begin{array}{ccc}
(\mathcal{M}, \omega) & \xrightarrow{\ \mathcal{P}_t\ } & (\mathcal{M}, \omega) \\
\Big\downarrow{\scriptstyle \pi} & & \Big\uparrow{\scriptstyle E_{\mathcal{M}}} \\
(\mathcal{M}', \omega') & \xrightarrow{\ \alpha_t\ } & (\mathcal{M}', \omega')
\end{array}
$$

then we call the tuplet $(\mathcal{M}', \omega', (\alpha_t)_{t \geq 0})$ a *dilation* of the *dynamical system* $(\mathcal{M}, \omega, (\mathcal{P}_t)_{t \geq 0})$. Such dilation is called *essentially commutative* if the relative commutant of $\pi(\mathcal{M})$ in $\mathcal{M}'$ is commutative.

Following [Kümmerer and Maassen, 1987], we now restrict our attention to the case when $(\mathcal{P}_t)_{t \geq 0}$ is defined on the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on the finite dimensional Hilbert space $\mathcal{H}$, and assume that the state ρ is invariant. In this case, any commutative dilation is characterized by a commutative von Neumann algebra $\mathcal{C}$ with faithful normal state $\mathbb{E}_\mu$ such that $\mathcal{M}' = \mathcal{B}(\mathcal{H}) \otimes \mathcal{C}$ and $\omega' = \omega \otimes \mathbb{E}_\mu$. Moreover, for any $X \in \mathcal{B}(\mathcal{H})$, $\pi(X) = X \otimes \mathbb{1}_\mathcal{C}$ and $E_{\mathcal{B}(\mathcal{H})}$ is uniquely defined by its action on tensor products: $E_{\mathcal{B}(\mathcal{H})}(X \otimes f) = \mathbb{E}_\mu(f) \, X$. Since any automorphism of $\mathcal{B}(\mathcal{H})$ leaves the trace invariant, it follows that one can take $\omega \equiv d_{\mathcal{H}}^{-1} \mathrm{Tr}$, without loss of generality. Then (see Theorem 1.1.1 of [Kümmerer and Maassen, 1987]):

Theorem 5.5.6. *The following are equivalent:*

(i) *The dynamical system $(\mathcal{B}(\mathcal{H}), \frac{\mathbb{1}_{\mathcal{H}}}{d_{\mathcal{H}}}, (\mathcal{P}_t)_{t \geq 0})$ admits an essentially commutative Markov dilation.*

(ii) *There exists a weak*-continuous convolution semigroup $(\mu_t)_{t \geq 0}$ on the space $\mathcal{M}(\mathrm{Aut}(\mathcal{B}(\mathcal{H})))$ of probability measures on the group of automorphisms of $\mathcal{B}(\mathcal{H})$ such that for all $X \in \mathcal{B}(\mathcal{H})$ and $t \geq 0$:*

$$\mathcal{P}_t(X) = \int_{\alpha \in \mathrm{Aut}(\mathcal{B}(\mathcal{H}))} \alpha(X) \, \mu_t(d\alpha)$$

(iii) *The generator $\mathcal{L}$ of $(\mathcal{P}_t)_{t \geq 0}$ is of the form:*

$$\mathcal{L}(X) = i\,[H,X] + \frac{1}{2} \sum_{j=1}^{k} 2\,A_j X A_j - A_j^2 X - X A_j^2 + \sum_{i=1}^{l} c_i\,(U_i^* X U_i - X)\,, \qquad (5.51)$$

where H and A_j, $j \in \{1,...,k\}$ are self-adjoint elements of $\mathcal{B}(\mathcal{H})$, U_i, $i \in \{1,...,l\}$, are unitaries in $\mathcal{B}(\mathcal{H})$ and c_i, $i \in \{i,...,l\}$ are positive numbers.

(iv) *The generator $\mathcal{L}$ of $(\mathcal{P}_t)_{t \geq 0}$ is in the closure of the cone generated by $\{\alpha - \mathrm{id};\ \alpha \in \mathrm{Aut}(\mathcal{B}(\mathcal{H}))\}$*

(v) *For all $t \geq 0$, $\mathcal{P}_t$ lies in the convex hull of $\mathrm{Aut}(\mathcal{B}(\mathcal{H}))$.*

We come back to this decomposition in Chapter 9, where a simpler version of the theorem is presented.

Remark 5.5.7. Not every unital QMS on $\mathcal{B}(\mathcal{H})$ can be seen as a quantum convolution semigroup, as shown in [Kümmerer and Maassen, 1987]. This becomes the case when the QMS satisfies ω-DBC with respect to the completely mixed state (see e.g. [Frigerio and Gorini, 1984]). However, there exist QMS that are not $d_{\mathcal{H}}^{-1}\mathbb{1}_{\mathcal{H}}$-DBC and still admit a Markov dilation [Kümmerer and Maassen, 1987].

As we will see in Chapter 9, there may be several convolution semigroups of probability measures leading to the same QMS $(\mathcal{P}_t)_{t \geq 0}$ on $\mathcal{B}(\mathcal{H})$. However, Theorem 2.5.1 allows us to write a canonical form for the generator of such a semigroup in terms of the form (2.33) of the generator of $(\mu_t)_{t \geq 0}$. This is the content of Theorem 1.5.1 of [Kümmerer and Maassen, 1987]:

Theorem 5.5.8. *Let $X_1,...,X_{n^2-1}$ be a fixed basis of the Lie algebra $\mathfrak{aut}(\mathcal{B}(\mathcal{H}))$ of $\mathrm{Aut}(\mathcal{B}(\mathcal{H}))$ that is orthonormal with respect to $\langle .,.\rangle_{\mathrm{HS}}$. Choose functions $g_1,...,g_{n^2-1}$ on $C_2(\mathrm{Aut}(\mathcal{B}(\mathcal{H})))$ with the properties $g_i(\mathrm{id}) = 0$, $D_i g_j(\mathrm{id}) = \delta_{ij}$ where D_i is the derivation on $C^2(\mathrm{Aut}(\mathcal{B}(\mathcal{H})))$ induced by X_i. Moreover, let $(\mathcal{P}_t)_{t \geq 0}$ be a unital QMS on $\mathcal{B}(\mathcal{H})$. Then there exists a one-to-one correspondence between:*

(i) *essentially commutative minimal Markov dilations of $(\mathcal{B}(\mathcal{H}), d_{\mathcal{H}}^{-1}\mathbb{1}_{\mathcal{H}}, (\mathcal{P}_t)_{t \geq 0})$;*

(ii) *convolution semigroups $(\mu_t)_{t \geq 0}$ of measures on $\mathrm{Aut}(\mathcal{B}(\mathcal{H}))$ such that*

$$\mathcal{P}_t(X) = \int_{\mathrm{Aut}(\mathcal{B}(\mathcal{H}))} \alpha(X)\,\mu_t(d\alpha)\,;$$

(iii) *triples $(H, \{b_{ij}\}_{1 \leq i,j \leq n^2-1}, \nu)$ where H is a self-adjoint element of $\mathcal{B}(\mathcal{H})$, $\{b_{ij}\}$ a real positive semidefinite matrix and ν a Lévy measure such that for all $X \in \mathcal{B}(\mathcal{H})$,*

$$\mathcal{L}(X) = i\,[H,X] + \sum_{ij=1}^{n^2-1} b_{ij}\,[X_i,[X_j,X]] + \int_{\mathrm{Aut}(\mathcal{B}(\mathcal{H}))} \left(\alpha(X) - X + \sum_{i=1}^{n^2-1} g_i(\alpha)\,[X_i,X] \right) \nu(d\alpha)\,.$$

Chapter 6.

Decoherence

Quantum and classical systems exhibit very different behaviors, and it is not clear in practice how the former should be thought of as generalizations of the latter. For example, there is no real parameter governing the smooth transition between systems that behave according to the superposition principle and those who don't. This is in total contrast with the other major extension of classical physics, namely special relativity, whose laws converge to Newtonian kinematics and dynamics when taking the speed of light to infinity. This simple observation lead Bohr to formulate the Copenhagen interpretation of quantum mechanics according to which classical and quantum descriptions of a physical system are complementary to each other. The Copenhagen interpretation is somewhat unsatisfactory for two main reasons: Firstly, it imposes the necessary cohabitation of two very distinct physical theories with apparently little logical relation. Secondly, it does not provide any way to decide when one should use either one or the other theory to describe a physical system.

The program of environment induced decoherence was first proposed as a way to solve this issue. Roughly speaking, it states that the universe is fundamentally quantum, but most macroscopic systems acquire classical properties, and hence effectively behave like classical systems, due to unavoidable interactions with their environment. The measurement-like interactions between a system and its environment lead to a strong entanglement between them. This in turn results in the dynamical destruction of certain initial quantum superpositions between vectors belonging to different subspaces in the state of the system under consideration.

An early contribution along these lines is the paper [Zeh, 1970]. Decoherence was popularized by Zurek's articles [Zurek, 1981, Zurek, 1982] (see also [Joos et al., 2003, Blanchard et al., 2000] and the references therein). Perhaps the simplest example of environment induced decoherence is the one treated by [Joos and Zeh, 1985]. There, they showed that the "non-diagonal" elements (in the position basis) of the reduced density matrix of a particle subject to scattering particles of its environment vanish exponentially fast. Similarly, in our Markovian setting, typical decoherent evolutions will be non-primitive. In Chapter 8, we introduce a new set of functional inequalities adapted to the study of the asymptotic properties of such QMS.

Layout of the chapter: We briefly review some general features of the mathematical theory of decoherence in Section 6.1. Section 6.2 deals with the finite dimensional case. We end the chapter by mentioning some important examples of QMS that display decoherence in Section 6.4.

6.1. General theory

One of the first steps towards a rigorous algebraic approach to the study of decoherence was made by [Blanchard and Olkiewicz, 2003] (the slightly different definition given here comes from [Hellmich, 2011], see also [Olkiewicz, 1999, Olkiewicz, 2000] for former special cases previously studied, and [Rebolledo, 2005] for an alternative definition)[1]: consider a von Neumann algebra $\mathcal{M}$ together with an irreversible (reduced) dynamics $(\mathcal{P}_t)_{t \geq 0}$, that is a family of positive, normal unital maps such that $t \mapsto \mathcal{P}_t(X)$ is weak* continuous for every $X \in \mathcal{M}$[2]. $(\mathcal{P}_t)_{t \geq 0}$ is said to display *decoherence* if the following holds: there exists a von Neumann subalgebra $\mathcal{N}_0 \subseteq \mathcal{M}$ and a weak* continuous group of *-automorphisms $(\alpha_t)_{t \geq 0}$ on $\mathcal{N}_0$ such that $\mathcal{P}_t(\mathcal{N}_0) \subseteq \mathcal{N}_0$, $\mathcal{P}_t|_{\mathcal{N}_0} = \alpha_t$ for any $t \geq 0$, as well as a *-invariant, weak* closed subspace $\mathcal{V}_0 \subset \mathcal{M}$ with $\mathcal{P}_t(\mathcal{V}_0) \subseteq \mathcal{V}_0$ for all $t \geq 0$ such that

$$\mathcal{M} = \mathcal{N}_0 \oplus \mathcal{V}_0 \,, \tag{EID}$$

and such that for all $X \in \mathcal{V}_0$ and any $\omega \in \mathcal{M}_*$,

$$\lim_{t \to \infty} \omega \left(\mathcal{P}_t(X) \right) = 0 \,.$$

The algebra $\mathcal{N}_0$ is chosen as the maximal von Neumann subalgebra on which $(\mathcal{P}_t)_{t \geq 0}$ reduces to a group of *-automorphisms, and is referred to as the *algebra of effective observables*. On the other hand, the subspace $\mathcal{V}_0$ is known as the space of *non-detectable observables*. In other words, decoherence as defined here demands that the non-automorphic part of the quantum evolution vanishes in time. Hence, after a sufficiently long time, the system is effectively characterized by observables belonging to the algebra $\mathcal{N}_0$ and the reversible time evolution is given by the group $(\alpha_t)_{t \geq 0}$. This means that it effectively evolves like a closed system whose properties may differ from the ones of the original system. Note that the above definition makes no assumptions about the rate at which decoherence occurs. In Chapters 8 and 9 we devise some functional analytic techniques to estimate such rates for finite dimensional systems.

The problem of finding minimal conditions for the occurrence of environment induced decoherence is a difficult one. Even when assuming that decoherence occurs, uniqueness of the EID decomposition is not known in general. Here, we briefly review some of the mains results in the general von Neumann algebraic setting. In the next section, we will specialize to the case of QMS defined on the algebra $\mathcal{B}(\mathcal{H})$ of linear operators on a finite dimensional Hilbert space.

First of all, we introduce the *decoherence free algebra* $\mathcal{N}_0 \subseteq \mathcal{N}$ and the space $\mathcal{V}_0 \subseteq \mathcal{V}$ of weakly vanishing operators in $\mathcal{M}$:

$$\mathcal{N} := \left\{ X \in \mathcal{M} \,|\, \forall t \geq 0, \ \mathcal{P}_t(X^*X) = \mathcal{P}_t(X)^* \mathcal{P}_t(X), \mathcal{P}_t(XX^*) = \mathcal{P}_t(X)\mathcal{P}_t(X)^* \right\} \,,$$
$$\mathcal{V} := \left\{ X \in \mathcal{M}; \ \text{weak}^* \lim \mathcal{P}_t(X) = 0 \text{ as } t \to \infty \right\} \,.$$

$\mathcal{N}$ and $\mathcal{V}$ typically constitute natural candidates for $\mathcal{N}_0$ and $\mathcal{V}_0$ when EID occurs. The inclusion $\mathcal{N}_0 \subseteq \mathcal{N}$ follows from the following result, the proof of which can be found in Proposition 1 of [Carbone et al., 2011]:

Proposition 6.1.1. $\mathcal{N}$ *is the largest von Neumann subalgebra of* $\mathcal{M}$ *for which* $\mathcal{P}_t|_{\mathcal{N}}$ *acts as a*

---

[1]A similar definition of decoherence in the C^*-algebraic setting was also considered in [Blanchard et al., 2007].
[2]Remark that we do not yet require $(\mathcal{P}_t)_{t \geq 0}$ to form a QMS at this stage.

$*$-homomorphism. In particular, for all $X, Y \in \mathcal{M}$, if either X or Y belongs to $\mathcal{N}$, then for all $t \geq 0$:

$$\mathcal{P}_t(X^*Y) = \mathcal{P}_t(X^*)\mathcal{P}_t(Y). \tag{6.1}$$

Proof. Obviously, $\mathcal{N}^* = \mathcal{N}$. Moreover, since $\mathcal{P}_t$ is completely positive unital for any $t \geq 0$, it satisfies the Schwarz inequality (1.55):

$$\mathcal{P}_t(X^*X) \geq \mathcal{P}_t(X^*)\mathcal{P}_t(X). \tag{6.2}$$

Therefore the map $D_t : (X, Y) \mapsto \mathcal{P}_t(X^*Y) - \mathcal{P}_t(X^*)\mathcal{P}_t(Y)$ is positive and sesquilinear, so that $D_t(X, X) = 0$ is equivalent to $D_t(X, Y) = 0$ for all Y. Now, since $D_t(X, X) = 0$ for all $X \in \mathcal{N}$, Equation (6.1) follows. Next, we prove that $\mathcal{N}$ is an algebra. Linearity is obvious. Now, let $X, Y \in \mathcal{N}$:

$$\begin{aligned}
\mathcal{P}_t((XY)^*(XY)) &= \mathcal{P}_t(Y^*X^*XY) \\
&= \mathcal{P}_t(Y^*)\mathcal{P}_t(X^*XY) \\
&= \mathcal{P}_t(Y^*)\mathcal{P}_t(X^*)\mathcal{P}_t(XY) \\
&= \mathcal{P}_t(Y^*X^*)\mathcal{P}_t(XY) \\
&= \mathcal{P}_t((XY)^*)\mathcal{P}_t(XY),
\end{aligned}$$

and $\mathcal{P}_t((XY)(XY)^*) = \mathcal{P}_t(XY)\mathcal{P}_t((XY)^*)$ follows similarly. Therefore, $XY \in \mathcal{N}$. $\mathcal{P}_t$ invariance follows from

$$\mathcal{P}_t(\mathcal{P}_s(X)^*\mathcal{P}_s(X)) = \mathcal{P}_t(\mathcal{P}_s(X^*X)) = \mathcal{P}_{t+s}(X^*X) = \mathcal{P}_t(\mathcal{P}_s(X^*))\mathcal{P}_t(\mathcal{P}_s(X)),$$

for any $X \in \mathcal{N}$, and $s, t \geq 0$. Finally, the weak* closure comes from the following observation: define for any $Y \in \mathcal{M}$:

$$\varphi_Y : \mathcal{M} \ni X \mapsto \mathcal{P}_t(Y^*X) - \mathcal{P}_t(Y^*)\mathcal{P}_t(X) \in \mathcal{M}, \tag{6.3}$$

Then $\mathcal{N} = \cap_{Y \in \mathcal{M}} \varphi_Y^{-1}(\{0\})$. The weak* closure follows from the fact that $\mathcal{P}_t$ is normal and $X \mapsto YX$ is weak* continuous. $\qquad\square$

It is further known that $\mathcal{N} = \mathcal{N}_0$ (i.e. $\mathcal{N}$ is the largest von Neumann subalgebra of $\mathcal{M}$ on which the maps $\mathcal{P}_t$ act as $*$-automorphisms) if any of the following holds [Robinson, 1982] (see also Proposition 3 of [Carbone et al., 2015]):

- $(\mathcal{P}_t)_{t \geq 0}$ possesses a faithful normal invariant state;

- $(\mathcal{P}_t)_{t \geq 0}$ is uniformly continuous on $\mathcal{M} := \mathcal{B}(\mathcal{H})$, $\mathcal{H}$ separable;

- $\mathcal{N} \subseteq \mathcal{F}$, where $\mathcal{F}$ is the set of fixed points of $(\mathcal{P}_t)_{t \geq 0}$:[3]

$$\mathcal{F} := \{X \in \mathcal{M}; \ \mathcal{P}_t(X) = X \ \forall t \geq 0\}.$$

In the uniformly continuous case, this trivially holds since the generator $\mathcal{L}$ is bounded, so that $\mathcal{P}_t$ is invertible for any $t \geq 0$, of inverse $\mathcal{P}_{-t} := \mathrm{e}^{-t\mathcal{L}}$. Moreover, $\mathcal{N} \cap \mathcal{V} = \{0\}$, so that if EID holds one necessarily has $\mathcal{N} = \mathcal{N}_0$ and $\mathcal{V} = \mathcal{V}_0$ (see Proposition 2 of [Carbone et al., 2013], Proposition 6

---

[3]In fact, the link between $\mathcal{N}$ and $\mathcal{F}$ was already investigated in the 70s and used to study ergodic properties of quantum Markov semigroups (see [Evans, 1977, Frigerio, 1978, Frigerio and Verri, 1982, Robinson, 1982, Fagnola and Rebolledo, 2008]).

of [Carbone et al., 2015]). This fact is no longer true in general (see Example 9 of [Carbone et al., 2015]).

One natural condition that one can require is the existence of a faithful normal invariant state $\omega \in \mathcal{M}_*$.[4] In this case, the set $\mathcal{F}$ of fixed points is a von Neumann subalgebra (see [Spohn, 1977, Frigerio, 1978], Theorem 6.12 of [Wolf, 2012]). In fact, the following was proved in [Frigerio and Verri, 1982]:

Theorem 6.1.2. *Given a QMS $(\mathcal{P}_t)_{t\geq 0}$ with faithful normal invariant state ω, the limit*

$$E_{\mathcal{F}}[X] := \text{weak}^* \lim_{t\to 0} \frac{1}{t} \int_0^t \mathcal{P}_s(X) ds$$

exists for all $X \in \mathcal{B}(\mathcal{H})$ and defines a $(\mathcal{P}_t)_{t\geq 0}$ invariant normal conditional expectation $E_{\mathcal{F}}$ onto the algebra $\mathcal{F}$ of fixed points of $(\mathcal{P}_t)_{t\geq 0}$.

Under the condition of existence of a faithful normal invariant state ω, EID was moreover proved to hold in the following cases:

- $\mathcal{N} \subseteq \mathcal{F}$ (and in fact $\mathcal{N} = \mathcal{F}$);

- $\mathcal{M} = \mathcal{B}(\mathcal{H})$, $\mathcal{H}$ finite dimensional;

- $\mathcal{M} = \mathcal{B}(\mathcal{H})$, $\mathcal{H}$ separable, $\mathcal{N}$ atomic, $(\mathcal{P}_t)_{t\geq 0}$ uniformly continuous.

with $\mathcal{V}_0$ being the orthogonal complement $\mathcal{N}^\perp$ of $\mathcal{N}$ under the GNS inner product $\langle A, B \rangle_{1,\omega} := \omega(A^* B)$. Moreover, in these cases, there exists a conditional expectation $E_{\mathcal{N}}$ onto $\mathcal{N}$, compatible with ω, with $\ker(E_{\mathcal{N}}) = \mathcal{N}^\perp$ (see Theorem 1.17 of [Carbone et al., 2014], Theorem 9 of [Carbone et al., 2013], Theorem 22 of [Deschamps et al., 2016] respectively). In general, $\mathcal{M}$ might not be complete with respect to $\langle .,. \rangle_{1,\omega}$, so that $\mathcal{N}$ might not be closed with respect to the corresponding norm and one cannot define $E_{\mathcal{N}}$ as the projection onto $\mathcal{N}$. In this case, [Carbone et al., 2015] considered the completion $\overline{\mathcal{M}}^{\langle .,. \rangle_\omega}$ together with the splitting

$$\overline{\mathcal{M}}^{\langle .,. \rangle_\omega} = \overline{\mathcal{N}}^{\langle .,. \rangle_\omega} \oplus (\overline{\mathcal{N}}^{\langle .,. \rangle_\omega})^\perp .$$

In general, it is not true that $\mathcal{M} = \mathcal{N} \oplus (\overline{\mathcal{N}}^{\langle .,. \rangle_\omega \perp} \cap \mathcal{M})$. The situation simplifies if one further assumes the existence of a conditional expectation $E_{\mathcal{N}}$ onto $\mathcal{N}$ (cf. Theorem 19 of [Carbone et al., 2015]). A proof of it in the case when $\mathcal{M} = \mathcal{B}(\mathcal{H})$, $\mathcal{H}$ finite, is provided at the end of Section 6.2 for sake of completeness.

Theorem 6.1.3. *Assume there exists a faithful normal invariant state ω and a conditional expectation $E_{\mathcal{N}}$ compatible with ω. Then EID occurs with $\mathcal{N}_0 = \mathcal{N}$ and $\mathcal{V}_0 = \ker E_{\mathcal{N}} = \overline{\mathcal{N}}^{\langle .,. \rangle_\omega \perp} \cap \mathcal{M}$ being $\mathcal{P}_t$-invariant for all $t \geq 0$. Moreover, $\mathcal{P}_t$ commutes with $E_{\mathcal{N}}$.*

Remark 6.1.4. The sufficient condition given in the above theorem is equivalent to the invariance of $\mathcal{N}$ under the modular automorphism group $(\sigma_s^\omega)_{s\in\mathbb{R}}$ (cf. Theorem 0.1.4). This in particular happens when each $\mathcal{P}_t$ commutes with $(\sigma_s^\omega)_{s\in\mathbb{R}}$ (see Remark 20 of [Carbone et al., 2015]). The former fact always holds in finite dimensions, as we see in the next section.

Other splittings of the algebra $\mathcal{M}$ into a unitarily evolving part and a decaying part have been extensively studied and compared to EID (see [Carbone et al., 2015, Hellmich, 2009] and references therein for more details).

---

[4]This assumption was actually part of the original definition of EID by [Blanchard and Olkiewicz, 2003]. It is however not necessary for EID to hold (see Proposition 2.5 of [Carbone et al., 2014]).

Primitive quantum Markov semigroups: In the case when $\mathcal{N} = \mathcal{F}$, the semigroup is said to be *ergodic*. In this case, one can use Frigerio and Verri's sufficient condition for the convergence of a QMS towards a steady state[5]: under the condition of existence of a faithful normal invariant state, the following convergence holds:

$$E_{\mathcal{F}}[X] = \text{weak}^* \lim_{t \to \infty} \mathcal{P}_t(X).$$

where $E_{\mathcal{F}}$ is the conditional expectation of Theorem 6.1.2.

In the case when $\mathcal{N} = \mathcal{F} = \mathbb{C}\,\mathbb{1}$, the conditional expectation simply reduces to the average with respect to the unique invariant state ω:

$$\text{weak}^* \lim_{t \to \infty} \mathcal{P}_t(X) = \omega(X)\,\mathbb{1} \quad \Leftrightarrow \quad \text{strong} \lim_{t \to \infty} \mathcal{P}_{t*} = \omega.$$

In this case, the semigroup is said to be *primitive*. A sufficient condition for primitivity is then the existence of a unique faithful invariant state (see also [Burgarth et al., 2013, Theorem 14]).

6.2. The case $\mathcal{M} = \mathcal{B}(\mathcal{H})$, $\mathcal{H}$ finite

In this section, we restrict our attention to the case of norm continuous QMS defined on the algebra $\mathcal{B}(\mathcal{H})$ of linear operators acting on a finite dimensional Hilbert space $\mathcal{H}$ (see [Carbone et al., 2011, Deschamps et al., 2016, Carbone et al., 2013]). The following result can be found in [Fagnola and Rebolledo, 2008] (for an extension to minimal QMS whose generator is represented by a generalised GKLS form, see [Dhahri et al., 2010]):

Theorem 6.2.1. *Given a QMS $(\mathcal{P}_t)_{t \geq 0}$ with GKLS form given by Equation (5.5), the decoherence free algebra $\mathcal{N}$ is characterized by*

$$\{\nabla_H^k(L_j), \nabla_H^k(L_j^*);\ k \in \mathbb{N}, j \in I\}',$$

where $\nabla_H(L) \coloneqq [H, L]$. This is independent of the choice of Lindblad operators.

As already discussed in Section 0.1.4, a basic result from the theory of $*$-algebras on finite dimensional Hilbert spaces states that $\mathcal{N}$ can always be decomposed into a direct sum of subparts where it restricts to a factor [Kadison and Ringrose, 2015]. More precisely, up to a unitary transformation, the Hilbert space $\mathcal{H}$ admits the following decomposition

$$\mathcal{H} = \bigoplus_{i \in I} \mathcal{H}_i \otimes \mathcal{K}_i, \tag{6.4}$$

such that $\mathcal{N}$ is unitarily isomorphic to a matrix algebra: there exists a unitary $U \in \mathcal{U}(\mathcal{H})$ such that[6]

$$\mathcal{N} = U \bigoplus_{i \in I} \mathcal{B}(\mathcal{H}_i) \otimes \mathbb{1}_{\mathcal{K}_i} U^* \equiv U \sum_{i \in I} P_i \mathcal{N} P_i U^*, \tag{6.5}$$

where the projections P_i are mutually orthogonal projections in $\mathcal{N}$ onto $\bigoplus_{j=1}^{\dim(\mathcal{K}_i)} \mathcal{H}_i$, $\sum_{i \in I} P_i = \mathbb{1}_{\mathcal{H}}$, and clearly commute with any elements of $\mathcal{N}$. They are minimal in the sense that for any other

---

[5]This condition turns out to be necessary in the finite dimensional setting, see [Frigerio and Verri, 1982].

[6]From now on, we forget the rotation induced by the unitary operator U for sake of simplicity, this can be done after redefining the semigroup as $(U \mathcal{P}_t(U^*(.)U)U^*)_{t \geq 0}$.

projection $P \in \mathcal{N}$,

$$P \leq P_i \quad \Rightarrow \quad P = 0 \text{ or } P = P_i. \tag{6.6}$$

In the primitive case, we recall that $\mathcal{N} = \mathbb{C} \mathbb{1}_{\mathcal{H}}$ so that there exists a unique minimal projection which is equal to $\mathbb{1}_{\mathcal{H}}$. More generally, for each $i \in I$, P_i is in the centralizer $\mathcal{Z}(\mathcal{N})$ of $\mathcal{N}$[7].

Since we are in finite dimensions, we know from last section that EID occurs under the condition of existence of a full-rank invariant state σ, and that, in this case, $\mathcal{N}_0 = \mathcal{N}$ and $\mathcal{V}_0 = \mathcal{V}$. Moreover, there exists a conditional expectation $E_{\mathcal{N}} : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ (see end of this section). We call such semigroups *faithful*. In order get an expression for $E_{\mathcal{N}}$, we first need to study the structure of invariant states of $(\mathcal{P}_t)_{t \geq 0}$. This is taken from Theorem 21 of [Deschamps et al., 2016]:

Theorem 6.2.2. *Assume that $(\mathcal{P}_t)_{t \geq 0}$ admits a full-rank invariant state σ. Then, there exists a family $\{\tau_i\}_{i \in I}$ of full-rank invariant states on each $\mathcal{K}_i$ such that any other $(\mathcal{P}_t)_{t \geq 0}$-invariant state ρ can be written as*

$$\rho = \sum_{i \in I} \operatorname{Tr}_{\mathcal{K}_i} (P_i \rho P_i) \otimes \tau_i. \tag{6.7}$$

The proof consists in a reduction to the evolution in each factor $P_i \mathcal{N} P_i = \mathcal{B}(\mathcal{H}_i) \otimes \mathbb{1}_{\mathcal{K}_i}$. We first state the following technical lemma that is also of independent interest:

Lemma 6.2.3. *The projections P_i belong to the set of fixed points $\mathcal{F}$. Moreover, for any $i, j \in I$, any $X \in \mathcal{B}(\mathcal{H})$ and any $t \geq 0$:*

$$\mathcal{P}_t(P_j X P_i) = P_j \mathcal{P}_t(X) P_i.$$

In particular, each factor $P_i \mathcal{N} P_i = \mathcal{B}(\mathcal{H}_i) \otimes \mathbb{1}_{\mathcal{K}_i}$ of $\mathcal{N}$ is invariant under the action of $(\mathcal{P}_t)_{t \geq 0}$.

Proof. The proof proceeds in two steps. First, we show that any projection P_i is contained in the space $\mathcal{F}$ of fixed points of $(\mathcal{P}_t)_{t \geq 0}$. This is due to the fact that $(\mathcal{P}_t)_{t \geq 0}$ acts as a $*$-automorphism on $\mathcal{N}$ (cf. Section 6.1). Therefore, since each $P_i \in \mathcal{Z}(\mathcal{N})$, and since for any $X \in \mathcal{N}$, there exists $Y \in \mathcal{N}$ such that $X = \mathcal{P}_t(Y)$:

$$X \mathcal{P}_t(P_i) = \mathcal{P}_t(Y) \mathcal{P}_t(P_i) = \mathcal{P}_t(Y P_i) = \mathcal{P}_t(P_i Y) = \mathcal{P}_t(P_i) \mathcal{P}_t(Y) = \mathcal{P}_t(P_i) X,$$

i.e. $\mathcal{P}_t(P_i) \in \mathcal{Z}(\mathcal{N})$. Hence, $\{\mathcal{P}_t(P_i)\}_{i \in I}$ is a family of mutually orthogonal projections in $\mathcal{Z}(\mathcal{N})$. Similarly, for any $i, j \in I$, it is easy to show that $P_j \mathcal{P}_t(P_i) P_j$ is a projection in $\mathcal{Z}(\mathcal{N})$, that is $(P_j \mathcal{P}_t(P_i) P_j)^2 = P_j \mathcal{P}_t(P_i) P_j$. Moreover, since $P_j \mathcal{P}_t(P_i) P_j \leq P_j$, $P_j \mathcal{P}_t(P_i) P_j = 0$ or $P_j \mathcal{P}_t(P_i) P_j = P_j$ (cf. (6.6)). By continuity of $t \to \mathcal{P}_t(X)$ for all $X \in \mathcal{B}(\mathcal{H})$, we actually have that $P_j \mathcal{P}_t(P_i) P_j = 0$ for $i \neq j$ and $P_i \mathcal{P}_t(P_i) P_i = P_i$. Therefore

$$P_i \mathcal{P}_t(\mathbb{1} - P_i) P_i = \sum_{j \neq i} P_i \mathcal{P}_t(P_j) P_i = 0 \quad \Rightarrow \quad \mathcal{P}_t(P_i) = P_i.$$

We proved that for any $i \in I$, $P_i \in \mathcal{F}$. This implies that, for all $t \geq 0$ and any $X \in \mathcal{N}$:

$$\mathcal{P}_t(P_i X P_j) = \mathcal{P}_t(P_i) \mathcal{P}_t(X) \mathcal{P}_t(P_j) = P_i \mathcal{P}_t(X) P_j,$$

---

[7]We recall that the centralizer of a von Neumann algebra $\mathcal{M}$ is the von Neumann algebra that consists in the elements X of $\mathcal{M}$ that commute with all $Y \in \mathcal{M}$.

where the first identity above comes from the fact that, under the condition of existence of an invariant full-rank state, $\mathcal{F}$ is an algebra (and hence a subalgebra of $\mathcal{N}$), so that Equation (6.1) holds.[8] $\qquad\square$

The next lemma is key in the reduction of the problem to the case of a factor algebra $\mathcal{N}$, i.e. when $|I| = 1$. It can be found in Proposition 14 of [Deschamps et al., 2016].

Lemma 6.2.4. *Under the assumption of existence of a full-rank invariant state, $P_i \rho P_j = 0$ for any $i \ne j$ and any invariant state ρ.*

Proof. We showed in Lemma 6.2.3 that for any $i, j \in I$ and any $t \ge 0$, $\mathcal{P}_t(P_j X P_i) = P_j \mathcal{P}_t(X) P_i$. By duality, this implies that for any trace-class operator T, $\mathcal{P}_{t*}(P_j T P_i) = P_j \mathcal{P}_{t*}(T) P_i$. Hence, for $T \equiv \rho$ being an invariant state, for any $s \ge 0$:

$$\mathrm{Tr}(P_i \, \rho \, P_j \, X) = \mathrm{Tr}(P_i \mathcal{P}_{s*}(\rho) \, P_j \, X) = \mathrm{Tr}(\rho \, \mathcal{P}_s(P_j \, X \, P_i)) = \mathrm{Tr}(\rho \, P_j \, \mathcal{P}_s(X) \, P_i) \,.$$

Integrating the above from 0 to $t \ge 0$, we get

$$\mathrm{Tr}\,(P_i \, \rho \, P_j \, X) = \mathrm{Tr}\left(\rho \, P_j \left(t^{-1} \int_0^t \mathcal{P}_s(X) ds\right) P_i\right) \,.$$

Then, it follows from Theorem 6.1.2 that, after taking the limit $t \to 0$ in the above right hand side:

$$\mathrm{Tr}(P_i \, \rho \, P_j \, X) = \mathrm{Tr}(\rho \, P_j \, E_{\mathcal{F}}[X] \, P_i) \,.$$

Since in this case $\mathcal{F} \subseteq \mathcal{N}$, and since $P_i, P_j \in \mathcal{Z}(\mathcal{N})$, $P_j E_{\mathcal{F}}[X] P_i = P_j P_i E_{\mathcal{F}}[X] = 0$, which implies that, for any $i \ne j$:

$$\mathrm{Tr}(P_i \rho P_j X) = 0 \,.$$

Since this is true for any $X \in \mathcal{B}(\mathcal{H})$, the result follows. $\qquad\square$

From Lemma 6.2.4, it is enough to prove that, for any QMS $(\mathcal{P}_t)_{t \ge 0}$ on $\mathcal{B}(\mathcal{H}_0 \otimes \mathcal{K}_0)$ of associated decoherence-free algebra $\mathcal{N} = \mathcal{B}(\mathcal{H}_0) \otimes \mathbb{1}_{\mathcal{K}_0}$ possessing a full-rank invariant state σ, any other invariant state ρ can be written as

$$\rho = \mathrm{Tr}_{\mathcal{K}_0}(\rho) \otimes \tau, \tag{6.8}$$

for some given full-rank state τ on $\mathcal{K}_0$. This is done by means of the following structure theorem (see Theorem 11 of [Deschamps et al., 2016]):

Theorem 6.2.5. *In the case when $\mathcal{N}(\mathcal{P}) = \mathcal{B}(\mathcal{H}_0) \otimes \mathbb{1}_{\mathcal{K}_0}$, the semigroup $(\mathcal{P}_t)_{t \ge 0}$ can be decomposed as follows:*

$$(\mathcal{P}_t)_{t \ge 0} = (\mathcal{P}_t^{\mathcal{H}_0} \otimes \mathcal{P}_t^{\mathcal{K}_0})_{t \ge 0} \,,$$

where $(\mathcal{P}_t^{\mathcal{H}_0})_{t \ge 0}$ is a semigroup of automorphisms of associated Hamiltonian H_W, and $(\mathcal{P}_t^{\mathcal{K}_0})_{t \ge 0}$ is primitive.

Proof. By Theorem 6.2.1, we know that L_k, L_k^* have to commute with the algebra $\mathcal{N} = \mathcal{B}(\mathcal{H}_0) \otimes \mathbb{1}_{\mathcal{K}_0}$. This implies the existence of operators M_k on $\mathcal{K}_0$, such that $L_k = \mathbb{1}_{\mathcal{H}_0} \otimes M_k$. Next, given $X \equiv X_0 \otimes \mathbb{1}_{\mathcal{K}_0} \in$

---

[8]In fact, the identity holds even without the assumption that $\mathcal{F}$ forms a $*$-algebra by showing that any projection belonging to $\mathcal{F}$ commutes with the Lindblad operators of any GKLS form of the generator $\mathcal{L}$, so that $\mathcal{L}^n(P_i X P_j) = P_i \mathcal{L}^n(X) P_j$ for any $n \ge 0$, see Lemma 7 of [Deschamps et al., 2016].

$\mathcal{N}$, since $\mathcal{P}_t(X) = e^{itH}(X_0 \otimes \mathbb{1}_{\mathcal{K}_0}) e^{-itH} \in \mathcal{N}$, it must be of the form $(W_t \otimes \mathbb{1}_{\mathcal{K}_0})(X_0 \otimes \mathbb{1}_{\mathcal{K}_0})(W_t \otimes \mathbb{1}_{\mathcal{K}_0})^*$ for some one parameter group $(W_t)_{t \geq 0}$ of unitary operators on $\mathcal{H}_0$. Therefore,

$$(W_t \otimes \mathbb{1}_{\mathcal{K}_0})^* \, e^{itH}(X_0 \otimes \mathbb{1}_{\mathcal{K}_0}) = (X_0 \otimes \mathbb{1}_{\mathcal{K}_0})(W_t \otimes \mathbb{1}_{\mathcal{K}_0})^* \, e^{itH} \,.$$

From the above identity, $(W_t \otimes \mathbb{1}_{\mathcal{K}_0})^* \, e^{itH}$ commutes with $\mathcal{B}(\mathcal{H}_0) \otimes \mathbb{1}_{\mathcal{K}_0}$, and therefore needs to be of the form $\mathbb{1}_{\mathcal{H}_0} \otimes R_t$ for each $t \geq 0$. Moreover, since $(e^{itH})_{t \geq 0} = (W_t \otimes R_t)_{t \geq 0}$ is a group of unitaries, so must be $(R_t)_{t \geq 0}$. Differentiating, we end up with $H = H_W \otimes \mathbb{1}_{\mathcal{K}_0} + \mathbb{1}_{\mathcal{H}_0} \otimes H_R$, where H_W, resp. H_R, denotes the generator of $(W_t)_{t \geq 0}$, resp. $(R_t)_{t \geq 0}$. Therefore, we showed that

$$\mathcal{L} = \mathcal{L}_{\mathcal{H}_0} \otimes \mathrm{id}_{\mathcal{K}_0} + \mathrm{id}_{\mathcal{H}_0} \otimes \mathcal{L}_{\mathcal{K}_0} \,,$$

where for any $X \in \mathcal{B}(\mathcal{H}_0)$, $Y \in \mathcal{B}(\mathcal{K}_0)$:

$$\mathcal{L}_{\mathcal{H}_0}(X) := i\,[H_W, X]\,,$$
$$\mathcal{L}_{\mathcal{K}_0}(Y) := i\,[H_R, Y] - \frac{1}{2} \sum_k M_k^* M_k Y + Y M_k^* M_k - 2 M_k^* Y M_k \,.$$

Hence, $(\mathcal{P}_t)_{t \geq 0} = (\mathcal{P}_t^{\mathcal{H}_0} \otimes \mathcal{P}_t^{\mathcal{K}_0})_{t \geq 0}$, where $(\mathcal{P}_t^{\mathcal{H}_0})_{t \geq 0}$, resp. $(\mathcal{P}_t^{\mathcal{K}_0})_{t \geq 0}$, is the semigroup corresponding to $\mathcal{L}_{\mathcal{H}_0}$, resp. $\mathcal{L}_{\mathcal{K}_0}$. Since for each $t \geq 0$, $\mathcal{P}_t^{\mathcal{H}_0}$ is an automorphism on $\mathcal{B}(\mathcal{H}_0)$, its decoherence free algebra $\mathcal{N}(\mathcal{P}^{\mathcal{H}_0})$ is $\mathcal{B}(\mathcal{H}_0)$. On the other hand, the decoherence free algebra $\mathcal{N}(\mathcal{P}^{\mathcal{K}_0})$ is equal to the commutant of $\{\nabla_{H_R}^k(M_j), \nabla_{H_R}^k(M_j^*);\ k \in \mathbb{N}, j \in I\}$ by Theorem 6.2.1. This in particular implies that for any $X \in \mathcal{N}(\mathcal{P}^{\mathcal{K}_0})$, $\mathbb{1}_{\mathcal{H}_0} \otimes X$ belongs to the commutant of $\{\nabla_H^k(L_j), \nabla_H^k(L_j^*);\ k \in \mathbb{N}, j \in I\}$, that is $\mathcal{N}(\mathcal{P}) = \mathcal{B}(\mathcal{H}) \otimes \mathbb{1}_{\mathcal{K}_0}$. Hence, $\mathcal{N}(\mathcal{P}^{\mathcal{K}_0}) = \mathbb{C}\,\mathbb{1}_{\mathcal{K}_0}$ and $(\mathcal{P}_t^{\mathcal{K}_0})_{t \geq 0}$ is primitive $\qquad\square$

Proof of Theorem 6.2.2 : It remains to prove Equation (6.8). The proof, that can be found in Theorem 20 of [Deschamps et al., 2016], is restated here for sake of completeness. Let $\{e_j\}_{j \geq 1}$ be a basis of eigenvectors of the Hamiltonian H_W (cf. Theorem 6.2.5) so that $H_W e_j = \kappa_j\, e_j$ for some eigenvalues $\kappa_j \in \mathbb{R}$. Now, let ρ be an invariant state that we write as follows,

$$\rho = \sum_{jk} |e_j\rangle\langle e_k| \otimes \rho_{jk}$$

where $\rho_{jk} = \mathrm{Tr}_{\mathcal{H}_0}((|e_j\rangle\langle e_k| \otimes \mathbb{1}_{\mathcal{K}_0})\rho)$. Then, since ρ is invariant:

$$\rho = \mathcal{P}_{t*}(\rho) = \sum_{kj} e^{i(\kappa_k - \kappa_j)t} |e_j\rangle\langle e_k| \otimes \mathcal{P}_{t*}^{\mathcal{K}_0}(\rho_{jk}) \,.$$

Since the operators $|e_j\rangle\langle e_k|$ are linearly independent, $e^{i(\kappa_j - \kappa_k)t} \rho_{jk} = \mathcal{P}_{t*}^{\mathcal{K}_0}(\rho_{jk})$ for all j, k. Since $(\mathcal{P}_t^{\mathcal{K}_0})_{t \geq 0}$ is primitive, with unique full-rank invariant state τ, $\mathcal{P}_{t*}^{\mathcal{K}_0}(\rho_{jk}) \to \mathrm{Tr}(\rho_{jk})\tau$. Therefore, if $\kappa_j \neq \kappa_k$, we have $\mathrm{Tr}(\rho_{jk}) = 0$, while if $\kappa_j = \kappa_k$, $\mathrm{Tr}(\rho_{jk})\tau = \rho_{jk}$. It follows that

$$\rho = \sum_{jk}(\mathrm{Tr}(\rho_{jk})|e_j\rangle\langle e_k|) \otimes \tau = \mathrm{Tr}_{\mathcal{K}_0}(\rho) \otimes \tau \,.$$

$$\square$$

Since $\mathrm{Tr}(\rho_{jk}) = 0$ if $\kappa_j \neq \kappa_k$, one can easily derive the following corollary:

Corollary 6.2.6. *The Hamiltonian H_W associated to the asymptotic evolution $(\mathcal{P}_t|_{\mathcal{N}})_{t \geq 0}$ admits the decomposition $U H_W U^* = \sum_i H_i \otimes \mathbb{1}_{\mathcal{K}_i}$, where for each $i \in I$, $H_i \in \mathcal{B}_{sa}(\mathcal{H}_i)$. Moreover given any invariant state ρ of decomposition $U \rho U^* = \sum_i \mathrm{Tr}_{\mathcal{K}_i}(P_i \rho P_i) \otimes \tau_i$, $[H_i, \mathrm{Tr}_{\mathcal{K}_i}(P_i \rho P_i)] = 0$.*

We are now ready to prove that EID occurs under the unique assumption of existence of a full-rank invariant state σ. To do so, we construct a conditional expectation $E_{\mathcal{N}}$ onto $\mathcal{N}$ as the orthogonal projection onto $\mathcal{N}$ under the GNS inner product associated to the (and in fact, any) invariant state σ. Theorem 6.1.3, for which we provide a proof in our special finite dimensional case below, allows us to conclude.

Construction and properties of the conditional expectation: Define $E_{\mathcal{N}} : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ as follows:

$$E_{\mathcal{N}}[X] := \sum_{i \in I} \mathrm{Tr}_{\mathcal{K}_i}((\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i) P_i X P_i) \otimes \mathbb{1}_{\mathcal{K}_i}, \tag{6.9}$$

One can simply verify from the definition of $E_{\mathcal{N}}$, Lemma 6.2.3 and Theorem 6.2.5 that $E_{\mathcal{N}}$ commutes with $\mathcal{P}_t$ for any $t \geq 0$. Moreover, by simple duality, we also get the following expression for $E_{\mathcal{N}*}$:

$$\rho_{\mathcal{N}} \equiv E_{\mathcal{N}*}(\rho) = \sum_{i \in I} \mathrm{Tr}_{\mathcal{K}_i}(P_i \rho P_i) \otimes \tau_i. \tag{6.10}$$

We show that this superoperator is a valid conditional expectation as well as the orthogonal projection onto $\mathcal{N}$ with respect to $\langle.,.\rangle_{1,\rho}$. Indeed,

- For any invariant state ρ, $E_{\mathcal{N}*}(\rho) = \rho$ (simply compare Equation (6.10) with Equation (6.7));

- For any $X \in \mathcal{N}$, $E_{\mathcal{N}}[X] = X$;

- For any $X \in \mathcal{B}(\mathcal{H})$,

$$\begin{aligned}
\|E_{\mathcal{N}}[X]\|_\infty &= \sum_{i \in I} \| \mathrm{Tr}_{\mathcal{K}_i}((\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i) P_i X P_i) \otimes \mathbb{1}_{\mathcal{K}_i} \|_\infty \\
&= \sum_{i \in I} \| \mathrm{Tr}_{\mathcal{K}_i}((\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i) P_i X P_i) \|_\infty \\
&\leq \sum_{i \in I} \| P_i X P_i \|_\infty \\
&= \|X\|_\infty.
\end{aligned}$$

By Theorem 0.1.4, $E_{\mathcal{N}}$ is the unique conditional expectation with respect to any full-rank invariant state ρ, and the modular automorphism group associated to any such state leaves the algebra $\mathcal{N}$ invariant (we remark that this could directly be verified from the structure (6.7) of these states). In other words:

$$\Delta_\rho(\mathcal{N}) = \rho \mathcal{N} \rho^{-1} = \mathcal{N}. \tag{6.11}$$

As claimed above, the conditional expectation $E_{\mathcal{N}}$ is the orthogonal projection onto $\mathcal{N}$ with respect to $\langle.,.\rangle_{1,\rho}$, for any full-rank invariant state ρ. The only property left to be proved is the self-adjointness of $E_{\mathcal{N}}$ with respect to this inner product: for any $X, Y \in \mathcal{B}(\mathcal{H})$,

$$\begin{aligned}
\mathrm{Tr}(\rho E_{\mathcal{N}}[X] Y) &= \sum_{i \in I} \mathrm{Tr}(\rho [\mathrm{Tr}_{\mathcal{K}_i}((\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i) P_i X P_i) \otimes \mathbb{1}_{\mathcal{K}_i}] Y) \\
&= \sum_{i \in I} \mathrm{Tr}\left(\mathrm{Tr}_{\mathcal{K}_i}(\rho) \, \mathrm{Tr}_{\mathcal{K}_i}((\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i) P_i X P_i) \, \mathrm{Tr}_{\mathcal{K}_i}((\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i) P_i Y P_i) \right) \\
&= \sum_{i \in I} \mathrm{Tr}\left((\mathrm{Tr}_{\mathcal{K}_i}(\rho) \otimes \tau_i)(\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i) P_i X P_i \, [\mathrm{Tr}_{\mathcal{K}_i}((\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i) P_i Y P_i) \otimes \mathbb{1}_{\mathcal{K}_i}] \right) \\
&= \mathrm{Tr}(\rho X E_{\mathcal{N}}[Y]).
\end{aligned}$$

From Lemma 5.2.5, we directly get for free that $E_{\mathcal{N}}$ is also self-adjoint with respect to the KMS inner product $\langle .,.\rangle_\rho$. Equivalently:

$$\Gamma_\rho \circ E_{\mathcal{N}} = E_{\mathcal{N}*} \circ \Gamma_\rho. \tag{6.12}$$

The following theorem is then a direct consequence of the existence of $E_{\mathcal{N}}$ together with Theorem 6.1.3:

Theorem 6.2.7. *When $\mathcal{M} = \mathcal{B}(\mathcal{H})$, $\mathcal{H}$ finite dimensional, and under the condition of existence of a full-rank invariant state,* EID *occurs with $\mathcal{N}_0 = \mathcal{N}$ and $\mathcal{V}_0 = \mathcal{V} = \mathcal{N}^{\perp,\rho}$.*

Proof of Theorem 6.1.3 for $\mathcal{M} = \mathcal{B}(\mathcal{H})$, $\mathcal{H}$ finite dimensional : Given a full-rank invariant state ρ, decompose $\mathcal{B}(\mathcal{H})$ into $\mathcal{N} \oplus \mathcal{N}^{\perp,\rho} = \mathrm{im}(E_{\mathcal{N}}) \oplus \ker(E_{\mathcal{N}})$. Now, it is obvious that $\mathcal{N} \subset \mathcal{N}_0$, and equality holds since $\mathcal{N}$ is the largest subalgebra of $\mathcal{B}(\mathcal{H})$ on which $\mathcal{P}_t$ acts as a $*$-homomorphism for any $t \geq 0$, and since $\mathcal{P}_t|_{\mathcal{N}}$ is an automorphism. Moreover, $\mathcal{V} \cap \mathcal{N} = \{0\}$ and $\mathcal{V} \subseteq \mathcal{N}^{\perp,\rho}$. Indeed, for any $X \in \mathcal{N}$ and $Y \in \mathcal{V}$, by Equation (6.1):

$$\mathrm{Tr}(\rho\, XY) = \mathrm{Tr}(\rho\, \mathcal{P}_t(XY)) = \mathrm{Tr}(\rho\, \mathcal{P}_t(X)\mathcal{P}_t(Y)) \underset{t\to\infty}{\to} 0.$$

As for the reverse inclusion $\mathcal{N}^{\perp,\rho} \subset \mathcal{V}$, let $X \in \mathcal{N}^{\perp,\rho}$. Since $[E_{\mathcal{N}}, \mathcal{P}_t] = 0$ for any $t \geq 0$, $\mathcal{P}_t(X) \in \mathcal{N}^{\perp,\rho}$ for all time. The net $(\mathcal{P}_t(X))_{t\geq 0}$ being bounded, it admits a cluster point X_∞ by Bolzano–Weierstrass theorem. However, any cluster point of $(\mathcal{P}_t(X))_{t\geq 0}$ belongs to $\mathcal{N}$ (see the proof of Theorem 3.1 of [Frigerio, 1978]). Hence $X_\infty \in \mathcal{N} \cap \mathcal{N}^{\perp,\rho} = \{0\}$. Therefore, $\mathcal{V}_0 = \mathcal{V} = \mathcal{N}^{\perp,\rho}$, and the proof follows.

$\square$

An invariant state that will play an important role in Chapter 8 is

$$\sigma_{\mathrm{Tr}} := E_{\mathcal{N}*}\left(\frac{\mathbb{1}}{d_{\mathcal{H}}}\right) = \frac{1}{d_{\mathcal{H}}}\sum_{i\in I} d_{\mathcal{K}_i} \mathbb{1}_{\mathcal{H}_i} \otimes \tau_i. \tag{6.13}$$

The relevance of σ_{Tr} comes from the fact that σ_{Tr} is *tracial* on $\mathcal{N}$, that is, for all $X \in \mathcal{N}$ and all $Y \in \mathcal{B}(\mathcal{H})$,

$$\mathrm{Tr}(\sigma_{\mathrm{Tr}} XY) = \mathrm{Tr}(\sigma_{\mathrm{Tr}} YX).$$

We close this section with a discussion of the relevance of the assumption of existence of a full-rank invariant state. In [Carbone et al., 2011], decoherence was shown to occur for any QMS when $\dim(\mathcal{H}) = 2$. However, this is not the case in general, even in finite dimensions, as it was showed by a counterexample in [Carbone et al., 2013] even when there exists a (non full-rank) invariant state. In Theorem 6 of the same article, the authors showed some necessary and sufficient conditions: EID holds if and only if the algebra $\mathcal{N}$ contains all the eigenvectors of the generator $\mathcal{L}$ associated with eigenvalues with null real part. This is equivalent to the fact that $\mathcal{N}$ coincides with the linear space generated by these eigenvectors, which always holds true under the existence of a full-rank invariant state.

6.3. Quantum Markovian evolutions in discrete time

Similarly to the classical setting, the theory of quantum Markov chains in discrete time is well-established and shares some similarities with its continuous time counterpart described above. Here, we regroup some basic facts concerning quantum Markov chains on finite dimensional Hilbert spaces that are going to be useful in Chapter 14. For more details on these and other basic facts of quantum information theory, we refer e.g. to [Wolf, 2012].

First, we recall that given a linear map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, $\mathcal{H}$ finite dimensional, its spectrum $\mathrm{sp}(\Phi)$ coincides with the set of λ's such that there exists $X \in \mathcal{B}(\mathcal{H})$ for which

$$\Phi(X) = \lambda X. \tag{6.14}$$

Moreover, $\mathrm{sp}(\Phi) = \mathrm{sp}(\Phi^*)$, where Φ^* corresponds to the dual map with respect to the Hilbert Schmidt inner product $\langle A, B \rangle_{\mathrm{HS}} := \mathrm{Tr}(A^*B)$. Now, like any linear operator, Φ admits a Jordan decomposition:

$$\Phi = \Phi_P + \Phi_Q, \quad \Phi_P = \sum_{k:\, \lambda_k \text{ peripheral}} \lambda_k P_k + N_k, \quad \Phi_Q = \sum_{k:\, \lambda_k \text{ not peripheral}} \lambda_k P_k + N_k, \tag{6.15}$$

where λ_k are the eigenvalues of Φ, P_k the associated (not necessarily orthogonal) eigenprojections, and $N_k^{d_k} = 0$, where $d_k := \mathrm{Tr}(P_k)$, so that $\sum_k P_k = \mathbb{1}$. For any k, $\lambda_k P_k + N_k$ constitutes the k-th Jordan block of Φ. The linear span $\tilde{\mathcal{N}}(\Phi)$ of the peripheral points is called *phase subspace*, and we denote by $\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}$ the projection onto it. As we will see in Section 6.3.1, the phase subspace is what takes the role of the decoherence-free subalgebra for discrete time evolutions.

In particular, if Φ is hermiticity preserving, (6.14) implies that the eigenvalues of Φ either are real, or come in conjugate pairs. If, moreover, Φ is positive unital ($\Phi(\mathbb{1}) = \mathbb{1}$) or trace preserving ($\mathrm{Tr}\,\Phi(A) = \mathrm{Tr}(A)$ for all $A \in \mathcal{B}(\mathcal{H})$), $1 \in \mathrm{sp}(\Phi)$ and all the other eigenvalues of Φ lie in the unit disc of the complex plane, and the eigenvalues lying on the peripheral spectrum are associated to one-dimensional Jordan blocks, so that $|\lambda_k| = 1 \Rightarrow N_k = 0$.

6.3.1. Decoherence for discrete time quantum Markov chains

Given a general quantum channel $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, its phase subspace introduced in Section 6.3 is known to possess the following structure (Theorem 6.16 of [Wolf, 2012], Theorem 8 of [Wolf and Perez-Garcia, 2010]): there exists a decomposition of $\mathcal{H}$ as $\mathcal{H} = \bigoplus_{j \in \mathcal{J}} \mathcal{H}_j \otimes \mathcal{K}_j \oplus \mathcal{K}_0$ such that

$$\tilde{\mathcal{N}}(\Phi) := \bigoplus_{i \in \mathcal{J}} \mathcal{B}(\mathcal{H}_i) \otimes \tau_i \oplus 0_{\mathcal{K}_0}, \quad \mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}(\rho \oplus 0_{\mathcal{K}_0}) = \sum_{i \in \mathcal{J}} \mathrm{Tr}_{\mathcal{K}_i}(p_i \rho p_i) \otimes \tau_i, \tag{6.16}$$

where p_i is the orthogonal projector onto the i-th subspace, for some fixed full-rank states $\tau_i \in \mathcal{D}(\mathcal{K}_i)$. Assuming that $\mathcal{K}_0 = \{0\}$, the range of the projector $\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}^*$, which is called the *decoherence-free subalgebra* of $\{\Phi^n\}_{n \in \mathbb{N}}$ and denoted by $\mathcal{N}(\Phi^*)$, has the form

$$\mathcal{N}(\Phi^*) := \bigoplus_{i \in \mathcal{J}} \mathcal{B}(\mathcal{H}_i) \otimes \mathbb{1}_{\mathcal{K}_i}. \tag{6.17}$$

This occurs when Φ is *faithful*, which means that it possesses a full-rank invariant state. For any such quantum Markov chain, the following convergence result is known: for any $\rho \in \mathcal{D}(\mathcal{H})$,

$$\Phi^n(\rho) - \Phi^n \circ \mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}(\rho) \to 0 \quad \text{as} \quad n \to \infty.$$

This is identical to the continuous time case. Here, $\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}$ plays the role of the (dual) conditional expectation $E_{\mathcal{N}_*}$ of Section 6.2. Coming back to the evolution of states, the map $\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}$ projecting states onto the peripheral subspace of Φ is such that $\sigma_{\mathrm{Tr}} := d_{\mathcal{H}}^{-1} \mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}(\mathbb{1})$ commutes with $\mathcal{N}(\Phi^*)$. The map $\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}^*$ is a conditional expectation and hence satisfies the following properties:

Lemma 6.3.1.

(i) $\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}^*$ *is adjoint preserving: for any* $X \in \mathcal{B}(\mathcal{H})$, $\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}^*(X^*) = \mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}^*(X)^*$;

(ii) for any $X \in \mathcal{B}(\mathcal{H})$ *and* $Y, Z \in \mathcal{N}(\Phi^*)$, $\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}^*(YXZ) = Y\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}^*(X)Z$;

(iii) for any $X \in \mathcal{B}(\mathcal{H})$, $\mathrm{Tr}(\sigma_{\mathrm{Tr}} X) = \mathrm{Tr}(\sigma_{\mathrm{Tr}} \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(X))$;

(iv) $\mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}$ is self-adjoint with respect to σ_{Tr}: for any $X, Y \in \mathcal{B}(\mathcal{H})$.

$$\langle X, \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(Y) \rangle_{\sigma_{\mathrm{Tr}}} = \langle \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(X), Y \rangle_{\sigma_{\mathrm{Tr}}} = \langle \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(X), \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(Y) \rangle_{\sigma_{\mathrm{Tr}}}.$$

Proof.

(i) is obvious since $\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}$ itself is adjoint-preserving (cf. Proposition 6.3 of [Wolf, 2012]).

(ii) For $Y, Z \in \mathcal{N}(\Phi^*)$, we have the decomposition $Y = \sum_i Y_i \otimes \mathbb{1}_{\mathcal{K}_i}$, where $p_i(Y_i \otimes \mathbb{1}_{\mathcal{K}_i})p_i = Y_i \otimes \mathbb{1}_{\mathcal{K}_i}$, and similarly with Z. Then,

$$\begin{aligned}
\mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(YXZ) &= \sum_i \mathrm{Tr}_{\mathcal{K}_i}((\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i)p_i YXZ p_i) \otimes \mathbb{1}_{\mathcal{K}_i} \\
&= \sum_i \mathrm{Tr}_{\mathcal{K}_i}((\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i)(Y_i \otimes \mathbb{1}_{\mathcal{K}_i})p_i X p_i(Z_i \otimes \mathbb{1}_{\mathcal{K}_i})) \otimes \mathbb{1}_{\mathcal{K}_i} \\
&= \sum_i [\mathrm{Tr}_{\mathcal{K}_i}(Y_i(\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i)p_i X p_i)Z_i] \otimes \mathbb{1}_{\mathcal{K}_i} \\
&= \sum_j Y_j \otimes \mathbb{1}_{\mathcal{K}_j} \left(\sum_i \mathrm{Tr}_{\mathcal{K}_i}((\mathbb{1}_{\mathcal{H}_i} \otimes \tau_i)p_i X p_i) \otimes \mathbb{1}_{\mathcal{K}_i} \right) \sum_j Z_j \otimes \mathbb{1}_{\mathcal{K}_j} \\
&= Y \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(X) Z.
\end{aligned}$$

(iii) follows from a simple computation: since Φ is trace preserving, so is $\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}$ (Proposition 6.3 of [Wolf, 2012]), and hence

$$\mathrm{Tr}(\sigma_{\mathrm{Tr}} \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(X)) = \mathrm{Tr}(\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)} \circ \Gamma_{\sigma_{\mathrm{Tr}}}(X)) = \mathrm{Tr}(\Gamma_{\sigma_{\mathrm{Tr}}}(X)) = \mathrm{Tr}(\sigma_{\mathrm{Tr}} X).$$

(iv) is a consequence of (i)–(iii):

$$\begin{aligned}
\langle X, \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(Y) \rangle_{\sigma_{\mathrm{Tr}}} &= \mathrm{Tr}(\sigma_{\mathrm{Tr}}^{1/2} X^* \sigma_{\mathrm{Tr}}^{1/2} \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(Y)) \\
&= \mathrm{Tr}(\sigma_{\mathrm{Tr}} X^* \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(Y)) \\
&= \mathrm{Tr}(\sigma_{\mathrm{Tr}} \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(X^* \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(Y))) \\
&= \mathrm{Tr}(\sigma_{\mathrm{Tr}} \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(X^*) \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(Y)) \\
&= \mathrm{Tr}(\sigma_{\mathrm{Tr}} \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(X)^* \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(Y)) \\
&= \mathrm{Tr}(\sigma_{\mathrm{Tr}}^{1/2} \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(X)^* \sigma_{\mathrm{Tr}}^{1/2} \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(Y)) \\
&= \langle \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(X), \mathcal{P}^*_{\tilde{\mathcal{N}}(\Phi)}(Y) \rangle_{\sigma_{\mathrm{Tr}}}
\end{aligned}$$

where we used the commutativity of σ_{Tr} with $\mathcal{N}(\Phi^*)$ in the second and sixth lines, (i) in the fifth line, (ii) in the fourth line and (iii) in the third line. The first identity in (iv) follows by symmetry. $\qquad\square$

6.4. Examples

6.4.1. Simple semigroups associated to a conditional expectation

The depolarizing semigroup introduced in Section 5.5.1 corresponds to a very symmetric situation where the noise added to the quantum system is assumed to be isotropic. Assuming that $\mathcal{H} = \mathbb{C}^2$, this

is clear from the following reformulation:

$$\mathcal{P}_*^{\mathrm{depol}}(\rho) = (1-p)\rho + p\frac{\mathbb{1}}{2}$$

$$= (1-p)\rho + \frac{p}{4}\sum_{i=0}^{3}\sigma_i\,\rho\,\sigma_i,$$

where $\sigma_0 = \frac{\mathbb{1}}{2}$ and

$$\sigma_1 := \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 := \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 := \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

are the Pauli noise operators in each direction of a computational basis that is assumed to be fixed beforehand. Hence,

$$\mathcal{P}_*^{\mathrm{depol}}(\rho) = \left(1 - \frac{3p}{4}\right)\rho + \frac{p}{4}\left(\sigma_x\,\rho\,\sigma_x + \sigma_y\,\rho\,\sigma_y + \sigma_z\,\rho\,\sigma_z\right).$$

Hence, the depolarizing channel can be interpreted as follows: with probability $\frac{p}{4}$, the noise σ_i in the direction i is applied to the state uniformly at random. Otherwise, the state remains untouched.

On the contrary, the quantum system might get affected in a single direction, say along the z axis. The resulting quantum channel, known as *dephasing channel* (or *phase flip channel*), takes the following simpler form:

$$\mathcal{P}_*^{\mathrm{deph}}(\rho) = (1-p)\rho + p\,\sigma_z\,\rho\,\sigma_z.$$

More generally, one considers the so-called *Pauli channel*:

$$\mathcal{P}_*^{\mathrm{Pauli}}(\rho) = \sum_{i=0}^{3} p(i)\,\sigma_i\,\rho\,\sigma_i$$

for a given probability mass function p on $\mathbb{Z}_4$. Assuming that the probability that an error occurs is uniform among the subset of the Pauli matrices, with associated index set I, the Pauli channel reduces to the following:

$$\mathcal{P}_*^I(\rho) = (1-p)\,\rho + \frac{p}{|I|}\sum_{i\in I}\sigma_i\,\rho\,\sigma_i.$$

As in the case of the depolarizing semigroup, this quantum channel has a simple semigroup analogue, when assuming that the probability $1-p$ vanishes exponentially fast: for all $X \in \mathcal{B}(\mathbb{C}^2)$:

$$\mathcal{P}_t^I(X) = \mathrm{e}^{-t}\,X + \frac{1-\mathrm{e}^{-t}}{|I|}\sum_{i\in I}\sigma_i\,X\,\sigma_i \qquad \Rightarrow \qquad \mathcal{L}^I(X) = \frac{1}{|I|}\sum_{i\in|I|}\sigma_i\,X\,\sigma_i - X.$$

The QMS $(\mathcal{P}_t^I)_{t\geq 0}$ is a perfect example of a non primitive semigroup. First of all, one easily reads its associated Lindblad operators (cf. Equation (5.5)) from the above expression of its generator: $H = 0$ and $L_j = |I|^{-\frac{1}{2}}\sigma_i$. Theorem 6.2.1 leads to

$$\mathcal{N}(\mathcal{P}^I) = \{\sigma_i,\ i\in I\}'.$$

In fact, the completely positive map $X \mapsto |I|^{-1}\sum_{i\in I}\sigma_i\,X\,\sigma_i$ is a conditional expectation associated to the algebra $\mathcal{N}(\mathcal{P}^I)$ with respect to the completely mixed state $\frac{\mathbb{1}}{2}$.

More generally, let $\mathcal{N}$ be a $*$-subalgebra of $\mathcal{B}(\mathcal{H})$, $\mathcal{H}$ being finite dimensional. As we saw in Section 0.1.4, $\mathcal{N}$ can be written as

$$\mathcal{N} \coloneqq \bigoplus_{i \in I} \mathcal{B}(\mathcal{H}_i) \otimes \mathbb{1}_{\mathcal{K}_i}, \qquad \mathcal{H} \coloneqq \bigoplus_{i \in I} \mathcal{H}_i \otimes \mathcal{K}_i.$$

Choosing σ_{Tr} as in Equation (6.13) and letting $E_{\mathcal{N}} : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ be its associated conditional expectation, the following is a valid generator of a QMS of associated decoherence free subalgebra $\mathcal{N}$:

$$\mathcal{L}^{\mathcal{N}}(X) = E_{\mathcal{N}}[X] - X.$$

We call the associated QMS $(\mathcal{P}_t^{\mathcal{N}})_{t \geq 0}$ the *simple QMS associated to* $\mathcal{N}$.

Lemma 6.4.1. *For any $*$-subalgebra $\mathcal{N}$ of $\mathcal{B}(\mathcal{H})$, and any conditional expectation $E_{\mathcal{N}}$ with respect to the state σ_{Tr}, the simple QMS $(\mathcal{P}_t^{\mathcal{N}})_{t \geq 0}$ satisfies σ_{Tr}-DBC. Moreover,*

$$E_{\mathcal{N}} \circ \Delta_{\sigma_{\mathrm{Tr}}} = \Delta_{\sigma_{\mathrm{Tr}}} \circ E_{\mathcal{N}} \qquad and \qquad \Gamma_{\sigma_{\mathrm{Tr}}} \circ E_{\mathcal{N}} = E_{\mathcal{N}*} \circ \Gamma_{\sigma_{\mathrm{Tr}}}. \tag{6.18}$$

Proof. In view of Lemma 5.2.5, it suffices to prove that $\mathrm{Tr}(\sigma_{\mathrm{Tr}} E_{\mathcal{N}}[X]Y) = \mathrm{Tr}(\sigma_{\mathrm{Tr}} X E_{\mathcal{N}}[Y])$ for any $X, Y \in \mathcal{B}(\mathcal{H})$. But this follows by a direct computation using the expression (6.9) for $E_{\mathcal{N}}$. $\qquad \square$

6.4.2. Collective decoherence

Here, we describe an important type of non primitive quantum Markovian evolution that has found application in quantum error correction, since it is believed to be a good candidate for the implementation of fault tolerant universal quantum computation devices [Lidar et al., 1998]. Here, the system under consideration is constituted of n qubits, that is $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$, which experience the exact same noise. This leads to an overall invariance of the evolution under permutation of the qubits that is referred to as *collective decoherence*. In the case when the system is subject to a collective dephasing along the z-axis, the generator of the resulting *weak collective decoherence* is given by

$$\mathcal{L}_n^{\mathrm{wcd}}(X) \coloneqq \Sigma_z^n \, X \, \Sigma_z^n - \frac{1}{2}((\Sigma_z^n)^2 X + X(\Sigma_z^n)^2), \qquad \text{where} \qquad \Sigma_z^n \coloneqq \sum_{i=1}^n \mathbb{1}_{\mathbb{C}^2}^{\otimes i-1} \otimes \sigma_z \otimes \mathbb{1}_{\mathbb{C}^2}^{n-i}.$$

More generally, if any collective Pauli noise is allowed, with associated index set $I \subset \{1, 2, 3\}$:

$$\mathcal{L}_n^{\mathrm{cd},I}(X) \coloneqq \sum_{i \in I} \Sigma_i^n \, X \, \Sigma_i^n - \frac{1}{2}((\Sigma_i^n)^2 X + X(\Sigma_i^n)^2), \qquad \text{where} \qquad \Sigma_i^n \coloneqq \sum_{k=1}^n \mathbb{1}_{\mathbb{C}^2}^{\otimes k-1} \otimes \sigma_i \otimes \mathbb{1}_{\mathbb{C}^2}^{n-k}.$$

The case when $I = \{1, 2, 3\}$ is commonly referred to as *strong collective decoherence*, and in this case we denote the generator by $\mathcal{L}_n^{\mathrm{scd}}$.

Decoherence free subsystems and quantum error protection As yet another approach to decoherence, [Lidar et al., 1998] (see also [Ticozzi and Viola, 2008, Lidar and Whaley, 2003, Lidar, 2014, Knill et al., 2000, Viola et al., 2001, Kempe et al., 2001]) looked at the evolution in the Schrödinger picture. In particular, they identified parts of the Hilbert space where the evolution is unitary. These so-called *decoherence-free subsystems* have been proposed as possible candidates to encode quantum information that is protected from the environmental noise. More precisely, given a decomposition of $\mathcal{H}$ as follows:

$$\mathcal{H} = \mathcal{H}_s \otimes \mathcal{H}_f \oplus \mathcal{H}_r.$$

the space $\mathcal{H}_s$ is said to support a decoherence-free/noiseless subsystem for a given QMS $(\mathcal{P}_t)_{t\geq 0}$ if for any initial state $\rho = \rho_s \otimes \rho_f$ supported on $\mathcal{H}_s \otimes \mathcal{H}_f$, and any $t \geq 0$:

$$\mathcal{P}_t(\rho) = U_t \rho_s U_t^* \otimes \mathcal{P}_t^f(\rho_f),$$

where U_t is a unitary operator on $\mathcal{H}_s$ and $(\mathcal{P}_t^f)_{t\geq 0}$ is a QMS on $\mathcal{B}(\mathcal{H}_f)$. In the case when $\dim(\mathcal{H}_f) = 1$, $\mathcal{H}_s$ is said to support a *decoherence-free subspace*.

The following theorem, a proof of which can be found in [Deschamps et al., 2016] in the uniformly continuous case, provides an identification of decoherence-free subsystems of a QMS $(\mathcal{P}_t)_{t\geq 0}$ in terms of the structure of its decoherence-free algebra (in the case of unbounded generators, see [Agredo et al., 2014]):

Theorem 6.4.2. *Given the decomposition (6.5) of the decoherence-free agebra of a QMS $(\mathcal{P}_t)_{t\geq 0}$, each $\mathcal{H}_i$ supports a decoherence-free subsystem. Moreover, any direct sum $\bigoplus_{i \in I'} \mathcal{H}_i$, where $I' \subset I$ is such that for all $i \in I'$, $\dim(\mathcal{K}_i) = 1$, supports a decoherence-free subspace.*

Part IV.

Quantum functional inequalities

Chapter 7.

Primitive quantum Markov semigroups

In Part II, we saw that functional inequalities represent a powerful tool to estimate mixing times of classical Markov semigroups. Generalizations of these tools were initiated in [Gross, 1975a, Lindsay, 1990, Carlen and Loss, 1993, Carlen and Lieb, 1993, Biane, 1997] for specific semigroups. The notion of hypercontractivity and logarithmic Sobolev inequalities was fully extended in [Olkiewicz and Zegarlinski, 1999] to arbitrary reversible primitive semigroups on non-commutative $\mathbb{L}_p$-spaces defined over approximately finite C^*-algebras. The exploration of the asymptotic behavior of a QMS satisfying a functional inequality has since become an area of focus in the quantum information and quantum probability communities [Carbone and Sasso, 2008, Temme et al., 2010, Montanaro, 2012, Carbone and Martinelli, 2015, Kastoryano and Temme, 2013, Temme et al., 2014, Cubitt et al., 2015, Kastoryano and Temme, 2016, Müller-Hermes and França, 2018, Delgosha and Beigi, 2014, Beigi and King, 2016, Junge et al., 2015, Carbone, 2004, Carbone and Fagnola, 2000].

Layout of the chapter: In this chapter, we survey the quantum versions of the functional inequalities of Chapter 3, namely Poincaré inequality (Section 7.1), logarithmic Sobolev inequalities and hypercontractivity (Section 7.2), Nash inequalities and ultracontractivity (Section 7.5), in the case of a primitive quantum Markov semigroup. Apart from Theorem 7.2.4 which provides the exact logarithmic Sobolev constant of order 2 for the generalized quantum depolarizing semigroup, the results discussed in Sections 7.1 to 7.3 and 7.5.1 are not new. In contrast, the notion of quantum reverse hypercontractivity for non doubly-stochastic QMS as discussed in Section 7.4 extends the one of [Cubitt et al., 2015]. We end this chapter with the statement and proof of a quantum Nash inequality for the quantum heat semigroup (cf. Section 5.5.2) in Section 7.5.2. The functional inequalities described in this chapter, and in particular Section 7.2, will be extended to the case of non-primitive QMS in Chapter 8.

7.1. Quantum Poincaré inequality

Let $\mathcal{H}$ be a separable Hilbert space, and fix a faithful normal state σ on $\mathcal{H}$. An $\mathbb{L}_2(\sigma)$ QMS $(\mathcal{P}_t)_{t \geq 0}$ satisfies a *quantum Poincaré inequality* with constant $\lambda > 0$, if for all $X \in \mathrm{dom}(\mathcal{L}_2)$,[1]

$$\lambda \mathrm{Var}_\sigma(X) \leq \mathcal{E}_{2,\mathcal{L}}(X), \tag{PI(λ)}$$

---

[1] The following discussion is very general and would apply to any strongly continuous contraction semigroup $(\mathcal{P}_t)_{t \geq 0}$ on a Hilbert space $\mathcal{K}$ with an invariant vector ψ.

where $\mathrm{Var}_\sigma(X) := \|X\|^2_{\mathbb{L}_2(\sigma)} - \mathrm{Tr}(\sigma X)^2 = \|X - \mathrm{Tr}(\sigma X)\,\mathbb{1}\|^2_{\mathbb{L}_2(\sigma)}$. The *spectral gap* of $\mathcal{L}$ is defined by the following nonnegative number:

$$\lambda(\mathcal{L}) = \inf_{X \in \mathrm{dom}(\mathcal{L}_2)} -\frac{\mathrm{Re}\langle X, \mathcal{L}_2 X\rangle_\sigma}{\|X - \langle \mathbb{1}, X\rangle_\sigma \mathbb{1}\|^2_{\mathbb{L}_2(\sigma)}} \equiv \inf_{X \in \mathrm{dom}(\mathcal{L}_2)} \frac{\mathcal{E}_{2,\mathcal{L}}(X)}{\mathrm{Var}_\sigma(X)},$$

The next proposition, whose proof consists in a simple differentiation of (7.1), is a straightforward extension of Proposition 3.1.1. It can be found e.g. in [Carbone and Fagnola, 2000]:

Proposition 7.1.1. *Under the above conditions, the spectral gap of $\mathcal{L}_2$ is the maximum positive value λ such that the following exponential decay occurs in $\mathbb{L}_2(\sigma)$: for any $X \in \mathbb{L}_2(\sigma)$ and all $t \geq 0$,*

$$\|\mathcal{P}_t(X) - \mathrm{Tr}(\sigma X)\,\mathbb{1}\|_{\mathbb{L}_2(\sigma)} \leq \mathrm{e}^{-\lambda t}\,\|X - \mathrm{Tr}(\sigma X)\,\mathbb{1}\|_{\mathbb{L}_2(\sigma)}. \tag{7.1}$$

In other words, the spectral gap corresponds to the maximal positive value λ such that

$$\mathrm{Var}_\sigma(\mathcal{P}_t(X)) \leq \mathrm{e}^{-2\lambda t}\,\mathrm{Var}_\sigma(X).$$

Finding the spectral gap of quantum Markov semigroups is already a challenging problem (see [Kastoryano and Brandão, 2016]). However, they usually don't provide the tightest bounds on the mixing time of a QMS. In the next section, we review the more refined noncommutative theory of hypercontractivity.

7.2. Quantum logarithmic Sobolev inequality, hypercontractivity

As mentioned in the preamble of the chapter, the theory of hypercontractivity for primitive QMS was fully formalized in [Olkiewicz and Zegarlinski, 1999], using Kosaki's theory of non-commutative interpolating weighted $\mathbb{L}_p$ spaces [Kosaki, 1984, Majewski and Zegarlinski, 1996], where the weights here are given in terms of a faithful invariant state of the evolution. This study was further pursued by different authors [Temme et al., 2014, Carbone and Martinelli, 2015] and applied to the problem of estimating mixing times in [Kastoryano and Temme, 2013]. We recall that, given a subalgebra $\mathcal{A}_0$ satisfying Condition 5.2.2, $\mathcal{A}_0^+$ denotes the subset $\{X + c\,\mathbb{1}, X \in \mathcal{A}_0, X \geq 0, c > 0\}$ of positive definite operators whose spectrum is uniformly bounded away from 0. Then, for $p \in \mathbb{R}\backslash\{0\}$, define the *entropy function of order p* as follows: for all $X \in \mathcal{A}_0^+$,

$$\mathrm{Ent}_{p,\sigma}(X) := \mathrm{Tr}\left(\Gamma_\sigma^{\frac{1}{p}}(X)\right)^p \ln\left(\Gamma_\sigma^{\frac{1}{p}}(X)\right)^p - \mathrm{Tr}\left(\Gamma_\sigma^{\frac{1}{p}}(X)\right)^p \ln\sigma - \|X\|^p_{\mathbb{L}_p(\sigma)} \ln\|X\|^p_{\mathbb{L}_p(\sigma)}, \tag{7.2}$$

where the family $\|.\|_p$ can be naturally extended on $\mathcal{A}_0^+$ for $p \leq 1$ (see Section 7.4). Similarly, extending the definition of the maps $\mathrm{I}_{q,p}$ on $\mathcal{A}_0^+$ to any $q, p \in \mathbb{R}\backslash\{0\}$, the following well-known properties of the entropy function are easy to verify.

Lemma 7.2.1. *Let $X \in \mathcal{A}_0^+$. Then, for any $p, q \in \mathbb{R}\backslash\{0\}$:*

   (i) $\mathrm{Ent}_{p,\sigma}(\mathrm{I}_{p,2}(X)) = \mathrm{Ent}_{q,\sigma}(\mathrm{I}_{q,2}(X))$;

   (ii) $\mathrm{Ent}_{p,\sigma}(cX) = c^p\,\mathrm{Ent}_{p,\sigma}(X)$ *for any $c > 0$;*

   (iii) $\mathrm{Ent}_{p,\sigma}\left(X\right) = D(\rho\|\sigma) - \mathrm{Tr}\,\rho \ln \mathrm{Tr}\,\rho$, *where $\rho = \left(\Gamma_\sigma^{\frac{1}{p}}(X)\right)^p$, and $D(\rho\|\sigma) = \mathrm{Tr}(\rho \ln \rho) - \mathrm{Tr}(\rho \ln \sigma)$ is Umegaki's relative entropy.*

We are now ready to define quantum logarithmic Sobolev inequalities:

Definition 7.2.2. Given $p > 1$, $(\mathcal{P}_t)_{t\geq 0}$ is said to satisfy a *quantum logarithmic Sobolev inequality* of order p if for any $X \in \mathcal{A}_0^+$:

$$\text{Ent}_{p,\sigma}(X) \leq c\,\mathcal{E}_{p,\mathcal{L}}(X) + 2d\,\|X\|_{\mathbb{L}_p(\sigma)}^p. \qquad (\text{qLSI}_p(c,d))$$

In the case when $d \neq 0$, the logarithmic Sobolev inequality is said to be *defective*. In fact, one can always reduce to the case $d = 0$ in the primitive, finite dimensional setting[2]. This legitimates the definition of the *p-logarithmic Sobolev constant* as

$$\alpha_p(\mathcal{L}) := \inf_{X\in\mathcal{A}_0^+} \frac{\mathcal{E}_{p,\mathcal{L}}(X)}{\text{Ent}_{p,\sigma}(X)}. \qquad (7.3)$$

The $\mathbb{L}_2(\sigma)$-QMS $(\mathcal{P}_t)_{t\geq 0}$ is hypercontractive if for all $p \leq q \leq 1 + (p-1)\,e^{\frac{4t}{c}}$, the following holds:

$$\|\mathcal{P}_t : \mathbb{L}_p(\sigma) \to \mathbb{L}_q(\sigma)\| \leq \exp\left\{2d\left(\frac{1}{p} - \frac{1}{q}\right)\right\}. \qquad (\text{HC}_p(c,d))$$

Quantum logarithmic Sobolev inequalities reduce to $\text{LSI}_p(c, d)$ in the commutative case. In fact, $\text{qLSI}_p(c,d)$ is also well-defined for $p \in \mathbb{R}\backslash\{0,1\}$. The range $p \in (-\infty, 0) \cup (0, 1)$ will be studied in Section 7.4, whereas the limiting case $p = 1$ gives the modified logarithmic Sobolev inequality introduced in Section 7.3.

The following theorem was first proved by [Olkiewicz and Zegarlinski, 1999]:

Theorem 7.2.3. *Let $(\mathcal{P}_t)_{t\geq 0}$ be an $\mathbb{L}_2(\sigma)$-quantum Markov semigroup, $p > 1$, $c > 0$ and $d \geq 0$. Then*

(i) *If $\text{HC}_p(c,d)$ holds, then $\text{qLSI}_p(c,d)$ holds.*

(ii) *If $\text{qLSI}_q(c,d)$ holds for all $q \geq p$, then $\text{HC}_p(c,d)$ holds.*

Moreover, if $(\mathcal{P}_t)_{t\geq 0}$ satisfies ω-DBC with respect to σ, then $\text{HC}_p(c,d)$ holds for any $p > 1$ with $d = 0$ and $c = \alpha_2(\mathcal{L})^{-1}$.

Proof. Items (i) and (ii) were proved in [Olkiewicz and Zegarlinski, 1999]. The equivalence under the condition of ω-DBC in a easy consequence of the quantum Stroock-Varopoulos inequality as well as part (i) of Proposition 7.2.1, from which we get that $p \mapsto \alpha_p(\mathcal{L})$ is non-increasing on $[1, 2]$ and that $\alpha_p(\mathcal{L}) = \alpha_{\hat{p}}(\mathcal{L})$ (see Proposition 7.4.6). $\qquad\square$

Example: the generalized depolarizing semigroup   We now give the exact expression of the 2-logarithmic Sobolev constant of the generalized depolarizing semigroup introduced in Section 5.5.1. To prove Theorem 7.2.4 we need to show that a certain function of qubit density matrices is optimized over diagonal ones. Once we show this, the explicit expression for the 2-logarithmic Sobolev constant is obtained from the associated classical constant derived in [Diaconis and Saloff-Coste, 1996a].

Theorem 7.2.4. *Let $\sigma \in \mathcal{D}_+(\mathcal{H})$ be arbitrary and let $\mathcal{L}_\sigma(X) = \text{Tr}(\sigma X)\mathbb{1} - X$. Then we have*

$$\alpha_2(\mathcal{L}_\sigma) = \frac{1 - 2\|\sigma^{-1}\|_\infty^{-1}}{\ln(\|\sigma^{-1}\|_\infty - 1)}. \qquad (7.4)$$

Proof. Since both $\text{Ent}_{2,\sigma}(X)$ and $\mathcal{E}_{2,\mathcal{L}_\sigma}(X)$ are homogenous of degree two, to prove a log-Sobolev inequality, without loss of generality we can assume that X is of the form $X = \Gamma_\sigma^{-1/2}(\sqrt{\rho})$ where ρ is a

---

[2]This is in sharp contrast with the non primitive case that we introduce in Chapter 8.

density matrix. In this case

$$\text{Ent}_{2,\sigma}(X) = D(\rho\|\sigma), \qquad -\langle X, \mathcal{L}_\sigma X\rangle_\sigma = 1 - \left[\text{Tr}\left(\sqrt{\sigma}\sqrt{\rho}\right)\right]^2.$$

Let $\sigma = \sum_{i=1}^d s_i \,|i\rangle\langle i|$ and $\rho = \sum_{k=1}^d r_k \,|\tilde{k}\rangle\langle\tilde{k}|$ be the eigen-decompositions of σ and ρ. Then

$$\text{Ent}_{2,\sigma}(X) = \sum_{k=1}^d r_k \ln r_k - \sum_{i,k=1}^d |\langle i|\tilde{k}\rangle|^2 r_k \ln s_i\,,$$

and

$$-\langle X, \mathcal{L}_\sigma X\rangle_\sigma = 1 - \left(\sum_{i,k=1}^d |\langle i|\tilde{k}\rangle|^2 \sqrt{s_i r_k}\right)^2.$$

Let $A = (a_{ik})_{d\times d}$ be a $d \times d$ matrix whose entries are given by

$$a_{ik} = |\langle i|\tilde{k}\rangle|^2.$$

Observe that, fixing the eigenvalues s_i's and r_k's, the entropy $\text{Ent}_{2,\sigma}(X)$ is a linear function of A and $\mathcal{E}_{2,\mathcal{L}_\sigma}(X)$ is a concave function of A. On the other hand, since both $\{|1\rangle, \ldots, |d\rangle\}$ and $\{|\tilde{1}\rangle, \ldots, |\tilde{d}\rangle\}$ form orthonormal bases, A is a doubly stochastic matrix. Then by Birkhoff's theorem, A can be written as a convex combination of permutations matrices. We conclude that if an inequality of the form

$$\alpha\left(\sum_{k=1}^d r_k \ln r_k - \sum_{i,k=1}^d a_{ik} r_k \ln s_i\right) \le 1 - \left(\sum_{i,k=1}^d a_{ik}\sqrt{s_i r_k}\right)^2,$$

holds for all permutation matrices A, then it holds for all doubly stochastic A, and then for all σ, ρ with the given eigenvalues. We note that A is a permutation matrix when $\{|1\rangle, \ldots, |d\rangle\}$ and $\{|\tilde{1}\rangle, \ldots, |\tilde{d}\rangle\}$ are the same bases (under some permutation) which means that σ and ρ commute. Therefore, a logarithmic Sobolev inequality of the form

$$\alpha\,\text{Ent}_{2,\sigma}\left(\Gamma_\sigma^{-1/2}(\rho)\right) \le \mathcal{E}_{2,\mathcal{L}_\sigma}\left(\Gamma_\sigma^{-1/2}(\rho)\right)$$

holds for all ρ if and only if it holds for all ρ that commute with σ. That is, to find the log-Sobolev constant

$$\alpha_2(\mathcal{L}_\sigma) = \inf_\rho \frac{\mathcal{E}_{2,\mathcal{L}_\sigma}\left(\Gamma_\sigma^{-1/2}(\rho)\right)}{\text{Ent}_{2,\sigma}\left(\Gamma_\sigma^{-1/2}(\rho)\right)},$$

we may restrict to those ρ that commute with σ. This optimization problem over such ρ is equivalent to computing the 2-log-Sobolev constant of the *classical* depolarizing Lindblad generator, and has been solved in Theorem A.1 of [Diaconis and Saloff-Coste, 1996a]. □

Remark 7.2.5. The above result should be compared with the expression for the MLSI constant derived in [Müller-Hermes et al., 2016] (see Section 7.3):

$$\alpha_1(\mathcal{L}_\sigma) = \min_{x\in[0,1]} \frac{1}{4}\left\{1 + \frac{D_{\text{bin}}(\|\sigma^{-1}\|_\infty^{-1}\|x)}{D_{\text{bin}}(x\|\|\sigma^{-1}\|_\infty^{-1})}\right\} \ge \frac{1}{4}, \tag{7.5}$$

where $D_{\text{bin}}(x\|y) := x \ln\frac{x}{y} + (1-x)\ln\frac{1-x}{1-y}$ denotes the binary relative entropy between the distributions $\{x, 1-x\}$ and $\{y, 1-y\}$.

7.3. Quantum modified logarithmic Sobolev inequality

Here, we restrict ourselves to a finite dimensional Hilbert space $\mathcal{H}$.

Theorem 7.3.1. *Let $(\mathcal{P}_t)_{t\geq0}$ be a primitive QMS on $\mathcal{B}(\mathcal{H})$ with invariant state σ. Then, for any $\rho \in \mathcal{D}(\mathcal{H})$, the following holds:*

$$\frac{d}{dt} D(\mathcal{P}_{*t}(\rho)\|\sigma) = \mathrm{Tr}\left(\mathcal{L}_*(\rho_t)(\ln\rho_t - \ln\sigma)\right).$$

The result follows by a simple application of the integral representation for the logarithm (see [Kastoryano and Temme, 2013]).

Theorem 7.3.1 provides a justification for the following definition:

Definition 7.3.2. For any $\rho = \Gamma_\sigma(X)$, $X > 0$, the *quantum entropy production* of ρ along the semigroup $(\mathcal{P}_t)_{t\geq0}$ is defined as:

$$\mathrm{EP}_\sigma(\rho) = -\left.\frac{d}{dt}\right|_{t=0} D(\rho_t\|\sigma) = -\mathrm{Tr}(\mathcal{L}_*(\rho)(\ln\rho - \ln\sigma)).$$

Then, the QMS $(\mathcal{P}_t)_{t\geq0}$ is said to satisfy a *quantum modified log-Sobolev inequality* with constant $\alpha_1 > 0$ if for all $\rho \in \mathcal{D}_+(\mathcal{H})$,

$$4\alpha_1 D(\rho\|\sigma) \leq \mathrm{EP}_\sigma(\rho). \tag{MLSI(α_1)}$$

which implies the following convergence:

$$\|\rho_t - \sigma\|_1 \leq \sqrt{2\ln\|\sigma^{-1}\|_\infty}\, \mathrm{e}^{-2\alpha_1(\mathcal{L})t}. \tag{7.6}$$

7.4. Quantum reverse hypercontractivity

Until now, we restricted our analysis to p-logarithmic Sobolev inequalities and hypercontractivity in the range $p > 1$. It turns out that, for $p < 1$, the contractivity properties of the maps $\mathcal{P}_t$ are flipped. This was first observed in the classical literature in [Borell, 1982, Mossel et al., 2006, Mossel et al., 2013], and extended to the quantum case for finite dimensional primitive semigroups whose invariant state is completely mixed in [Cubitt et al., 2015]. In this section we develop the theory of *quantum reverse hypercontractivity* beyond this so-called doubly stochastic case. The analysis carried out here is analogous to Section 7.2.

First, given a normal faithful state on $\mathcal{B}(\mathcal{H})$, $\mathcal{H}$ finite dimensional, we extend the definition of $\mathbb{L}_p(\sigma)$ norms to the range $p < 1$, $p \neq 0$: given $X \in \mathcal{B}(\mathcal{H})^+ := \{Y \in \mathcal{B}(\mathcal{H}), Y = Z + c\mathbb{1}, Z \geq 0, c > 0\}$:

$$\|X\|_{\mathbb{L}_p(\sigma)} := \mathrm{Tr}\left(|\Gamma_\sigma^{\frac{1}{p}}(X)|^p\right)^{\frac{1}{p}}.$$

Similarly to the classical case, $\|.\|_{\mathbb{L}_p(\sigma)}$ is simply a pseudo-norm for $p < 1$.

Lemma 7.4.1 (Reverse Hölder inequality). *Let $X, Y \in \mathcal{B}(\mathcal{H})^+$. Then, for any $0 < p < 1$ of Hölder conjugate $\hat{p}$, we have*

$$\langle X, Y\rangle_\sigma \geq \|X\|_{\mathbb{L}_p(\sigma)}\|Y\|_{\mathbb{L}_{\hat{p}}(\sigma)}.$$

Proof. The proof is a direct generalization of equation (32) of [Tomamichel et al., 2014]: for any two positive, definite operators $A, B \in \mathcal{B}(\mathcal{H})$:

$$\mathrm{Tr}(AB) \geq \|B\|_{\hat{p}} \|A\|_p. \tag{7.7}$$

From there, choosing $A := \Gamma_\sigma^{\frac{1}{p}}(X)$ and $B := \Gamma_\sigma^{\frac{1}{\hat{p}}}(Y)$,

$$\langle X, Y \rangle_\sigma := \mathrm{Tr}\left(\sigma^{\frac{1}{2p}} X \sigma^{\frac{1}{2\hat{p}}} \sigma^{\frac{1}{2\hat{p}}} Y \sigma^{\frac{1}{2\hat{p}}}\right) = \mathrm{Tr}(AB) \geq \|A\|_p \|B\|_{\hat{p}} = \|X\|_{\mathbb{L}_p(\sigma)} \|Y\|_{\mathbb{L}_{\hat{p}}(\sigma)}.$$

$\square$

Lemma 7.4.2 (Reverse Minkowski inequality). *For $p < 1$, $p \neq 0$, and $X, Y \in \mathcal{B}(\mathcal{H})^+$,*

$$\|X\|_{\mathbb{L}_p(\sigma)} + \|Y\|_{\mathbb{L}_p(\sigma)} \leq \|X + Y\|_{\mathbb{L}_p(\sigma)}.$$

Proof. This inequality in the special case of σ being the completely mixed state is proven in [Cubitt et al., 2015] but the generalization to arbitrary σ is trivial and follows from the reverse Hölder inequality:

$$\begin{aligned}
\|X + Y\|_{\mathbb{L}_p(\sigma)}^p &= \left\|\left(\Gamma_\sigma^{\frac{1}{p}}(X + Y)\right)^p\right\|_1 \\
&= \left\|\Gamma_\sigma^{\frac{1}{p}}(X + Y)\left(\Gamma_\sigma^{\frac{1}{p}}(X + Y)\right)^{p-1}\right\|_1 \\
&= \left\|\Gamma_\sigma^{\frac{1}{p}}(X)\left(\Gamma_\sigma^{\frac{1}{p}}(X + Y)\right)^{p-1}\right\|_1 + \left\|\Gamma_\sigma^{\frac{1}{p}}(Y)\left(\Gamma_\sigma^{\frac{1}{p}}(X + Y)\right)^{p-1}\right\|_1 \\
&\geq \left(\|X\|_{\mathbb{L}_p(\sigma)} + \|Y\|_{\mathbb{L}_p(\sigma)}\right)\|X + Y\|_{\mathbb{L}_p(\sigma)}^{p-1}.
\end{aligned}$$

$\square$

Now, given a full-rank state σ, a quantum Markov semigroup $(\mathcal{P}_t)_{t \geq 0}$ for which σ is invariant is called *reverse $\mathbb{L}_p(\sigma)$-contractive* if for any $p < 1$, all $t \geq 0$ and $X \in \mathcal{B}(\mathcal{H})^+$:

$$\|\mathcal{P}_t(X)\|_{\mathbb{L}_p(\sigma)} \geq \|X\|_{\mathbb{L}_p(\sigma)}.$$

We already saw in Proposition 5.2.1 that any quantum Markov semigroup is $\mathbb{L}_p(\sigma)$-contractive for any $p \geq 1$ whenever σ is an invariant state. Next theorem is an extension of this to the range $p < 1$:

Proposition 7.4.3. *(i) Any quantum Markov semigroup $(\mathcal{P}_t)_{t \geq 0}$ with invariant state σ is reverse $\mathbb{L}_p(\sigma)$-contractive for $p \in (-\infty, -1] \cup [1/2, 1)$.*

(ii) A doubly stochastic quantum Markov semigroup is reverse $\mathbb{L}_p(\mathbb{1}/\dim(\mathcal{H}))$-contractive for all $p < 1$.

Proof. (i) First let $p = -q \in (-\infty, -1]$, and $X > 0$. We note that

$$\|\mathcal{P}_t(X)\|_{\mathbb{L}_p(\sigma)} = \|\mathcal{P}_t(X)^{-1}\|_{\mathbb{L}_q(\sigma)}^{-1}.$$

On the other hand, $\mathcal{P}_t$ is completely positive and unital, and $z \mapsto z^{-1}$ is operator convex. Therefore, $\mathcal{P}_t(X^{-1}) \geq \mathcal{P}_t(X)^{-1}$ and $\|\mathcal{P}_t(X)^{-1}\|_{\mathbb{L}_q(\sigma)} \leq \|\mathcal{P}_t(X^{-1})\|_{\mathbb{L}_q(\sigma)}$. We conclude that

$$\|\mathcal{P}_t(X)\|_{\mathbb{L}_p(\sigma)} \geq \|\mathcal{P}_t(X^{-1})\|_{\mathbb{L}_q(\sigma)}^{-1} \geq \|X^{-1}\|_{\mathbb{L}_q(\sigma)}^{-1} = \|X\|_{\mathbb{L}_p(\sigma)},$$

where for the second inequality we use q-contractivity of $\mathcal{P}_t$ for $q \geq 1$. Now suppose that $p \in [1/2, 1)$. We note that $\hat{p} \in (-\infty, -1]$, and that the conjugate $\hat{\mathcal{P}}_t$ of $\mathcal{P}_t$ with respect to $\langle ., . \rangle_\sigma$ is reverse $\mathbb{L}_{\hat{p}}$-contractive.

Then using Hölder's duality, for $X > 0$ we have

$$
\begin{aligned}
\|\mathcal{P}_t(X)\|_{\mathbb{L}_p(\sigma)} &= \inf_{Y>0:\, \|Y\|_{\mathbb{L}_{\hat{p}}(\sigma)} \geq 1} |\langle Y, \mathcal{P}_t(X) \rangle_\sigma| \\
&= \inf_{Y>0:\, \|Y\|_{\mathbb{L}_{\hat{p}}(\sigma)} \geq 1} |\langle \hat{\mathcal{P}}_t(Y), X \rangle_\sigma| \\
&\geq \inf_{Z>0:\, \|Z\|_{\mathbb{L}_{\hat{p}}(\sigma)} \geq 1} |\langle Z, X \rangle_\sigma| \\
&= \|X\|_{\mathbb{L}_p(\sigma)}.
\end{aligned}
$$

Here the inequality follows from the reverse $\mathbb{L}_{\hat{p}}$-contractivity of $\hat{\mathcal{P}}_t$, i.e, $\|\hat{\mathcal{P}}_t(Y)\|_{\mathbb{L}_{\hat{p}}(\sigma)} \geq \|Y\|_{\mathbb{L}_{\hat{p}}(\sigma)} \geq 1$.

(ii) As worked out in Lemma 8 of [Cubitt et al., 2015] this is an immediate consequence of the operator Jensen inequality. $\qquad\square$

Remark 7.4.4. The (reverse) contractivity for the range of parameters $p \geq 1/2$ can be shown to hold from the data processing inequality of sandwiched p-Rényi relative entropy, which is known to hold [Frank and Lieb, 2013, Beigi, 2013, Müller-Lennert et al., 2013] for $p \geq 1/2$. Here, we gave a proof of part (i) for the range $p \in (-\infty, -1] \cup [1/2, 1)$ based on new ideas which may be of independent interest.

By (ii) and (iii) of Lemma 7.2.1, for any $p \in \mathbb{R} \backslash \{0, 1\}$ and $X > 0$,

$$
\mathrm{Ent}_{p,\sigma}(X) \geq 0.
$$

In the case when $p > 1$, we already saw that $\mathcal{E}_{p,\mathcal{L}}(X) \geq 0$ for any $X > 0$. This can be extended to the case $p \in \mathbb{R} \backslash \{0, 1\}$ so long as $(\mathcal{P}_t)_{t \geq 0}$ is reverse $\mathbb{L}_p(\sigma)$-contractive:

Proposition 7.4.5. *Given $p < 1$, $p \neq 0$, suppose that $\mathcal{L}$ is the generator of a reverse $\mathbb{L}_p(\sigma)$-contractive QMS $(\mathcal{P}_t)_{t \geq 0}$. Then for any $X > 0$,*

$$
\mathcal{E}_{p,\mathcal{L}}(X) \geq 0.
$$

Proof. From Lemma 5.2.8 (i), it is enough to consider the range $0 < p < 1$. Then, define

$$
g(t) \coloneqq \hat{p} \|\mathcal{P}_t(X)\|^p_{\mathbb{L}_p(\sigma)} - \hat{p} \|X\|^p_{\mathbb{L}_p(\sigma)}.
$$

By assumption of reverse $\mathbb{L}_p(\sigma)$-contractivity, for all $t \geq 0$ we have $g(t) \leq 0$. We note that $g(0) = 0$. Therefore, $g'(0) \leq 0$. We compute

$$
\begin{aligned}
0 \geq g'(0) \\
&= \frac{d}{dt} \hat{p} \|\mathcal{P}_t(X)\|^p_{p,\sigma} \Big|_{t=0} \\
&= \frac{d}{dt} \hat{p} \, \mathrm{Tr}\left(\Gamma_\sigma^{\frac{1}{p}} \circ \mathcal{P}_t(X)^p \right) \Big|_{t=0} \\
&= p\hat{p} \, \mathrm{Tr}\left(\Gamma_\sigma^{\frac{1}{p}} \circ \mathcal{L}(X) \cdot \Gamma_\sigma^{\frac{1}{p}}(X)^{p-1} \right) \\
&= p\hat{p} \, \mathrm{Tr}\left(\mathcal{L}(X) \cdot \Gamma_\sigma^{\frac{1}{p}} \left(\Gamma_\sigma^{\frac{1}{p}}(X)^{p-1} \right) \right) \\
&= p\hat{p} \, \langle I_{\hat{p},p}(X), \mathcal{L}(X) \rangle_\sigma \\
&= -4 \, \mathcal{E}_{p,\mathcal{L}}(X).
\end{aligned}
$$

$\qquad\square$

Extending the notion of a *p-logarithmic Sobolev constant* as defined in (Equation (7.3)) to the range $p \in \mathbb{R}\backslash\{0,1\}$, the following proposition is easily verified from the properties of the entropy function and of the Dirichlet form:

Proposition 7.4.6. *Let $(\mathcal{P}_t)_{t\geq 0}$ be a QMS on the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators over a finite dimensional Hilbert space $\mathcal{H}$. Then, for any $p \in \mathbb{R}\backslash\{0,1\}$ of Hölder conjugate $\hat{p}$,*

$$\alpha_p(\mathcal{L}) = \alpha_{\hat{p}}(\mathcal{L}).$$

Proof. Identifying X with $\mathrm{I}_{p,2}(Y)$ for some $Y \in \mathcal{B}(\mathcal{H})^+$, this is an immediate consequence of part (i) of Lemma 7.2.1 and part (i) of Lemma 5.2.8. $\qquad\qquad\square$

The above proposition allows us to restrict ourselves to log-Sobolev constants for values of $p \in [0,2]$. The proof of next proposition is standard [Olkiewicz and Zegarlinski, 1999, Kastoryano and Temme, 2013] and will be extended to weighted amalgamated norms in Chapter 8:

Proposition 7.4.7. *For a differentiable operator valued function $p \mapsto X_p$ we have, for any $p \in \mathbb{R}\backslash\{0\}$:*

$$\frac{d}{dp}\|X_p\|_{\mathbb{L}_p(\sigma)} = \frac{1}{p^2}\|X_p\|_{\mathbb{L}_p(\sigma)}^{1-p} \cdot \left(\frac{1}{2}\,\mathrm{Ent}_{p,\sigma}\left(\mathrm{I}_{p,p}(X_p)\right) + \frac{1}{2}\,\mathrm{Ent}_{p,\sigma}\left(\mathrm{I}_{p,p}(X_p^*)\right) + \gamma\right).$$

Here γ is given by

$$\gamma = \frac{p^2}{2}\left(\mathrm{Tr}\left[\Gamma_\sigma^{\frac{1}{p}}(Z_p^*)\cdot\Gamma_\sigma^{\frac{1}{p}}(X_p)\cdot\left|\Gamma_\sigma^{\frac{1}{p}}(X_p)\right|^{p-2}\right] + \mathrm{Tr}\left[\Gamma_\sigma^{\frac{1}{p}}(X_p^*)\cdot\Gamma_\sigma^{\frac{1}{p}}(Z_p)\cdot\left|\Gamma_\sigma^{\frac{1}{p}}(X_p)\right|^{p-2}\right]\right),$$

where $Z_p := \frac{d}{dp}X_p$.

Theorem 7.4.8. *Let $(\mathcal{P}_t)_{t\geq 0}$ be a primitive QMS and σ a full-rank invariant state. Then the QMS is reverse hypercontractive: Suppose that $\tilde{\alpha}_1 = \inf_{p\in[0,1]}\alpha_p(\mathcal{L}) > 0$. Then for $p \leq q < 1$ and $t \geq \frac{1}{4\tilde{\alpha}_1}\ln\frac{p-1}{q-1}$, and all $X > 0$:*

$$\|\mathcal{P}_t(X)\|_{\mathbb{L}_p(\sigma)} \geq \|X\|_{\mathbb{L}_q(\sigma)}. \tag{7.8}$$

Proof. It suffices to prove the theorem when $t = t(p) \equiv \frac{1}{4\tilde{\alpha}_1}\ln\frac{p-1}{q-1}$, for a given fixed q. Define

$$f(p) := \|\mathcal{P}_{t(p)}(X)\|_{\mathbb{L}_p(\sigma)} - \|X\|_{\mathbb{L}_q(\sigma)} = \|X_p\|_{\mathbb{L}_p(\sigma)} - \|X\|_{\mathbb{L}_q(\sigma)},$$

where $X_p := \mathcal{P}_{t(p)}(X) > 0$. To continue the proof we compute the derivative of $f(p)$ using Proposition 7.4.7.

$$f'(p) = \frac{d}{dp}\|X_p\|_{\mathbb{L}_p(\sigma)} = \frac{1}{p^2}\|X_p\|_{\mathbb{L}_p(\sigma)}^{1-p} \cdot \left(\mathrm{Ent}_{p,\sigma}(X_p) + p^2\,\mathrm{Tr}\left[\Gamma_\sigma^{\frac{1}{p}}(Z_p)\cdot\Gamma_\sigma^{\frac{1}{p}}(X_p)^{p-1}\right]\right),$$

where

$$Z_p = \frac{d}{dp}X_p = -t'(p)\mathcal{L}(X_p) = -\frac{1}{4\tilde{\alpha}_1(p-1)}\mathcal{L}(X_p).$$

Therefore,

$$f'(p) = \frac{1}{p^2}\|X_p\|_{\mathbb{L}_p(\sigma)}^{1-p} \cdot \left(\mathrm{Ent}_{p,\sigma}(X_p) - \frac{1}{\tilde{\alpha}_1}\mathcal{E}_{p,\mathcal{L}}(X_p)\right).$$

Now, assume that $q < 1$ and $\tilde{\alpha}_1 \leq \alpha_p(\mathcal{L})$ for all $p \in [0,1]$. Then for $p \leq q$ we have

$$\mathrm{Ent}_{p,\sigma}(X_p) \leq \frac{1}{\alpha_p(\mathcal{L})}\mathcal{E}_{p,\mathcal{L}}(X_p) \leq \frac{1}{\tilde{\alpha}_1}\mathcal{E}_{p,\mathcal{L}}(X_p),$$

where the second inequality holds since $p < 1$, so either p or its Hölder conjugate belongs to $[0, 1]$. Therefore, $f'(p) \leq 0$ for all $p \leq q < 1$, and since $f(q) = 0$, $f(p) \geq 0$ for all $p < q$.

$\qquad\square$

In a similar vein as in the setting of hypercontractivity, one gets a stronger result under the condition that $(\mathcal{P}_t)_{t \geq 0}$ satisfies σ-DBC, due to the Stroock Varopoulos inequality of Theorem 5.4.2. First, we get the reverse $\mathbb{L}_p(\sigma)$-contractivity for any value of $p < 1$:

Corollary 7.4.9. *Let the QMS $(\mathcal{P}_t)_{t \geq 0}$ satisfy σ-DBC. Then the followings hold:*

(i) *For all $p < 1$ and $X > 0$, $\mathcal{E}_{p,\mathcal{L}}(X) \geq 0$.*

(ii) *The associated QMS is reverse $\mathbb{L}_p(\sigma)$-contractive for all $p < 1$.*

As mentioned before, (reverse) $\mathbb{L}_p(\sigma)$-contractivity of $\mathcal{P}_t$ implies that Sandwiched $\mathbb{L}_p(\sigma)$-Rényi relative entropy is monotone under $\mathcal{P}_t$ [Frank and Lieb, 2013, Beigi, 2013, Müller-Lennert et al., 2013]. Therefore, when $\mathcal{P}_t$ comes from a QMS satisfying ω-DBC, the p-Rényi relative entropy is monotone under $\mathcal{P}_t$ not only for $p \geq 1/2$ but *all* values of p.

Proof. (i) By Theorem 5.4.2 (and part (i) of Lemma 5.2.8) for every p we have

$$\mathcal{E}_{p,\mathcal{L}}(\mathrm{I}_{p,2}(X)) \geq \mathcal{E}_{2,\mathcal{L}}(X).$$

On the other hand, by Proposition 7.4.5 and Proposition 5.22, we have $\mathcal{E}_{2,\mathcal{L}}(X) \geq 0$. Therefore, $\mathcal{E}_{p,\mathcal{L}}(\mathrm{I}_{p,2}(X)) \geq 0$.

(ii) Define $g(t)$ as in the proof of Proposition 7.4.5. By part (i) we have $g'(t) \leq 0$ for all $t \geq 0$ and $g(0) = 0$. Therefore, $g(t) \geq 0$ for all $t \geq 0$.

$\qquad\square$

The following corollary is an immediate consequence of the quantum Stroock-Varopoulos inequality as well as part (i) of Proposition 7.2.1:

Corollary 7.4.10. *Let $\mathcal{L}$ be the generator of a primitive QMS $(\mathcal{P}_t)_{t \geq 0}$ that satisfies σ-DBC. Then the QMS is reverse hypercontractive: For $p \leq q < 1$, $t \geq \frac{1}{4\alpha_1(\mathcal{L})} \ln \frac{p-1}{q-1}$ and all $X > 0$,*

$$\|\mathcal{P}_t(X)\|_{\mathbb{L}_p(\sigma)} \geq \|X\|_{\mathbb{L}_q(\sigma)}. \tag{7.9}$$

7.5. Quantum Nash inequality

7.5.1. Nash inequalities in finite dimensions

Nash inequalities were extended to the finite dimensional quantum setting in [Kastoryano and Temme, 2016] in the spirit of [Diaconis and Saloff-Coste, 1996b]. Here, we simply recall their main result showing the equivalence between a Nash inequality and ultracontractivity of the semigroup. Let $(\mathcal{P}_t = e^{t\mathcal{L}})_{t \geq 0}$ be a primitive QMS on $\mathcal{B}(\mathcal{H})$, $\mathcal{H}$ finite dimensional, with full-rank invariant state σ. The QMS is said to satisfy a *quantum Nash inequality* if there exist constants $C > 0$, $n \geq 2$ and $T \geq 0$ such that for any $X \in \mathcal{B}(\mathcal{H})$:

$$\|X\|_{\mathbb{L}_2(\sigma)}^{2+4/n} \leq C \left(\mathcal{E}_{2,\mathcal{L}}(X) + \frac{1}{T} \|X\|_{\mathbb{L}_2(\sigma)}^2 \right) \|X\|_{\mathbb{L}_1(\sigma)}^{4/n}. \tag{qNash(C,n,T)}$$

The following result was then proved in [Kastoryano and Temme, 2016]:

Theorem 7.5.1. *Assume that* $\mathrm{qNash}(C, n, T)$ *holds. Then, for all* $t \leq T$:

$$\|\mathcal{P}_t : \mathbb{L}_1(\sigma) \to \mathbb{L}_2(\sigma)\| \leq \mathrm{e}\left(\frac{nC}{4t}\right)^{n/4}.$$

On the other hand, assume that $(\mathcal{P}_t)_{t \geq 0}$ *is* KMS-*symmetric with respect to* σ. *Then* $\|\mathcal{P}_t : \mathbb{L}_1(\sigma) \to \mathbb{L}_2(\sigma)\| \leq \left(\frac{C_0}{4t}\right)^{n/4}$ *for any* $t \leq T$ *implies* $\mathrm{qNash}(C, n, T)$ *with* $C \equiv 2C_0(1 + 2/n)(1 + n/2)^{2/n}$.

7.5.2. Nash inequalities on the quantum phase space

As mentioned in Section 3.4, the original Nash inequality on $\mathbb{R}^n$ has the following form: for any smooth enough function $f : \mathbb{R}^n \to \mathbb{R}$ (e.g a Schwartz function),

$$\|f\|_{\mathbb{L}_2(\mathbb{R}^n)}^{2+4/n} \leq C_n \|f\|_{\mathbb{L}_1(\mathbb{R}^n)}^{4/n} \|\nabla f\|_{\mathbb{L}_2(\mathbb{R}^n)}^2, \tag{7.10}$$

where the sharp constant $C_n := \frac{2((n+2)/2)^{(n+2)/n}}{n \lambda(B^n)|B^n|^{2/n}}$ was found in [Carlen and Loss, 1993], with the constant $|B^n|$ denoting the volume of the unit ball B^n on $\mathbb{R}^n$, whereas $\lambda(B^n)$ is the spectral gap of the Laplacian on B^n. After a simple integration by parts, this inequality can be re-expressed in terms of the Dirichlet form of the heat semigroup:

$$\|f\|_{\mathbb{L}_2(\mathbb{R}^n)}^{2+4/n} \leq C_n \|f\|_{\mathbb{L}_1(\mathbb{R}^n)}^{4/n} \mathcal{E}(f), \qquad \mathcal{E}(f) := -\int_{\mathbb{R}^n} f \Delta f \, dx. \tag{7.11}$$

In this section, we prove that the quantum heat semigroup $(\mathcal{P}_t)_{t \geq 0}$ defined in Section 5.5.2 also satisfies a Nash inequality, at least on Schwartz operators of positive Wigner function. In the following we denote $\rho_t = \mathcal{P}_t^{\mathrm{qheat}}(\rho) = \mathcal{P}_{t*}^{\mathrm{qheat}}(\rho)$.

Theorem 7.5.2 (A non-commutative Nash inequality). *If* ρ *is a Schwartz state of positive Wigner function, then* $\rho \in \mathrm{dom}(\mathcal{L}^{\mathrm{qheat}})$ *and the following non-commutative Nash inequality holds:*

$$\|\rho\|_2^{2+2/n} \leq C_{2n} \, \mathcal{E}_{2, \mathcal{L}^{\mathrm{qheat}}}(\rho). \tag{7.12}$$

Remark 7.5.3. Upon replacing $\mathbb{L}_p(\mathbb{R}^n)$ norms by Schatten norms and n by $2n$, the similarity between (7.12) and (7.11) is obvious and unsurprising in view of Parseval's relation 0.2.3. The doubling of the dimension is also expected since the quantum heat semigroup is defined on a representation of a $2n$-dimensional phase space.

Proof. We prove that (7.12) is satisfied for any Schwartz state ρ with positive Wigner function. For such states $\mathcal{L}^{\mathrm{qheat}}(\rho)$ is well-defined and trace class (see e.g. Propositions 3.14 and 3.15 of [Keyl et al., 2016]), and using the non-commutative Parseval relation, Theorem 0.2.3, and the Dirichlet form $\mathcal{E}_{2, \mathcal{L}^{\mathrm{qheat}}}$ takes the following form:

$$\mathcal{E}_{2, \mathcal{L}^{\mathrm{qheat}}}(\rho) := -\operatorname{Tr}(\rho \mathcal{L}^{\mathrm{qheat}}(\rho)) = -(2\pi)^{-n} \int_{\mathcal{Z}} \overline{\mathcal{F}_\rho^{\mathrm{q}}(z)} \mathcal{F}_{\mathcal{L}^{\mathrm{qheat}}(\rho)}^{\mathrm{q}}(z) dz. \tag{7.13}$$

Now, using Equation (0.29), we find that $\mathcal{F}_{\mathcal{L}^{\mathrm{qheat}}(\rho)}^{\mathrm{q}}(z) = -\frac{1}{4}|z|^2 \mathcal{F}_\rho^{\mathrm{q}}(z)$. Substituting this into the right hand side of Equation (7.13), we get,

$$\mathcal{E}_{2, \mathcal{L}^{\mathrm{qheat}}}(\rho) = \frac{1}{4}(2\pi)^{-n} \int_{\mathcal{Z}} |z|^2 |\mathcal{F}_\rho^{\mathrm{q}}(z)|^2 dz. \tag{7.14}$$

Now, $z \mapsto \mathcal{F}_\rho^{\mathrm{q}}(z)$ is a Schwartz function, and therefore its inverse Fourier transform is also a Schwartz

function. Let us denote by $\mathcal{F}_f$ the Fourier transform of an integrable function f on $\mathbb{R}^{2n}$:

$$\mathcal{F}_f(z) := (2\pi)^{-n} \int_{\mathbb{R}^{2n}} f(x)\, e^{i(x,z)}\, dx\,. \tag{7.15}$$

We make use of the following basic facts. Firstly, the Fourier transform satisfies the following useful identity

$$\mathcal{F}_{x \mapsto \partial_{x_j} f(x)}(z) = -iz_j \mathcal{F}_f(z), \tag{7.16}$$

for any integrable, continuously differentiable function f, which has an integrable partial derivative $\partial_{x_j} f$, where z_j denotes the jth component of the vector $z \in \mathcal{Z}$. Secondly, for any square integrable function h on $\mathbb{R}^{2n}$, we have that

$$\|\mathcal{F}_h\|_2 = \|h\|_2, \tag{7.17}$$

which is the classical Plancherel identity. We also employ the well-known polarization identity:

$$\int_{\mathbb{R}^{2n}} \overline{g(x)}\, h(x) dx = \frac{1}{4} \left(\|g+h\|_2^2 - \|g-h\|_2^2 + i\|g+ih\|_2^2 - i\|g-ih\|_2^2 \right), \tag{7.18}$$

for any two square integrable functions g and h, and the fact that the characteristic function $\mathcal{F}_\rho^{\mathrm{q}}$ is equal to $\mathcal{F}_{f_\rho}$, where f_ρ denotes the Wigner function of ρ and is a Schwartz function.

$$
\begin{aligned}
\mathcal{E}_{2,\mathcal{L}^{\mathrm{qheat}}}(\rho) &= \frac{1}{4}(2\pi)^{-n} \int_{\mathcal{Z}} \overline{\mathcal{F}_\rho^{\mathrm{q}}(z)} \left(|z|^2 \mathcal{F}_\rho^{\mathrm{q}}(z) \right) dz \\
&= -\sum_{i=1}^{2n} \frac{1}{4}(2\pi)^{-n} \int_{\mathcal{Z}} \overline{\mathcal{F}_{f_\rho}(z)} (-iz_j)^2 \mathcal{F}_{f_\rho}(z) dz \\
&= -\sum_{i=1}^{2n} \frac{1}{4}(2\pi)^{-n} \int_{\mathcal{Z}} \overline{\mathcal{F}_{f_\rho}(z)} \mathcal{F}_{x \mapsto \frac{\partial^2}{\partial x_i^2} f_\rho(x)}(z) dz \\
&= -\frac{1}{4} \sum_{i=1}^{2n} (2\pi)^{-n} \int_{\mathcal{Z}} f_\rho(x) \partial_{x_i^2}^2 f_\rho(x) dx \\
&= -(2\pi)^{-n} \int_{\mathcal{Z}} f_\rho(x) \frac{1}{4} \Delta f_\rho(x) dx \equiv \frac{1}{4(2\pi)^n} \mathcal{E}(f_\rho),
\end{aligned}
\tag{7.19}
$$

where $\mathcal{E}(.)$ is the Dirichlet form of the classical heat semigroup as given in Equation (7.11). The third line follows from two uses of Equation (7.16) and the fourth line follows from Equation (7.17) and Equation (7.18).

Moreover, by the non-commutative Parseval relation, Theorem 0.2.3, and Equation (7.17), we have

$$\|\rho\|_2^2 = \mathrm{Tr}(\rho^2) = (2\pi)^{-n} \int_{\mathcal{Z}} |\mathcal{F}_\rho^{\mathrm{q}}(z)|^2 dz = (2\pi)^{-n} \|\mathcal{F}_\rho^{\mathrm{q}}\|_2^2 = (2\pi)^{-n} \|f_\rho\|_{\mathbb{L}_2(\mathbb{R}^n)}^2. \tag{7.20}$$

Further,

$$(2\pi)^n \|\rho\|_1 = (2\pi)^n = \|f_\rho\|_{\mathbb{L}_1(\mathbb{R}^n)}, \tag{7.21}$$

which follows from the fact that the Wigner function of a state (as defined through Equation (0.35)) has integral equal to $(2\pi)^n$, and the assumption that f_ρ is positive. Hence,

$$\|\rho\|_2^{2+2/n} = (2\pi)^{-(n+1)} \|f_\rho\|_{\mathbb{L}_2(\mathbb{R}^n)}^{2+2/n} \le (2\pi)^{-n+1} C_{2n} \mathcal{E}(f_\rho) = 8\pi\, C_{2n}\, \mathcal{E}_{2,\mathcal{L}^{\mathrm{qheat}}}(\rho), \tag{7.22}$$

where we made use of the classical Nash inequality (7.12), with n replaced by $2n$. □

As expected, (7.12) easily implies ultracontractivity of $(\mathcal{P}_t)_{t \geq 0}$.

Theorem 7.5.4 (Ultracontractivity of the quantum heat semigroup). *If ρ is a Schwartz state of positive Wigner function, then there exists a positive constant κ_n such that for any $t > 0$*

$$\|\rho_t\|_\infty \leq \|\rho_t\|_2 \leq \left(\frac{nC_{2n}}{2t} \right)^{\frac{n}{2}} . \tag{7.23}$$

Moreover, for any such initial state ρ evolving under the action of the semigroup $(\mathcal{P}_t^{\mathrm{qheat}})_{t \geq 0}$, the following bounds give the rate of decay of its purity and the rate of increase of its von Neumann entropy, respectively:

$$\mathrm{Tr}\,\rho_t^2 \equiv \|\rho_t\|_2^2 \leq \left(\frac{nC_{2n}}{2t} \right)^n ; \tag{7.24}$$

$$S(\rho_t) \geq \frac{n}{2} \ln \left(\frac{2t}{n\,C_{2n}} \right) . \tag{7.25}$$

Proof. The proof is identical to the one given in [Kastoryano and Temme, 2016] in the case of finite dimensional QMS, which in turn closely follows the proof in the classical case studied in [Diaconis and Saloff-Coste, 1996b]. Let ρ be a Schwartz state with a positive Wigner function. By Proposition 5.5.3, $\rho_t := \mathcal{P}_t(\rho)$ is also such a state, and Theorem 7.5.2 applies. Let $u(t) := \|\rho_t\|_2^2$; using Theorem 0.2.3 and Equation (5.47), one can verify that the function u is differentiable and $\dot{u}(t) = -2\,\mathcal{E}_{2,\mathcal{L}^{\mathrm{qheat}}}(\rho_t)$. Theorem 7.5.2 implies that

$$u^{1+1/n}(t) = \|\rho_t\|_2^{2+2/n} \leq C_{2n}\,\mathcal{E}_{2,\mathcal{L}^{\mathrm{qheat}}}(\rho_t) = -\frac{C_{2n}}{2} \dot{u}(t)$$

so that

$$\frac{d}{dt} \frac{1}{u^{1/n}(t)} \frac{nC_{2n}}{2} = -\frac{C_{2n}}{2} \frac{\dot{u}(t)}{u(t)^{1+1/n}} \geq 1,$$

and by integrating both sides of the above inequality from 0 to t, one gets:

$$\frac{1}{u^{1/n}(t)} \geq \frac{2t}{nC_{2n}} + \frac{1}{u^{1/n}(0)} \geq \frac{2t}{nC_{2n}} ,$$

so that

$$\|\rho_t\|_2^2 \equiv u(t) \leq \left(\frac{nC_{2n}}{2t} \right)^n .$$

In particular, this implies that

$$\|\rho_t\|_\infty \leq \|\rho_t\|_2 \leq \left(\frac{nC_{2n}}{2t} \right)^{\frac{n}{2}} .$$

Therefore,

$$S(\rho_t) = -\mathrm{Tr}(\rho_t \ln \rho_t) \geq -\ln \|\rho_t\|_\infty \geq \frac{n}{2} \ln \left(\frac{2t}{n\,C_{2n}} \right) .$$

□

Chapter 8.

Non primitive functional inequalities for the study of decoherence

Largely inspired by [Diaconis and Saloff-Coste, 1996a], our goal in this chapter is to develop the theory of LSI and HC for non-primitive QMS and its use in proving rapid decoherence. This line of research started with [Bardet, 2017] who introduced non-primitive generalizations of the Poincaré and the modified logarithmic Sobolev inequality. Note that, due to the non-commutativity of quantum systems, typical quantum features arise in this situation that are absent from the classical theory.

Of course, the hypercontractive property depends highly on the choice of the interpolating family of $\mathbb{L}_p$ norms. In particular, a QMS which is hypercontractive for Kosaki's $\mathbb{L}_p$ norms will be primitive. The main contribution of the present work is to study hypercontractivity with respect to a generalisation of Kosaki's norms, called the amalgamated norms and defined by [Junge and Parcet, 2010]. We refer to Section 1.1.1 for a detailed analysis of amalgamated $\mathbb{L}_p$ norms. Using these norms, we will be able to reproduce the steps of [Diaconis and Saloff-Coste, 1996a] for non-primitive QMS.

Throughout this chapter, we consider quantum Markov semigroups defined on the algebra $\mathcal{B}(\mathcal{H})$ of linear operators acting on a finite dimensional Hilbert space $\mathcal{H}$. Following ideas from [Beigi and King, 2016], we derive a formula for the differential of the amalgamated $\mathbb{L}_p$ norms (see Theorem 8.3.3), with respect to the index p. This leads to the definition of the *weak decoherence-free logarithmic Sobolev inequality* (DF-wLSI) and the *weak decoherence-free hypercontractivity* (DF-wHC), and allows us to extend Gross' integration lemma to this setting (see Theorem 8.2.2). A first difference compared to the primitive case is that LSI implies HC but with a larger weak constant which depends on the structure of the DF-algebra.

In the primitive case, the uniform convexity of the $\mathbb{L}_p$ norms was used in [Olkiewicz and Zegarlinski, 1999] to show that wLSI together with PI give rise to a strong logarithmic Sobolev inequality (sLSI) (cf. Proposition 3.2.6). We show that a similar analysis can be performed in our extended framework, in order to derive universal upper bounds on the log-Sobolev constants (see Definition 8.2.1 and Corollary 8.2.6). We also prove that, except in the primitive case, the strong LSI does not hold and therefore neither does the related notion of strong hypercontractivity. This implies that the uniform convexity no longer holds for the amalgamated $\mathbb{L}_p$-norms.

Layout of the chapter: In Section 8.1, we sketch a succinct overview of the framework of [Bardet, 2017] and relate his decoherence-free Poincaré inequality for a quantum Markov semigroup to the spectral gap of a corresponding discrete-time quantum Markov chain. Next, we define and study the notions of a non-primitive logarithmic Sobolev inequality and hypercontractivity: our main results are stated in Section 8.2. The notions of decoherence-free log-Sobolev inequality and hypercontractivity

are studied in Section 8.4, where we prove Gross' integration Lemma as well as a universal upper bound on the constants. In Section 8.5 we prove that the strong LSI fails for non-trivially primitive QMS. Some applications of our framework to the derivation of decoherence rates are provided in Section 8.6.

8.1. Decoherence-free spectral gap and MLSI

Assume given a faithful QMS $(\mathcal{P}_t)_{t\geq0}$, that is one possessing a full-rank invariant state. We are in the setting of Section 6.2: the semigroup converges to its decoherence-free algebra $\mathcal{N}$ which we may decompose as in Equation (6.5). Moreover, there exists a conditional expectation $E_\mathcal{N}$ onto $\mathcal{N}$ and the state $\sigma_{\mathrm{Tr}} := E_{\mathcal{N}*}(d_\mathcal{H}^{-1}\mathbb{1})$ is left invariant under the evolution induced by $(\mathcal{P}_t)_{t\geq0}$. We are interested in estimating the speed of convergence to the following limit by means of non-commutative functional inequalities:

$$\mathcal{P}_t(X) \underset{t\to\infty}{\sim} \mathcal{P}_t \circ E_\mathcal{N}[X].$$
(8.1)

We first briefly review the results of [Bardet, 2017], where the Poincaré and modified logarithmic Sobolev inequalities were extended to the present non-primitive scenario.

Decoherence-free Poincaré inequality : The Poincaré inequality can be simply modified in order to take into account $\mathbb{L}_2$ convergence of non-primitive evolutions: we say that a DF-*Poincaré inequality* holds if for all $X \in \mathcal{B}_{sa}(\mathcal{H})$:

$$\lambda \operatorname{Var}_\mathcal{N}(X) \leq \mathcal{E}_{2,\mathcal{L}}(X).$$
(PI$_\mathcal{N}(\lambda)$)

where $\operatorname{Var}_\mathcal{N}(X) := \|X - E_\mathcal{N}[X]\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}^2$ is the DF-*variance* of X. Then, define the *decoherence-free Poincaré constant* as follows:

$$\lambda(\mathcal{L}) := \inf_{X \in \mathcal{B}_{sa}(\mathcal{H})} \frac{\mathcal{E}_{2,\mathcal{L}}(X)}{\operatorname{Var}_\mathcal{N}(X)}.$$
(8.2)

The Poincaré constant turns out to be the spectral gap of the operator $\frac{\mathcal{L}+\hat{\mathcal{L}}}{2}$, where $\hat{\mathcal{L}}$ is the adjoint of $\mathcal{L}$ with respect to $\langle.,.\rangle_{\sigma_{\mathrm{Tr}}}$, that is minus its second (negative) largest eigenvalue. In complete analogy with Proposition 7.1.1, we obtain the following $\mathbb{L}_2(\sigma_{\mathrm{Tr}})$ exponential decay:

$$\|\mathcal{P}_t(X - E_\mathcal{N}(X))\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})} \leq \mathrm{e}^{-\lambda(\mathcal{L})t} \|X - E_\mathcal{N}[X]\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}.$$
(8.3)

From this, bounds on the convergence in trace distance of the corresponding evolution of quantum states can be derived using

$$\|\mathcal{P}_{t*} \circ (\rho - E_{\mathcal{N}*}(\rho))\|_1 = \|\hat{\mathcal{P}}_t(X) - \hat{\mathcal{P}}_t \circ E_\mathcal{N}[X]\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} \leq \|\hat{\mathcal{P}}_t(X) - \hat{\mathcal{P}}_t \circ E_\mathcal{N}[X]\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})},$$
(8.4)

where $(\hat{\mathcal{P}}_t)_{t\geq0}$ denotes the semigroup associated with $\hat{\mathcal{L}}$.

In the case of a quantum Markov chain $\{\Phi^n\}_{n\in\mathbb{N}}$, we define the *absolute spectral gap* $\lambda(\Phi)$ as $1 - \max_{|\lambda_k|\neq1}|\lambda_k|$, where the maximization is taken over all the eigenvalues of Φ of absolute value less than 1. Convergence bounds in terms of the absolute spectral gap, as well as the Jordan structure of the map Φ were found in [Szehr et al., 2015].

Decoherence-free MLSI As expected, one can derive better bounds on the convergence (8.1) using more powerful methods. In [Bardet, 2017], the modified logarithmic Sobolev constant $\alpha_1(\mathcal{L})$ was defined as follows:

$$\alpha_1(\mathcal{L}) := \inf_{\rho \in \mathcal{D}(\mathcal{H})} \frac{\mathrm{EP}_{\sigma_{\mathrm{Tr}}}(\rho)}{4\, D(\rho \| E_{\mathcal{N}_*}(\rho))} \,.$$

The following stronger exponential decay in relative entropy is a straightforward consequence of the above definition: denoting $\rho_t := \mathcal{P}_{*t}(\rho)$ and $\rho_{\mathcal{N}} := E_{\mathcal{N}}(\rho)$:

$$D(\rho_t \| \mathcal{P}_{*t}(\rho_{\mathcal{N}})) \leq \mathrm{e}^{-4\alpha_1(\mathcal{L})t}\, D(\rho \| \rho_{\mathcal{N}}) \,.$$

As opposed to the primitive case, their exists no lower bound on the constant $\alpha_1(\mathcal{L})$ for a generic evolution for finite dimensional systems to this day. This is always the case for primitive evolutions satisfying a detailed balance condition with respect to their unique invariant state. This is due to the fact that, in this case, the constant $\alpha_1(\mathcal{L})$ can be lower bounded by $\alpha_2(\mathcal{L})$ using regularity of the Dirichet forms. On the other hand, general lower bounds on $\alpha_2(\mathcal{L})$ were found in [Temme et al., 2014] by standard interpolation techniques.

This constitutes a good reason for introducing a decoherence-free analogue of $\alpha_2(\mathcal{L})$ in the context of non-primitive evolutions. As it turns out, we will show in Section 8.5 that the constant $\alpha_2(\mathcal{L})$ is positive only in the case of a primitive QMS. Therefore, the problem of finding a lower bound to $\alpha_1(\mathcal{L})$ for non-primitive semigroups remains unanswered. Fortunately, the introduction and study of a weak decoherence-free logarithmic Sobolev inequality, which is the main topic of this chapter, remains interesting since it will allow us to derive similar decoherence times via hypercontractivity.

8.2. Decoherence-free hypercontractivity: main results

By Corollary 6 of [Temme et al., 2014], in finite dimensions, hypercontractivity of a QMS with respect to the $\mathbb{L}_p(\sigma_{\mathrm{Tr}})$ norms is equivalent to the primitivity of the QMS. In order to deal with non-primitive QMS, a possible choice of norms are the so-called *amalgamated norms* introduced in [Junge and Parcet, 2010] (cf. Section 1.1.1). We recall that these norms are defined as follows: for $1 \leq q \leq p \leq +\infty$ and $\frac{1}{r} = \frac{1}{q} - \frac{1}{p}$, define

$$\|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} := \inf_{\substack{A,B \in \mathcal{N}, Y \in \mathcal{B}(\mathcal{H}) \\ X = AYB}} \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|Y\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})} \,, \tag{8.5}$$

$$\|Y\|_{\mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}}))} := \sup_{A,B \in \mathcal{N}} \frac{\|AYB\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}{\|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})}} \,. \tag{8.6}$$

We shall prove that they are particularly well-suited to study the hypercontractivity of the QMS, since:

(i) they reduce to the $\mathbb{L}_p(\sigma)$ norms when the QMS is primitive with unique invariant state σ;

(ii) they reduce to the $\mathbb{L}_q(\sigma_{\mathrm{Tr}})$ norms when evaluated on $\mathcal{N}$;

(iii) the QMS is contractive with respect to these norms for all $p, q \geq 1$.

When differentiating an amalgamated norm with respect to p, some natural quantities will appear that we will connect with entropic ones in Lemma 8.3.2. Similarly to [Olkiewicz and Zegarlinski, 1999], we

thus introduce a decoherence-free generalisation of the $\mathbb{L}_p$ relative entropies as follows: define the map

$$S_p(X) = -p\,\partial_s\,\mathrm{I}_{p+s,p}(X)|_{s=0}\,,$$

referred to as *operator valued relative entropy*, where $\mathrm{I}_{q,p}$ is defined in Equation (1.12)[1]. It can be computed explicitly: when $X \geq 0$,

$$S_p(X) = \Gamma_{\sigma_{\mathrm{Tr}}}^{-\frac{1}{p}}[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X)\ln\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X)] - \frac{1}{2p}\{X, \ln\sigma_{\mathrm{Tr}}\}\,.$$

We then define the DF-$\mathbb{L}_p$ *relative entropy* associated with the algebra $\mathcal{N} \equiv \mathcal{N}(\mathcal{P})$ as follows: for $\frac{1}{p} + \frac{1}{\hat{p}} = 1$,

$$\mathrm{Ent}_{p,\mathcal{N}}(X) := p\,\langle\mathrm{I}_{\hat{p},p}(X), S_p(X)\rangle_{\sigma_{\mathrm{Tr}}} - \mathrm{Tr}\left[(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X))^p \ln E_{\mathcal{N}}[\Gamma_{\sigma_{\mathrm{Tr}}}^{-1}(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X))^p]\right]. \tag{8.7}$$

In the case of a primitive QMS where σ_{Tr} is the unique invariant state of the evolution, $E_{\mathcal{N}}[.] := \mathrm{Tr}(\sigma_{\mathrm{Tr}}\,.)\,\mathbb{1}$ and we get back the original definition of [Kastoryano and Temme, 2013], which we denoted by $\mathrm{Ent}_{p,\sigma_{\mathrm{Tr}}}(X)$ in Equation (7.2). In general, $\mathrm{Ent}_{p,\mathcal{N}}(X) \leq \mathrm{Ent}_{p,\sigma_{\mathrm{Tr}}}(X)$. In the important cases $p = 1$ and $p = 2$, Equation (8.7) reduces to

$$\mathrm{Ent}_{1,\mathcal{N}}(X) := \mathrm{Tr}\left[\Gamma_{\sigma_{\mathrm{Tr}}}(X)\left(\ln\frac{\Gamma_{\sigma_{\mathrm{Tr}}}(X)}{\mathrm{Tr}(\Gamma_{\sigma_{\mathrm{Tr}}}(X))} - \ln\sigma_{\mathrm{Tr}}\right)\right] - \mathrm{Tr}\left[\Gamma_{\sigma_{\mathrm{Tr}}}(X)\ln\frac{E_{\mathcal{N}}[X]}{\mathrm{Tr}(\Gamma_{\sigma_{\mathrm{Tr}}}(X))}\right],$$

$$\mathrm{Ent}_{2,\mathcal{N}}(X) := \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(X)\right]^2\left(\ln\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(X)\right]^2 - \ln E_{\mathcal{N}}\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{-1}\left(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(X)\right)^2\right] - \ln\sigma_{\mathrm{Tr}}\right)\right).$$

$$\tag{8.8}$$

Definition 8.2.1. We say that the QMS $(\mathcal{P}_t)_{t\geq0}$ of generator $\mathcal{L}$

(i) satisfies a *weak DF-q-log-Sobolev inequality* with positive *strong DF-q-log-Sobolev constant* $c > 0$ and *weak DF-q-log-Sobolev constant* $d \geq 0$ if for all $X > 0$,

$$\mathrm{Ent}_{q,\mathcal{N}}(X) \leq c\,\mathcal{E}_{q,\mathcal{L}}(X) + 2d\,\|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q\,; \tag{$\mathrm{LSI}_{q,\mathcal{N}}(c,d)$}$$

(ii) is *weakly q-DF-hypercontractive* for positive constants $c > 0$ and $d \geq 0$ if for any function $p : [0, +\infty) \to \mathbb{R}$ such that for any $t \geq 0$, $q \leq p(t) \leq 1 + (q-1)\,\mathrm{e}^{\frac{4t}{c}}$,

$$\|\mathcal{P}_t : \mathbb{L}_q(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))\| \leq \exp\left\{2d\left(\frac{1}{q} - \frac{1}{p(t)}\right)\right\}. \tag{$\mathrm{HC}_{q,\mathcal{N}}(c,d)$}$$

The first main result of this chapter is the following generalisation of Gross' integration lemma that establishes the equivalence between hypercontractivity and the log-Sobolev inequality for a faithful QMS:

Theorem 8.2.2. *Let $(\mathcal{P}_t)_{t\geq0}$ be a faithful QMS on $\mathcal{B}(\mathcal{H})$ and let $q \geq 1$, $c > 0$ and $d \geq 0$.*

(i) If $\mathrm{HC}_{q,\mathcal{N}}(c,d)$ holds, then $\mathrm{LSI}_{q,\mathcal{N}}(c,d)$ holds.

(ii) If $\mathrm{LSI}_{\tilde{q},\mathcal{N}}(c,d)$ holds for all $\tilde{q} \geq q$, then $\mathrm{HC}_{q,\mathcal{N}}(c, d + \ln\sqrt{|I|})$ holds, where $|I|$ denotes the number of blocks of $\mathcal{N}$ in (0.9).

This theorem is quite surprising compared to the (classical and quantum) primitive case, where there is an exact equivalence between hypercontractivity and the logarithmic Sobolev inequality (i.e.

---

[1]Here, the operators $\mathrm{I}_{q,p}$ are defined with respect to the reference state σ_{Tr}.

with the same constant). In general, the term $\ln\sqrt{|I|}$ appearing in the theorem is not optimal (see end of Section 8.4.1).

The case where $\mathcal{N}$ is a factor and where the QMS is unital and trace-preserving was proved in [Beigi and King, 2016], but only in the case $d = 0$. However, the authors failed to give an example where the constant c is finite. We shall actually prove in Section 8.5 that this is impossible. More generally, we prove that, as soon as the QMS is truly non-primitive and non-invertible (that is, not a unitary evolution), necessarily $c < +\infty$ implies $d > 0$.

Remark also that the last statement is weaker than in the classical case, when one only needs to assume that the weak LSI holds for $\tilde{q} = q$. We recall that this is due to the fact that the regularity conditions seen in Lemma 2.4.2 always hold in the commutative setting, which ensures that $\mathrm{LSI}_2(c, d)$ implies $\mathrm{LSI}_{q,\mathcal{N}}(2c, d)$ for all $q \geq 2$, and even $\mathrm{LSI}_{q,\mathcal{N}}(c, d)$ in the reversible case. This condition needs to be assumed in the general quantum setting, even in the primitive case. A generator $\mathcal{L}$ of a QMS $(\mathcal{P}_t)_{t\geq 0}$ is said to be *strongly $\mathbb{L}_p$-regular* if there exists $d_0 \geq 0$ such that for all $p \geq 1$ and all $X \in \mathcal{B}_{sa}(\mathcal{H})$,

$$d_0 \|X\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})}^p + \mathcal{E}_{p,\mathcal{L}}(X) \geq \mathcal{E}_{2,\mathcal{L}}(\mathrm{I}_{2,p}(X)). \qquad (\text{s-}\mathbb{L}_p(d_0))$$

With these definitions, we can prove the following theorem.

Theorem 8.2.3. *Assume that $\mathrm{LSI}_{2,\mathcal{N}}(c, d)$ holds. If the generator $\mathcal{L}$ is strongly $\mathbb{L}_p$-regular for some $d_0 \geq 0$, then $\mathrm{LSI}_{q,\mathcal{N}}(c, d + c\, d_0/2)$ holds for all $q \geq 1$, so that $\mathrm{HC}_{2,\mathcal{N}}(c, d + \ln\sqrt{|I|} + c\, d_0/2)$ holds.*

The last two theorems generalise Theorem 3.8 of [Olkiewicz and Zegarlinski, 1999] as well as Theorem 15 of [Kastoryano and Temme, 2013]. Moreover, it was conjectured in [Kastoryano and Temme, 2013] that KMS-symmetric QMS are strongly $\mathbb{L}_p$-regular, again with $d_0 = 0$. This was shown to hold in [Bardet, 2017] under the condition of σ_{Tr}-DBC and without the primitive assumption (cf. Section 5.4). For a KMS-symmetric QMS, a straightforward extension of the proof of Proposition 5.2 of [Olkiewicz and Zegarlinski, 1999] implies that the strong regularity of $\mathcal{L}$ always holds, with $d_0 = \|\mathcal{L} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\sigma_{\mathrm{Tr}})\| + 1$. These remarks motivate the following corollary of Theorem 8.2.3:

Corollary 8.2.4. *Assume that $\mathrm{LSI}_{2,\mathcal{N}}(c, d)$ holds. Then:*

(i) *If $(\mathcal{P}_t)_{t\geq 0}$ is KMS-symmetric, then $\mathrm{HC}_{2,\mathcal{N}}(c, d + \ln\sqrt{|I|} + c\,(\|\mathcal{L} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\sigma_{\mathrm{Tr}})\| + 1))$ holds.*

(ii) *If $\mathcal{L}$ satisfies σ_{Tr}-DBC, then $\mathrm{HC}_{2,\mathcal{N}}(c, d + \ln\sqrt{|I|})$ holds.*

We also prove that it is always possible to get a weak DF-2-log-Sobolev inequality with a universal weak DF-2-log-Sobolev constant from any weak DF-2-log-Sobolev inequality, hence extending Theorem 4.2 of [Olkiewicz and Zegarlinski, 1999] to the non-primitive case.

Theorem 8.2.5. *Assume that $\mathrm{LSI}_{2,\mathcal{N}}(c, d)$ holds and denote by $\lambda(\mathcal{L})$ the spectral gap of $(\mathcal{L} + \hat{\mathcal{L}})/2$. Then $\mathrm{LSI}_{2,\mathcal{N}}(c + 2\frac{d+1}{\lambda(\mathcal{L})}, d' = \ln\sqrt{2})$ holds.*

Finally, using the DF-hypercontractivity and complex interpolation methods, we derive the following universal DF-2-log-Sobolev constants:

Corollary 8.2.6. *Given a KMS-symmetric QMS $(\mathcal{P}_t)_{t\geq 0}$ with spectral gap $\lambda(\mathcal{L})$, $\mathrm{LSI}_{2,\mathcal{N}}(c, \ln\sqrt{2})$ holds, with*

$$c \leq \frac{\ln(\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty) + 2}{\lambda(\mathcal{L})}.$$

8.2.1. Application to decoherence rates

Given a QMS $(\mathcal{P}_t)_{t\geq 0}$, its *decoherence time* is defined as:

$$\tau_{deco}(\varepsilon) := \inf\left\{t \geq 0:\ \|\mathcal{P}_{*t}\left(\rho - E_{\mathcal{N}*}(\rho)\right)\|_1 \leq \varepsilon,\ \forall \rho \in \mathcal{D}(\mathcal{H})\right\}.$$

The standard method to obtain estimates for $\tau(\varepsilon)$ in the primitive case is to use Pinsker's inequality in order to upper bound the trace distance in terms of the relative entropy, which decays exponentially fast with decay rate provided by $\alpha_1(\mathcal{L})$ [Diaconis and Saloff-Coste, 1996a, Kastoryano and Temme, 2013] (cf. Chapters 3 and 7). An estimate on the constant $\alpha_1(\mathcal{L})$ can then be found in terms of $\alpha_2(\mathcal{L})$, under the condition that the corresponding weak LSI$_2$ constant is null. However we prove in Section 8.5 that this condition holds only for primitive and unitary evolution. In all the other cases when there is only access to a weak DF-log-Sobolev inequality, we can fortunately still derive bounds on the decoherence times by extending a technique already used in the classical case in [Zegarlinski, 1995, Diaconis and Saloff-Coste, 1996a], by combining Poincaré's inequality and the weak DF-hypercontractivity property of the semigroup (cf. Section 3.2).

Proposition 8.2.7. *Assume that a QMS $(\mathcal{P}_t)_{t\geq 0}$ satisfies $\mathrm{HC}_{2,\mathcal{N}}(c,d)$, and that $\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty \geq \mathrm{e}$. Then, given $t = \frac{c}{4} \ln\ln \|\sigma_{\mathrm{Tr}}^{-1}\|_\infty + \frac{\kappa}{\lambda(\mathcal{L})}$, $\kappa > 0$:*

$$\forall \rho \in \mathcal{D}(\mathcal{H}), \qquad \|\mathcal{P}_{*t}\left(\rho - E_{\mathcal{N}*}(\rho)\right)\|_1 \leq \max_{i\in I} \sqrt{d_{\mathcal{H}_i}}\, \mathrm{e}^{1+d-\kappa}, \tag{8.9}$$

where the coefficients $d_{\mathcal{H}_i}$ are the dimensions of the spaces $\mathcal{H}_i$ arising in the decomposition of $\mathcal{N}$ given by (0.9). The above inequality provides the following bound on the decoherence time of the QMS:

$$\tau_{\mathrm{deco}}(\varepsilon) \leq \frac{\ln\left(\max_{i\in I} \sqrt{d_{\mathcal{H}_i}}\, \varepsilon^{-1}\right) + 1 + d}{\lambda(\mathcal{L})} + \frac{c}{4}\ln\ln\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty.$$

Remark that the assumption on $\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty$ is not restrictive: it means that the lowest eigenvalue of σ_{Tr} has to be smaller than $1/\mathrm{e}$. In particular, it always holds when $d_{\mathcal{H}} \geq 3$.

We see that having a weak constant $d = \sqrt{2}$ has in practice no effect on the decoherence-time. Remark also that the constant $\max_{i\in I} \sqrt{d_{\mathcal{H}_i}}$ is again a signature of the non-primitive case.

8.3. Advanced properties of DF norms

In the following proposition, we gather properties of $\mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}}))$, when $\mathcal{N} \equiv \mathcal{N}(\mathcal{P})$ is the decoherence-free algebra of a faithful QMS $(\mathcal{P}_t)_{t\geq 0}$, that will be particularly useful throughout this chapter:

Proposition 8.3.1. *Fix $1 \leq q \leq p \leq +\infty$ and let $(\mathcal{P}_t)_{t\geq 0}$ be a faithful QMS, with $\mathcal{N} \equiv \mathcal{N}(\mathcal{P})$. Then the following properties hold:*

(i) *$(\mathcal{P}_t)_{t\geq 0}$ is contractive with respect to $\|.\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))}$ for all $1 \leq q, p \leq +\infty$.*

(ii) *For all $X \in \mathcal{N}(\mathcal{P})$, $\|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} = \|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}$.*

(iii) *Ordering of the norms: for fixed $q \geq 1$ and for $q \leq p_1 \leq p_2$, $\|.\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p_1}(\sigma_{\mathrm{Tr}}))} \leq \|.\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p_2}(\sigma_{\mathrm{Tr}}))}$.*

(iv) *In the case when $\mathcal{N} = \mathcal{N}(\mathcal{P}) \equiv \mathbb{C}\mathbb{1}$ and $\sigma_{\mathrm{Tr}} \equiv \sigma$ is the unique invariant state, equality holds in the second inequality of (1.20) as well as in the first inequality of (1.21).*

Proof.

(i) In Proposition 5.2.1, we showed that for any $p \geq 1$, $t \geq 0$ and any invariant state σ, $(\mathcal{P}_t)_{t \geq 0}$ is an $\mathbb{L}_p(\sigma)$ contraction. In particular, it means for $\sigma = \sigma_{\mathrm{Tr}}$ that for all $X \in \mathcal{B}(\mathcal{H})$,

$$\|\mathcal{P}_t(X)\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})} \leq \|X\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})} \ .$$

Assume now that $1 \leq q \leq p \leq +\infty$. We first prove that $(\mathcal{P}_t)_{t \geq 0}$ is contractive for the $\|.\|_{\mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}}))}$ norm. By definition,

$$
\begin{aligned}
\|\mathcal{P}_t(X)\|_{\mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}}))} &= \sup_{\substack{A, B \in \mathcal{N}(\mathcal{P}), \\ \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})}, \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} = 1}} \|\mathcal{P}_t(A)\mathcal{P}_t(X)\mathcal{P}_t(B)\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} \\
&= \sup_{\substack{A, B \in \mathcal{N}(\mathcal{P}), \\ \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})}, \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} = 1}} \|\mathcal{P}_t(AXB)\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} \\
&\leq \sup_{\substack{A, B \in \mathcal{N}(\mathcal{P}), \\ \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})}, \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} = 1}} \|AXB\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} = \|X\|_{\mathbb{L}_p(\mathcal{N}, \mathbb{L}_q(\sigma_{\mathrm{Tr}}))} \ ,
\end{aligned}
$$

where $\frac{1}{r} = \frac{1}{q} - \frac{1}{p}$. Here the first line follows from the fact that $(\mathcal{P}_t)_{t \geq 0}$ acts unitarily on $\mathcal{N}$, with associated Hamiltonian H_W that commutes with σ_{Tr} (cf. Corollary 6.2.6). The second line follows by definition of a decoherence-free algebra, and the third one from the contractivity of $\mathcal{P}_t$ as a map from $\mathbb{L}_q(\sigma_{\mathrm{Tr}})$ to $\mathbb{L}_q(\sigma_{\mathrm{Tr}})$. The case of $\|.\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))}$ follows by the duality of Proposition 1.1.8 (ii) and Hölder's inequality (Proposition 1.1.8 (i)):

$$
\begin{aligned}
\|\mathcal{P}_t(X)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} &= \sup_{\|Y\|_{\mathbb{L}_{\hat{q}}(\mathcal{N}, \mathbb{L}_{\hat{p}}(\sigma_{\mathrm{Tr}}))} \leq 1} |\langle Y, \mathcal{P}_t(X) \rangle_{\sigma_{\mathrm{Tr}}}| \\
&= \sup_{\|Y\|_{\mathbb{L}_{\hat{q}}(\mathcal{N}, \mathbb{L}_{\hat{p}}(\sigma_{\mathrm{Tr}}))} \leq 1} |\langle \hat{\mathcal{P}}_t(Y), X \rangle_{\sigma_{\mathrm{Tr}}}| \\
&\leq \sup_{\|Y\|_{\mathbb{L}_{\hat{q}}(\mathcal{N}, \mathbb{L}_{\hat{p}}(\sigma_{\mathrm{Tr}}))} \leq 1} \|\hat{\mathcal{P}}_t(Y)\|_{\mathbb{L}_{\hat{q}}(\mathcal{N}, \mathbb{L}_{\hat{p}}(\sigma_{\mathrm{Tr}}))} \|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} \ ,
\end{aligned}
$$

where $\frac{1}{p} + \frac{1}{\hat{p}} = 1$ and $\frac{1}{q} + \frac{1}{\hat{q}} = 1$. We conclude by using the above proof of DF-contractivity for $1 \leq \hat{p} \leq \hat{q} \leq +\infty$, applied to the dual QMS $(\hat{\mathcal{P}}_t)_{t \geq 0}$.

(ii) This is point (viii) of Proposition 1.1.8 for $\mathcal{N} \equiv \mathcal{N}(\mathcal{P})$.

(iii) This is point (iv) of Proposition 1.1.8 for $\mathcal{N} \equiv \mathcal{N}(\mathcal{P})$.

(iv) This is obvious by definition.

$\square$

8.3.1. Differentiation of the decoherence-free norms

As in the primitive case, the equivalence between hypercontractivity and the log-Sobolev inequality relies on a formula for the differentiation of the decoherence-free norms, commonly called Gross' integration Lemma. In the bipartite case, where $\mathcal{N}(\mathcal{P}) = \mathcal{B}(\mathcal{H}_A) \otimes \mathbb{1}_{\mathcal{H}_B}$ and the invariant state is the maximally mixed state, this differentiation was done in [Beigi and King, 2016]. Here we generalise this result to the case of the amalgamated $\mathbb{L}_p$ norms associated to a faithful QMS. The next lemma, which extends Lemma 5 of [Kastoryano and Temme, 2013], provides a physical interpretation of the DF-$\mathbb{L}_p$ relative entropies in terms of the *quantum relative entropy* of a state and its projection onto the decoherence-free subalgebra. Recall that the quantum relative entropy $D(\rho\|\sigma)$ of two states

$\rho, \sigma \in \mathcal{D}(\mathcal{H})$ is given by

$$D(\rho \| \sigma) := \begin{cases} \mathrm{Tr}\left(\rho \left(\ln \rho - \ln \sigma\right)\right) & \mathrm{supp}(\rho) \subset \mathrm{supp}(\sigma), \\ + \infty & \text{otherwise.} \end{cases} \tag{8.10}$$

Lemma 8.3.2. *Let $\rho \in \mathcal{D}_+(\mathcal{H})$ and $X \in \mathcal{B}(\mathcal{H})$ positive definite, then*

(i) $\mathrm{Ent}_{q,\mathcal{N}}(\Gamma_{\sigma_{\mathrm{Tr}}}^{-\frac{1}{q}}(\rho^{\frac{1}{q}})) = D(\rho \| \rho_{\mathcal{N}})$ *for any $q \geq 1$.*

(ii) *If $X \in \mathcal{N}(\mathcal{P})$, then for any $q \geq 1$, $\mathrm{Ent}_{q,\mathcal{N}}(X) = 0$.*

(iii) $\mathrm{Ent}_{p,\mathcal{N}}(X) = \mathrm{Ent}_{2,\mathcal{N}}(\mathrm{I}_{2,p}(X))$ *for any $p \geq 1$.*

Proof.

(i) For $X = \Gamma_{\sigma_{\mathrm{Tr}}}^{-\frac{1}{q}}(\rho^{\frac{1}{q}})$, Equation (8.7) reduces to

$$\mathrm{Ent}_{q,\mathcal{N}}(X) = \mathrm{Tr}(\rho \ln \rho) - \mathrm{Tr}\left(\rho \ln E_{\mathcal{N}}\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{-1}(\rho)\right]\right) - \mathrm{Tr}(\rho \ln \sigma_{\mathrm{Tr}}). \tag{8.11}$$

Now, $\mathrm{Tr}\left(\rho \ln E_{\mathcal{N}}\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{-1}(\rho)\right]\right) = \mathrm{Tr}\left(\rho_{\mathcal{N}} \ln E_{\mathcal{N}}\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{-1}(\rho)\right]\right)$. Using Equation (1.17) together with $[\sigma_{\mathrm{Tr}}, E_{\mathcal{N}_*}(\rho)] = 0$, we arrive at

$$\mathrm{Tr}\left(\rho_{\mathcal{N}} \ln E_{\mathcal{N}}\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{-1}(\rho)\right]\right) = D(\rho_{\mathcal{N}} \| \sigma_{\mathrm{Tr}}).$$

Substituting the above right hand side into (8.11), we finally arrive at (cf. [Bardet, 2017])

$$\mathrm{Ent}_{q,\mathcal{N}}(Y) = D(\rho \| \sigma_{\mathrm{Tr}}) - D(\rho_{\mathcal{N}} \| \sigma_{\mathrm{Tr}}) = D(\rho \| \rho_{\mathcal{N}}). \tag{8.12}$$

(ii) This is a simple consequence of (i) together with the fact that if $X \in \mathcal{N}(\mathcal{P})$, $[X, \sigma_{\mathrm{Tr}}] = 0$ and

$$\rho_{\mathcal{N}} := E_{\mathcal{N}_*}[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(X)^q] = E_{\mathcal{N}_*}[\Gamma_{\sigma_{\mathrm{Tr}}}(X^q)] = \Gamma_{\sigma_{\mathrm{Tr}}}(E_{\mathcal{N}}[X^q]) = \Gamma_{\sigma_{\mathrm{Tr}}}(X^q) = (\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(X))^q \equiv \rho.$$

(iii) follows by direct computation.

$\square$

The proof of next theorem follows closely the one of Theorem 7 of [Beigi and King, 2016], and is discussed in Appendix 8.A for sake of clarity. It can be seen as both a generalisation of the differentiation done in the primitive case in [Olkiewicz and Zegarlinski, 1999] to non-primitive QMS (see also Lemma 14 of [Kastoryano and Temme, 2013]), and the one carried out for the CB-norm in [Beigi and King, 2016] to the non-unital case.

Theorem 8.3.3. *Let $t \mapsto p(t)$ be a twice continuously differentiable increasing function in a neighborhood of 0, with $p(0) = q \geq 1$. Also let $t \mapsto Y(t) \in \mathcal{B}(\mathcal{H})$ be an operator-valued twice continuously differentiable function, where $Y(t)$ is positive definite in a neighborhood of 0, and denote $Y \equiv Y(0)$. Then*

$$\frac{d}{dt}\|Y(t)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))} \bigg|_{t=0} = \frac{p'(0)}{q^2 \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}} \left(\mathrm{Ent}_{q,\mathcal{N}}(Y) + \frac{q^2}{p'(0)} \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^{q-1} \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y'(0))\right)\right).$$

We shall apply this theorem to different situations. Perhaps the most relevant one is when $Y(t)$ models the evolution of an initial observable $X \in \mathcal{B}(\mathcal{H})$ under the QMS $(\mathcal{P}_t)_{t \geq 0}$. We state it as a corollary.

Corollary 8.3.4. *For any positive definite $X \in \mathcal{B}(\mathcal{H})$,*

$$\frac{d}{dt}\|\mathcal{P}_t(X)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))}\Big|_{t=0} = \frac{p'(0)}{q^2\|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}}\left(\mathrm{Ent}_{q,\mathcal{N}}(X) - \frac{4(q-1)}{p'(0)}\mathcal{E}_{q,\mathcal{L}}(X)\right). \tag{8.13}$$

8.3.2. Almost uniform convexity

In this subsection, we study an analogue of the well-known uniform convexity of the Schatten norms proved in [Ball et al., 1994]. This analogue was proved in the context of weighted $\mathbb{L}_p(\sigma)$ norms by [Olkiewicz and Zegarlinski, 1999]. This will be an essential tool when proving universal lower bounds on the weak DF-log-Sobolev constants. This inequality states that for all X positive semidefinite, any full-rank state σ, and all $p \in [1, 2]$,

$$\|X\|_{\mathbb{L}_p(\sigma)}^2 \geq (p-1)\|X - \mathrm{Tr}(\sigma X)\|_{\mathbb{L}_p(\sigma)}^2 + \mathrm{Tr}(\sigma X)^2. \tag{8.14}$$

For the Shattern norms, this inequality can be seen as a consequence of Clarkson inequalities (see [Pisier and Xu, 2003] for a discussion of this fact). It has many important applications in the theory of non-commutative $\mathbb{L}_p$ spaces, such as yielding the optimal constant for Fermionic hypercontractivity [Carlen and Lieb, 1993]. We shall prove however in Section 8.5 that the inequality does not hold for the amalgamated $\mathbb{L}_p$ spaces. Instead, in this section we prove a weaker form of it.

It will be useful to denote by $\mathcal{S}_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^+$ the set of positive definite operators on the sphere of radius 1 in $\mathbb{L}_1(\mathrm{Tr})$. Then, for $X \in \mathcal{B}(\mathcal{H})$, $A \in \mathcal{N}(\mathcal{P}) \cap \mathcal{S}_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^+$ and $p \geq 1$, we define

$$\Phi(X, A, p) := \left\|\Gamma_A^{-1/r}(X)\right\|_p = \mathrm{Tr}\left[\left|A^{-\frac{1}{2r}} X A^{-\frac{1}{2r}}\right|^p\right]^{\frac{1}{p}}, \tag{8.15}$$

where we recall that $1/r = |1/2 - 1/p|$. Remark that for all positive semidefinite $X \in \mathcal{B}(\mathcal{H})$ and all $A \in \mathcal{N}(\mathcal{P}) \cap \mathcal{S}_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^+$, $\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(X), A, 2) = \|X\|_{2, \sigma_{\mathrm{Tr}}}$. We shall prove that a similar inequality as (8.14) holds for Φ, which we subsequently refer to as *almost uniform convexity*.

Lemma 8.3.5. *The two following properties hold:*

(i) For all $X \in \mathcal{B}_{sa}^+(\mathcal{H})$, $A \in \mathcal{N}(\mathcal{P}) \cap \mathcal{S}_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^+$ and all $1 \leq p \leq 2$,

$$\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X), A, p)^2 \geq (p-1)\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X - E_{\mathcal{N}}[X]), A, p)^2 + \Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(E_{\mathcal{N}}[X]), A, p)^2. \tag{8.16}$$

(ii) For all $X \in \mathcal{B}_{sa}^+(\mathcal{H})$ and $A \in \mathcal{N}(\mathcal{P}) \cap \mathcal{S}_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^+$,

$$\begin{aligned}
\frac{\partial}{\partial p}\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X), A, p)^2\Big|_{p=2} &\leq \frac{\partial}{\partial p}\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X - E_{\mathcal{N}}[X]), A, p)^2\Big|_{p=2} \\
&\quad + \frac{\partial}{\partial p}\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(E_{\mathcal{N}}[X]), A, p)^2\Big|_{p=2} \\
&\quad + \|X - E_{\mathcal{N}}[X]\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}^2.
\end{aligned} \tag{8.17}$$

Proof. We follow the proof of Lemma 2.9 in [Olkiewicz and Zegarlinski, 1999] in order to prove the

first claim. We adopt the following notations. For $0 \le t \le 1$, define

$$X(t) = E_{\mathcal{N}}[X] + t\left(X - E_{\mathcal{N}}[X]\right),$$

$$\varphi(t) = \Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X(t)), A, p)^2,$$

$$h = \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}} \circ \Gamma_A^{-\frac{1}{r}}(X - E_{\mathcal{N}}[X]).$$

Then, Equation (8.16) reduces to:

$$\varphi(1) \ge (p-1)\|h\|_p^2 + \varphi(0). \tag{8.18}$$

This inequality follows directly from:

1. $\varphi'(0) = 0$;

2. $\varphi''(t) \ge 2(p-1)\|h\|_p^2$ for all $0 \le t \le 1$.

We start by computing $\varphi'(t)$. Writting $Z(t) = \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}} \circ \Gamma_A^{-\frac{1}{r}}(X(t))$, we have by integral representation that for all $0 \le t \le 1$

$$\varphi'(t) = 2\,\mathrm{Tr}[h\,Z(t)^{p-1}]\,\mathrm{Tr}[Z(t)^p]^{2/p-1}, \tag{8.19}$$

$$\varphi''(t) \ge 2\frac{\partial}{\partial t}\left(\mathrm{Tr}[h\,Z(t)^{p-1}]\right)\mathrm{Tr}[Z(t)^p]^{2/p-1}. \tag{8.20}$$

We start by proving claim 1. First remark that, since elements of $\mathcal{N}$ commute with σ_{Tr},

$$Z(0)^{p-1} = \sigma_{\mathrm{Tr}}^{\frac{1}{2}-\frac{1}{2p}}\,E_{\mathcal{N}}\left[\Gamma_A^{-\frac{1}{r}}(X)\right]^{p-1}\sigma_{\mathrm{Tr}}^{\frac{1}{2}-\frac{1}{2p}}.$$

Therefore $\mathrm{Tr}[h\,Z(0)^{p-1}] = \langle X - E_{\mathcal{N}}[X], B\rangle_{\sigma_{\mathrm{Tr}}}$ where $B = \Gamma_A^{-\frac{1}{r}}\left(E_{\mathcal{N}}\left[\Gamma_A^{-\frac{1}{r}}(X)\right]^{p-1}\right) \in \mathcal{N}$. By Equation (1.16) we get that $\mathrm{Tr}[h\,Z^{p-1}] = 0$ which results in $\varphi'(0) = 0$. The proof of claim 2 is a direct copy of the proof of Lemma 2.9 in [Olkiewicz and Zegarlinski, 1999] and we omit it. Hence, (8.16) holds.

In order to prove (ii), we rearrange the terms in (8.16) to get

$$(2-p)\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X - E_{\mathcal{N}}[X]), A, p)^2$$

$$\ge \left(\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(X), A, 2) + \Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X), A, p)\right)\left(\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(X), A, 2) - \Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X), A, p)\right)$$

$$- \left(\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(X - E_{\mathcal{N}}[X]), A, 2) + \Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X - E_{\mathcal{N}}[X]), A, p)\right) \times$$

$$\times \left(\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(X - E_{\mathcal{N}}[X]), A, 2) - \Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(X - E_{\mathcal{N}}[X]), A, p)\right)$$

$$- \left(\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(E_{\mathcal{N}}[X]), A, 2) + \Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(E_{\mathcal{N}}[X]), A, p)\right) \times$$

$$\times \left(\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(E_{\mathcal{N}}[X]), A, 2) - \Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(E_{\mathcal{N}}[X]), A, p)\right),$$

where we used that

$$\Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(X), A, 2)^2 = \Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(X - E_{\mathcal{N}}[X]), A, 2)^2 + \Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{2}}(E_{\mathcal{N}}[X]), A, 2)^2.$$

(8.17) follows by dividing this inequality by $2 - p$ and taking the limit $p \to 2$. $\qquad\square$

8.4. DF-hypercontractivity and the log-Sobolev inequality

In this section we state and prove the main results of this article. In Section 8.4.1, we prove the equivalence between hypercontractivity for the amalgamated norms and the DF-log-Sobolev inequality. In Section 8.4.2, we prove that the weak constant in the DF-log-Sobolev inequality can always be upper bounded by a universal constant, namely $\ln \sqrt{2}$. In Section 8.4.3, we show how to derive estimates on the log-Sobolev constants using interpolation techniques. Finally, we combine these two last results in order to obtain generic bounds on both constants.

8.4.1. Equivalence between HC and LSI

Here, we state and prove the main result of this section, that is, the equivalence between the DF-log-Sobolev inequality and DF-hypercontractivity.

Theorem 8.4.1. *Let $(\mathcal{P}_t)_{t\geq 0}$ be a faithful QMS on $\mathcal{B}(\mathcal{H})$ with associated generator $\mathcal{L}$, and let $q \geq 1$, $d \geq 0$ and $p(t) = 1 + (q-1) e^{\frac{4t}{c}}$ for some constant $c > 0$. Then*

(i) *If $\mathrm{HC}_{q,\mathcal{N}}(c,d)$ holds, then $\mathrm{LSI}_{q,\mathcal{N}}(c,d)$ holds.*

(ii) *If $\mathrm{LSI}_{p(t),\mathcal{N}}(c,d)$ holds for all $t \geq 0$, then $\mathrm{HC}_{q,\mathcal{N}}(c,d + \ln\sqrt{|I|})$ holds, where $|I|$ denotes the number of blocks in the decomposition of $\mathcal{N}$ as given in Equation (0.9).*

Remark 8.4.2. For primitive evolution, $|I| = 1$ and we recover the equivalence between hypercontractivity and the logarithmic Sobolev inequality of [Olkiewicz and Zegarlinski, 1999]. The equivalence is also achieved in the more general situation where $\mathcal{N}$ is a factor, that is, in the situation of Section 10.2. Below, we discuss why the term $\ln\sqrt{|I|}$ may not be optimal.

Proof. We first prove (i). For $X > 0$, define the function

$$F : [0, +\infty) \ni t \mapsto \exp\left\{-2d\left(\frac{1}{q} - \frac{1}{p(t)}\right)\right\} \|\mathcal{P}_t(X)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))},$$

where $p(t) := 1 + (q-1) e^{\frac{4t}{c}}$. $\mathrm{HC}_{q,\mathcal{N}}(c,d)$ implies that $\ln F(t) \leq \ln F(0)$ for all $t \geq 0$, with equality at $t = 0$. Therefore,

$$\frac{d \ln F(t)}{dt}\bigg|_{t=0^+} = -2d\, \frac{p'(0)}{q^2} + \frac{d}{dt} \ln \|\mathcal{P}_t(X)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))}\bigg|_{t=0^+} \leq 0.$$

Using Equation (8.13), the above inequality reduces to

$$-2d + \frac{1}{\|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q} \left(\mathrm{Ent}_{q,\mathcal{N}}(X) - c\,\mathcal{E}_{q,\mathcal{L}}(X)\right) \leq 0,$$

which yields $\mathrm{LSI}_{q,\mathcal{N}}(c,d)$.

To prove (ii), we proceed by contradiction, similarly to [Beigi and King, 2016]. The main difference resides in the replacement of the norm by an auxillary quantity that allows to control a remainder term that does not appear in the case where $\mathcal{N}$ is a factor. Without loss of generality, we assume that the evolution on elements of the decoherence-free algebra is trivial, and refer to Lemma 8.4.3 for a justification of this fact.

Next, assume that there exists an $X \in \mathcal{B}(\mathcal{H})$ such that hypercontractivity fails for this X. Following the same proof as Theorem 12 of [Devetak et al., 2006], we can show that it is sufficient to

consider that X is a positive definite operator. Indeed, for fixed $q \le p$, if there exists $C > 0$ such that for any X positive definite,

$$\|\mathcal{P}_t(X)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} \le C\|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})},$$

then the inequality remains true for any $X \in \mathcal{B}(\mathcal{H})$. Without loss of generality, we also assume that $\|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} = 1$. Then, suppose that there exists some time $t_0 > 0$ such that

$$\|\mathcal{P}_{t_0}(X)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{q(t_0)}(\sigma_{\mathrm{Tr}}))} > \exp\left\{2(d + \ln\sqrt{|I|})\left(\frac{1}{q} - \frac{1}{q(t_0)}\right)\right\}.$$

Define, for $\varepsilon > 0$,

$$\tilde{\varphi}(t) := \|\mathcal{P}_t(X)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))} \exp\left\{-2(d + \ln\sqrt{|I|})\left(\frac{1}{q} - \frac{1}{p(t)}\right)\right\} - \varepsilon t,$$

where $\|.\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))}$ is defined in Section 8.B. By definition, $\|\mathcal{P}_{t_0}(X)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t_0)}(\sigma_{\mathrm{Tr}}))} \ge \|\mathcal{P}_{t_0}(X)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t_0)}(\sigma_{\mathrm{Tr}}))}$ so that $\tilde{\varphi}(t_0) > 1$ for ε small enough. Define the set $U := \{t \in [0, t_0] : \tilde{\varphi}(t) \le 1\}$. Since $\mathcal{P}_0 = \mathrm{id}$ and $p(0) = q$, we have $\tilde{\varphi}(0) = \|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} = 1$, so that $U \ne \varnothing$. Let u be the supremum of the set U. By continuity of $t \mapsto \|\mathcal{P}_t(X)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))}$ (cf. Lemma 8.B.4), $\tilde{\varphi}$ is continuous and therefore $u \in U$ and $u < t_0$. Now, by definition of u, for all $t \in (u, t_0]$, $\tilde{\varphi}(t) > 1 = \tilde{\varphi}(u)$. For $t > 0$, let $\tilde{A}(t)$ be the unique minimiser of

$$\mathcal{N}(\mathcal{P}) \cap \tilde{\mathcal{S}}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} \ni A \mapsto \|A^{-s(t)/2}\mathcal{P}_t(X)A^{-s(t)/2}\|_{\mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}})},$$

as characterised in Lemma 8.B.3, where $\tilde{\mathcal{S}}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$ is defined in Section 8.B and $s(t) = \frac{1}{q} - \frac{1}{p(t)}$. Define

$$\mu(t) := \|\tilde{A}(u)^{-s(t)/2}\mathcal{P}_t(X)\tilde{A}(u)^{-s(t)/2}\|_{\mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}})} \exp\left\{-2(d + \ln\sqrt{|I|})\left(\frac{1}{q} - \frac{1}{p(t)}\right)\right\} - \varepsilon t.$$

Therefore, for all $t \ge u$,

$$\mu(t) \ge \inf_{A \in \mathcal{N}(\mathcal{P}) \cap \tilde{\mathcal{S}}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}} \|A^{-s(t)/2}\mathcal{P}_t(X)A^{-s(t)/2}\|_{\mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}})} \exp\left\{-2(d + \ln\sqrt{|I|})\left(\frac{1}{q} - \frac{1}{p(t)}\right)\right\} - \varepsilon t = \tilde{\varphi}(t)$$

and $\tilde{\varphi}(u) = \mu(u)$. Now, the derivative of $\mu(t)$ at $t = u$ can be computed using Equation (8.40) with $X(t) = \Gamma^{1/p(t)}_{\sigma_{\mathrm{Tr}}} \circ \mathcal{P}_t(X)$ and $A = \tilde{A}(u)$. Given $M(t) := \tilde{A}(u)^{-s(t)/2}\mathcal{P}_t(X)\tilde{A}(u)^{-s(t)/2}$, one finds

$$\frac{\partial}{\partial t}\bigg|_{t=u} \|\tilde{A}(u)^{-s(t)/2}\mathcal{P}_t(X)\tilde{A}(u)^{-s(t)/2}\|_{\mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}})} \tag{8.21}$$

$$= \frac{1}{\|M(u)\|^{p(u)-1}_{p(u),\sigma_{\mathrm{Tr}}}} \left(-\frac{p'(u)}{p(u)^2} \mathrm{Tr}\left(\Gamma^{\frac{1}{p(u)}}_{\sigma_{\mathrm{Tr}}}(M(u))^{p(u)}\right) \ln \mathrm{Tr}\left[\Gamma^{\frac{1}{p(u)}}_{\sigma_{\mathrm{Tr}}}(M(u))^{p(u)}\right]\right.$$

$$+ \frac{p'(u)}{p(u)^2} \mathrm{Tr}\left[\Gamma^{\frac{1}{p(u)}}_{\sigma_{\mathrm{Tr}}}(M(u))^{p(u)} \ln\Gamma^{\frac{1}{p(u)}}_{\sigma_{\mathrm{Tr}}}(M(u))^{p(u)}\right] - \frac{p'(u)}{p(u)^2} \mathrm{Tr}\left[\Gamma^{\frac{1}{p(u)}}_{\sigma_{\mathrm{Tr}}}(M(u))^{p(u)} \ln\tilde{A}(u)\right]$$

$$+ \mathrm{Tr}\left[\Gamma^{\frac{1}{p(u)}}_{\sigma_{\mathrm{Tr}}}(M(u))^{p(u)-1}\tilde{A}(u)^{-\frac{s(u)}{2}}\left\{\Gamma^{\frac{1}{p(u)}}_{\sigma_{\mathrm{Tr}}}\mathcal{L}(\mathcal{P}_u(X)) - \frac{p'(u)\{\ln\sigma_{\mathrm{Tr}}, \Gamma^{\frac{1}{p(u)}}_{\sigma_{\mathrm{Tr}}}(\mathcal{P}_u(X))\}}{2p(u)^2}\right\}\tilde{A}(u)^{-\frac{s(u)}{2}}\right]\right).$$

$$\tag{8.22}$$

Define $\rho(u) \coloneqq \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p(u)}}(M(u))^{p(u)}$. Then, Equation (8.22) simplifies into

$$
\left.\frac{\partial}{\partial t}\right|_{t=u} \|\tilde{A}(u)^{-s(t)/2}\mathcal{P}_t(X)\tilde{A}(u)^{-s(t)/2}\|_{\mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}})} \tag{8.23}
$$

$$
= \frac{p'(u)}{p(u)^2 \|M(u)\|_{\mathbb{L}_{p(u)}(\sigma_{\mathrm{Tr}})}^{p(u)-1}} \Big(D(\rho(u)\|E_{\mathcal{N}*}(\rho(u))) - c\,\mathcal{E}_{p(u),\mathcal{L}}(M(u))
$$

$$
- \mathrm{Tr}(\rho(u))\ln\mathrm{Tr}\,\rho(u) + \mathrm{Tr}\left(\rho(u)\ln E_{\mathcal{N}*}(\rho(u))\right) - \mathrm{Tr}\left(\rho(u)\ln\tilde{A}(u)\right) - \mathrm{Tr}(\rho(u)\ln\sigma_{\mathrm{Tr}})\Big)
$$

$$
= \frac{p'(u)}{p(u)^2 \|M(u)\|_{\mathbb{L}_{p(u)}(\sigma_{\mathrm{Tr}})}^{p(u)-1}} \Big\{ \mathrm{Ent}_{p(u),\mathcal{N}}(M(u)) - c\,\mathcal{E}_{p(u),\mathcal{L}}(M(u))
$$

$$
+ \Big(-\mathrm{Tr}(\rho(u))\ln\mathrm{Tr}\,\rho(u) + \mathrm{Tr}\left(\rho(u)\ln E_{\mathcal{N}*}(\rho(u))\right) - \mathrm{Tr}\left(\rho(u)\ln\tilde{A}(u)\right) - \mathrm{Tr}\left(\rho(u)\ln\sigma_{\mathrm{Tr}}\right)\Big)\Big\},
$$

where, in order to get the Dirichlet form, we also used that for any $A \in \mathcal{N}$ and $X \in \mathcal{B}(\mathcal{H})$, $\mathcal{P}_t(AXA) = A\mathcal{P}_t(X)A \to \mathcal{L}(AXA) = A\mathcal{L}(X)A$, since we assume the evolution to be trivial on $\mathcal{N}$. Using the expression for $\tilde{A}(u)$ derived in Equation (8.62),

$$
\ln\tilde{A}(u) = \ln\sum_{i\in I} P_i \tilde{A}(u) P_i
$$

$$
= \sum_{i\in I} \ln P_i \tilde{A}(u) P_i
$$

$$
= \sum_{i\in I} -\ln|I|\,P_i - \ln\mathrm{Tr}(P_i\,\rho(u)\,P_i)P_i - \ln P_i\sigma_{\mathrm{Tr}}P_i + \ln P_i\,E_{\mathcal{N}*}(\rho(u))P_i
$$

$$
\geq -\ln|I| - \ln\sigma_{\mathrm{Tr}} - \ln\mathrm{Tr}(\rho(u)) + \ln E_{\mathcal{N}*}(\rho(u)).
$$

Using this expression, Equation (8.23) leads to:

$$
\left.\frac{\partial}{\partial t}\right|_{t=u} \|\tilde{A}(u)^{-s(t)/2}\mathcal{P}_t(X)\tilde{A}(u)^{-s(t)/2}\|_{\mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}})}
$$

$$
\leq \frac{p'(u)}{p(u)^2 \|M(u)\|_{\mathbb{L}_{p(u)}(\sigma_{\mathrm{Tr}})}^{p(u)-1}} \Big\{ \mathrm{Ent}_{p(u),\mathcal{N}}(M(u)) - c\,\mathcal{E}_{p(u),\mathcal{L}}(M(u)) + \ln|I|\,\|M(u)\|_{\mathbb{L}_{p(u)}(\sigma_{\mathrm{Tr}})}^{p(u)}\Big\}.
$$

Then, using the assumption that $\mathrm{LSI}_{p(u),\mathcal{N}}(c,d)$ holds, we find that

$$
\mu'(u) \leq -\varepsilon.
$$

Therefore, there exists $\delta > 0$ such that $u + \delta \leq t_0$ and $\mu(u+\delta) \leq \mu(u)$. We then have

$$
\tilde{\varphi}(u+\delta) \leq \mu(u+\delta) \leq \mu(u) = \tilde{\varphi}(u) \leq 1,
$$

which is in contradiction with the very definition of u. $\qquad\square$

Lemma 8.4.3. *Let $(\mathcal{P}_t)_{t\geq 0}$ be a faithful QMS of generator $\mathcal{L}$ and associated decoherence-free Hamiltonian H_W (cf. Corollary 6.2.6). Then $\mathrm{HC}_{q,\mathcal{N}}(c,d)$ holds for $(\mathcal{P}_t)_{t\geq 0}$ if and only if it holds for the semigroup $(\mathcal{P}_t')$ constructed from $(\mathcal{P}_t)_{t\geq 0}$ by removing the unitary evolution associated to H_W: $\mathcal{L}' = \mathcal{L} - i[H_W, .]$. Consequently, $\mathrm{LSI}_{q,\mathcal{N}}(c,d)$ holds for $(\mathcal{P}_t)_{t\geq 0}$ if and only if it holds for $(\mathcal{P}_t')_{t\geq 0}$.*

Proof. From Theorem 6.2.5, the following holds for any $X \in \mathcal{B}(\mathcal{H})$:

$$
\mathcal{P}_t'(X) = \mathrm{e}^{-iH_W t}\mathcal{P}_t(X)\,\mathrm{e}^{iH_W t}.
$$

Therefore, for any $1 \leq q \leq p \leq \infty$

$$
\begin{aligned}
\|\mathcal{P}_t(X)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} &= \inf_{\substack{A, B \in \mathcal{N}, Y \in \mathcal{B}(\mathcal{H}) \\ \mathcal{P}_t(X) = AYB}} \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|Y\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})} \\
&= \inf_{\substack{A, B \in \mathcal{N}, Y \in \mathcal{B}(\mathcal{H}) \\ \mathcal{P}_t'(X) = e^{-iH_W t} A Y B e^{iH_W t}}} \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|Y\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})} \\
&= \inf_{\substack{A, B \in \mathcal{N}, Y \in \mathcal{B}(\mathcal{H}) \\ \mathcal{P}_t'(X) = AYB}} \|A\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|B\|_{\mathbb{L}_{2r}(\sigma_{\mathrm{Tr}})} \|Y\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})} \\
&= \|\mathcal{P}_t'(X)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} \, .
\end{aligned}
$$

where the third line follows from the commutation of H_W with σ_{Tr} and the unitary invariance of Schatten norms. From this, the equivalence between the hypercontractivity properties of $(\mathcal{P}_t)_{t \geq 0}$ and $(\mathcal{P}_t')_{t \geq 0}$ follows directly. From Theorem 8.4.1 (i) we also get the correspondence of the logarithmic Sobolev constants between the two semigroups. $\qquad \square$

In the above theorem, one needs $\mathrm{LSI}_{\tilde{q}, \mathcal{N}}(c, d)$ to hold for any $\tilde{q} \geq q$ in order to conclude that $\mathrm{HC}_{q, \mathcal{N}}(c, d + \ln \sqrt{|I|})$ holds. Under the assumption of regularity of the Dirichlet forms, it is enough to assume that it holds for $q = 2$ only.

Theorem 8.4.4. *Assume that $\mathrm{LSI}_{2, \mathcal{N}}(c, d)$ holds. If the generator $\mathcal{L}$ is strongly $\mathbb{L}_p$-regular for some $d_0 \geq 0$, then $\mathrm{LSI}_{q, \mathcal{N}}(c, d + c \, d_0/2)$ holds for all $q \geq 1$, so that $\mathrm{HC}_{2, \mathcal{N}}(c, d + \ln \sqrt{|I|} + c \, d_0/2)$ holds.*

Proof. From Lemma 8.3.2 (iii),

$$
\begin{aligned}
\mathrm{Ent}_{q, \mathcal{N}}(X) &= \mathrm{Ent}_{2, \mathcal{N}}(\mathrm{I}_{2, q}(X)) \\
&\leq \left(c \, \mathcal{E}_{2, \mathcal{L}}(\mathrm{I}_{2, q}(X)) + 2d \, \|\mathrm{I}_{2, q}(X)\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}^2 \right) \\
&\leq c \mathcal{E}_{q, \mathcal{L}}(X) + 2 \left(d + \frac{c \, d_0}{2} \right) \|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q \, ,
\end{aligned}
$$

where in the last line we used that $\mathcal{E}_{2, \mathcal{L}}(\mathrm{I}_{2, q}(X)) \leq \mathcal{E}_{q, \mathcal{L}}(X) + d_0 \|X\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q$ by strong $\mathbb{L}_p$-regularity. $\qquad \square$

It was shown in [Bardet, 2017] that any generator satisfying σ_{Tr}-DBC is strongly regular with constant $d_0 = 0$ (cf. Section 5.4). Furthermore, under the condition of KMS-symmetry, a straightforward extension of the proof of Proposition 5.2 of [Olkiewicz and Zegarlinski, 1999] to the case of a non-primitive QMS implies that the strong $\mathbb{L}_p$-regularity of $\mathcal{L}$ always holds, with $d_0 = \|\mathcal{L} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\sigma_{\mathrm{Tr}})\| + 1$. The following corollary is a straightforward consequence of these two facts.

Corollary 8.4.5. *Assume that $\mathrm{LSI}_{2, \mathcal{N}}(c, d)$ holds. Then:*

(i) *If $(\mathcal{P}_t)_{t \geq 0}$ is KMS-symmetric, then $\mathrm{HC}_{2, \mathcal{N}}(c, d + \ln \sqrt{|I|} + c \, (\|\mathcal{L} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\sigma_{\mathrm{Tr}})\| + 1)/2)$ holds.*

(ii) *If $\mathcal{L}$ satisfies σ_{Tr}-DBC, then $\mathrm{HC}_{2, \mathcal{N}}(c, d + \ln \sqrt{|I|})$ holds.*

Bounds on the amalgamated $\mathbb{L}_p$ norms of the identity: One can potentially hope to improve the bound (1.28) by applying Theorem 8.4.1 to the trivial QMS $(\mathrm{id})_{t \geq 0}$:

Proposition 8.4.6. *Let $\mathcal{N}$ be a matrix algebra on $\mathcal{H}$ with decomposition (0.9), $E_{\mathcal{N}} : \mathcal{B}(\mathcal{H}) \to \mathcal{N}$ a conditional expectation onto $\mathcal{N}$ and $\sigma_{\mathrm{Tr}} = E_{\mathcal{N}*}(d_{\mathcal{H}}^{-1} \mathbb{1})$. Then:*

(i) *Assume that* $\|\mathrm{id} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))\| \le C^{\frac{1}{2} - \frac{1}{p}}$ *for some* $C > 0$ *and for all* $p \ge 2$. *Then* $D\left(\rho \| E_{\mathcal{N}*}(\rho)\right) \le \ln C$ *for any density matrix* $\rho \in \mathcal{D}(\mathcal{H})$.

(ii) *Conversely, assume that there exists a constant* $C > 0$ *such that* $D\left(\rho \| E_{\mathcal{N}*}(\rho)\right) \le \ln C$ *for all density matrix* $\rho \in \mathcal{D}(\mathcal{H})$. *Then for any* $p \ge 2$

$$\|\mathrm{id} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))\| \le (|I|C)^{\frac{1}{2} - \frac{1}{p}}$$

where $|I|$ *is the number of blocks in the decomposition (0.9) of* $\mathcal{N}$.

Remark 8.4.7. In the proposition we ask that $\|\mathrm{id} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))\| \le C^{\frac{1}{2} - \frac{1}{p}}$ for all $p \ge 2$. This is actually not needed, since by Riesz-Thorin interpolation Theorem this is equivalent to

$$\|\mathrm{id} : \mathbb{L}_1(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_1(\mathcal{N}, \mathbb{L}_\infty(\sigma_{\mathrm{Tr}}))\| \le C.$$

We see here that it is central that the norms we use form an interpolating family of norms.

The last proposition is not optimal, which indicates that point (ii) in Theorem 8.4.1 may also not be, even for a non-trivial evolution. To see this, consider the situation where $\mathcal{N}$ is the algebra of diagonal operator in some orthonormal basis. In this case $|I|$ is equal to the dimension $d_{\mathcal{H}}$ of the Hilbert space $\mathcal{H}$ (the converse is also true: if $|I| = d_{\mathcal{H}}$ then $\mathcal{N}$ is commutative maximal). In this case $D\left(\rho \| E_{\mathcal{N}*}(\rho)\right) \le \ln d_{\mathcal{H}}$, with equality for the maximally coherent state Ω:

$$\Omega = \frac{1}{d_{\mathcal{H}}} \sum_{i,j=1}^{d_{\mathcal{H}}} |e_i\rangle\langle e_j|,$$

where $(e_i)_{i=1,\dots,d_{\mathcal{H}}}$ is the orthonormal basis with respect to which the operators in $\mathcal{N}$ are diagonal. It means that Equation (1.28) saturates and that the implication in Proposition 8.4.6 (i) becomes an equivalence.

8.4.2. A universal upper bound on the weak log-Sobolev constant

Here and in the next section, we show how to get a DF log-Sobolev inequality with universal constants in terms of the spectral gap of the QMS. The first step is to prove that the weak log-Sobolev inequality together with the DF-Poincaré's inequality imply a universal weak log-Sobolev constant.

Theorem 8.4.8. *Assume that* $\mathrm{LSI}_{2,\mathcal{N}}(c,d)$ *holds and denote by* $\lambda(\mathcal{L})$ *the spectral gap of* $\mathcal{L}$. *Then* $\mathrm{LSI}_{2,\mathcal{N}}\left(c + 2\frac{d+1}{\lambda(\mathcal{L})}, d' = \ln\sqrt{2}\right)$ *holds.*

One can obtain from the uniform convexity (8.14) the following inequality

$$\mathrm{Ent}_{2,\sigma}(X) \le \mathrm{Ent}_{2,\sigma}\left(|X - \mathrm{Tr}[\sigma X]|_2\right) + 2\,\mathrm{Var}_\sigma(X),$$

where $\mathrm{Var}_\sigma(X) = \|X - \mathrm{Tr}(\sigma X)\|_{\mathbb{L}_2(\sigma)}^2$ is the variance of $X \in \mathcal{B}_{sa}(\mathcal{H})$ under the state σ, and for any $Z \in \mathcal{B}_{sa}(\mathcal{H})$,

$$|Z|_2 := \Gamma_\sigma^{-\frac{1}{2}} |\Gamma_\sigma^{\frac{1}{2}}(Z)|.$$

From this we can derive the analogue result of Theorem 8.4.8 in the primitive case (see Theorem 4.2 of [Olkiewicz and Zegarlinski, 1999]). The extension of this result is the subject of the next proposition.

Proposition 8.4.9. *For all* $X \in \mathcal{B}_{sa}^+(\mathcal{H})$,

$$\mathrm{Ent}_{2,\mathcal{N}}(X) \le \mathrm{Ent}_{2,\mathcal{N}}(|X - E_{\mathcal{N}}[X]|_2) + 2\,\mathrm{Var}_{\mathcal{N}}(X) + 2\ln\sqrt{2}\,\|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}^2. \tag{8.24}$$

Proof. We shall adopt the notations introduced in Section 8.3.2 and write for $Z \in \mathcal{B}_{sa}(\mathcal{H})$:

$$Z_{\mathcal{N}} = \frac{E_{\mathcal{N}}\left[I_{1,2}(Z)\right]}{\|Z\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}}\,.$$

Using Equation (8.42) with $q = 2$ as well as Lemma 8.A.2 and Equation (8.8), we find that

$$\frac{\partial}{\partial p}\Phi(\Gamma^{\frac{1}{p}}_{\sigma_{\mathrm{Tr}}}(|Z|_2), A, p)^2\Big|_{p=2} = \frac{1}{2}\,\mathrm{Ent}_{2,\mathcal{N}}(|Z|_2) + \frac{1}{2}\,\|Z\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}\,D\left(\Gamma_{\sigma_{\mathrm{Tr}}}(Z_{\mathcal{N}})\,\|\,\Gamma_{\sigma_{\mathrm{Tr}}}(A)\right),$$

where Φ is defined in Equation (8.15), and where we used $I_{1,2}(Z) = I_{1,2}(|Z|_2)$ and $\||Z|_2\|_{\mathbb{L}_2(\sigma)} = \|Z\|_{\mathbb{L}_2(\sigma)}$. Consequently, by Equation (8.17) we get that for all $A \in \mathcal{N}(\mathcal{P}) \cap \mathcal{S}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$ and for $p = \frac{\|E_{\mathcal{N}}[X]\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}}{\|X\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}}$,

$$\begin{aligned}
\mathrm{Ent}_{2,\mathcal{N}}(X) &\le \mathrm{Ent}_{2,\mathcal{N}}(|X - E_{\mathcal{N}}[X]|_2) + 2\,\mathrm{Var}_{\mathcal{N}}(X) \\
&\quad + \|X - E_{\mathcal{N}}[X]\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}\,D\left(\sigma^{\frac{1}{2}}_{\mathrm{Tr}}(X - E_{\mathcal{N}}[X])_{\mathcal{N}}\,\sigma^{\frac{1}{2}}_{\mathrm{Tr}}\,\|\,\sigma^{\frac{1}{2}}_{\mathrm{Tr}}A\sigma^{\frac{1}{2}}_{\mathrm{Tr}}\right) \\
&\quad + \|E_{\mathcal{N}}[X]\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}\,D\left(\sigma^{\frac{1}{2}}_{\mathrm{Tr}}(E_{\mathcal{N}}[X])_{\mathcal{N}}\,\sigma^{\frac{1}{2}}_{\mathrm{Tr}}\,\|\,\sigma^{\frac{1}{2}}_{\mathrm{Tr}}A\sigma^{\frac{1}{2}}_{\mathrm{Tr}}\right) \\
&\quad - \|X\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}\,D\left(\sigma^{\frac{1}{2}}_{\mathrm{Tr}}X_{\mathcal{N}}\,\sigma^{\frac{1}{2}}_{\mathrm{Tr}}\,\|\,\sigma^{\frac{1}{2}}_{\mathrm{Tr}}A\sigma^{\frac{1}{2}}_{\mathrm{Tr}}\right) \\
&= \mathrm{Ent}_{2,\mathcal{N}}(|X - E_{\mathcal{N}}[X]|_2) + 2\,\mathrm{Var}_{\mathcal{N}}(X) \\
&\quad + \|X\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})} \times \\
&\quad \times \Big\{ p\,D\left(\sigma^{\frac{1}{2}}_{\mathrm{Tr}}(E_{\mathcal{N}}[X])_{\mathcal{N}}\,\sigma^{\frac{1}{2}}_{\mathrm{Tr}}\,\|\,\sigma^{\frac{1}{2}}_{\mathrm{Tr}}A\sigma^{\frac{1}{2}}_{\mathrm{Tr}}\right) + (1-p)\,D\left(\sigma^{\frac{1}{2}}_{\mathrm{Tr}}(X - E_{\mathcal{N}}[X])_{\mathcal{N}}\,\sigma^{\frac{1}{2}}_{\mathrm{Tr}}\,\|\,\sigma^{\frac{1}{2}}_{\mathrm{Tr}}A\sigma^{\frac{1}{2}}_{\mathrm{Tr}}\right) \\
&\quad - D\left(\sigma^{\frac{1}{2}}_{\mathrm{Tr}}X_{\mathcal{N}}\,\sigma^{\frac{1}{2}}_{\mathrm{Tr}}\,\|\,\sigma^{\frac{1}{2}}_{\mathrm{Tr}}A\sigma^{\frac{1}{2}}_{\mathrm{Tr}}\right),\Big\}
\end{aligned} \tag{8.25}$$

since by definition $\mathrm{Var}_{\mathcal{N}}(X) = \|X - E_{\mathcal{N}}[X]\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}$ and $\mathrm{Ent}_{2,\mathcal{N}}(E_{\mathcal{N}}[X]) = 0$, and where we also used the fact that for any $Z \in \mathcal{B}_{sa}(\mathcal{H})$, $Z_{\mathcal{N}} = (|Z|_2)_{\mathcal{N}}$. Remark that, since $(E_{\mathcal{N}}[X])_{\mathcal{N}} = E_{\mathcal{N}}[X]^2/\|E_{\mathcal{N}}[X]\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}$,

$$\begin{aligned}
X_{\mathcal{N}} &= \frac{E_{\mathcal{N}}\left[\sigma^{-1/2}_{\mathrm{Tr}}|\sigma^{1/4}_{\mathrm{Tr}}X\sigma^{1/4}_{\mathrm{Tr}}|^2\sigma^{-1/2}_{\mathrm{Tr}}\right]}{\|X\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}} \\
&= \frac{E_{\mathcal{N}}\left[\sigma^{-1/2}_{\mathrm{Tr}}(\sigma^{1/4}_{\mathrm{Tr}}X\sigma^{1/4}_{\mathrm{Tr}})^2\sigma^{-1/2}_{\mathrm{Tr}}\right]}{\|X\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}} \\
&= \frac{E_{\mathcal{N}}\left[\sigma^{-1/2}_{\mathrm{Tr}}(\sigma^{1/4}_{\mathrm{Tr}}(X - E_{\mathcal{N}}[X] + E_{\mathcal{N}}[X])\sigma^{1/4}_{\mathrm{Tr}})^2\sigma^{-1/2}_{\mathrm{Tr}}\right]}{\|X\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}} \\
&= \frac{E_{\mathcal{N}}\left[\sigma^{-1/2}_{\mathrm{Tr}}(\sigma^{1/4}_{\mathrm{Tr}}E_{\mathcal{N}}[X]\sigma^{1/4}_{\mathrm{Tr}})^2\sigma^{-1/2}_{\mathrm{Tr}}\right]}{\|X\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}} + \frac{E_{\mathcal{N}}\left[\sigma^{-1/2}_{\mathrm{Tr}}(\sigma^{1/4}_{\mathrm{Tr}}(X - E_{\mathcal{N}}[X])\sigma^{1/4}_{\mathrm{Tr}})^2\sigma^{-1/2}_{\mathrm{Tr}}\right]}{\|X\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}} \\
&\quad + \frac{E_{\mathcal{N}}\left[\sigma^{-1/4}_{\mathrm{Tr}}(X - E_{\mathcal{N}}[X])\sigma^{1/2}_{\mathrm{Tr}}E_{\mathcal{N}}[X]\sigma^{-1/4}_{\mathrm{Tr}}\right]}{\|X\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}} + \frac{E_{\mathcal{N}}\left[\sigma^{-1/4}_{\mathrm{Tr}}E_{\mathcal{N}}[X]\sigma^{1/2}_{\mathrm{Tr}}(X - E_{\mathcal{N}}[X])\sigma^{-1/4}_{\mathrm{Tr}}\right]}{\|X\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}} \\
&= p\,(E_{\mathcal{N}}[X])_{\mathcal{N}} + (1-p)\,(X - E_{\mathcal{N}}[X])_{\mathcal{N}},
\end{aligned} \tag{8.26}$$

where we used Pythagoras theorem $\|X\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})} = \|X - E_{\mathcal{N}}[X]\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})} + \|E_{\mathcal{N}}[X]\|^2_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}$, and where the last two terms in Equation (8.26) can be shown to be equal to zero using Equations (0.5) and (6.18),

since

$$E_{\mathcal{N}}\left[\sigma_{\mathrm{Tr}}^{-1/4}(X - E_{\mathcal{N}}[X])\sigma_{\mathrm{Tr}}^{1/2}E_{\mathcal{N}}[X]\sigma_{\mathrm{Tr}}^{-1/4}\right] = E_{\mathcal{N}}\left[\sigma_{\mathrm{Tr}}^{-1/4}(X - E_{\mathcal{N}}[X])\sigma_{\mathrm{Tr}}^{1/2}E_{\mathcal{N}}[X]\sigma_{\mathrm{Tr}}^{-1/2}\sigma_{\mathrm{Tr}}^{1/4}\right]$$

$$= E_{\mathcal{N}}\left[\sigma_{\mathrm{Tr}}^{-1/4}(X - E_{\mathcal{N}}[X])E_{\mathcal{N}}[\sigma_{\mathrm{Tr}}^{1/2}X\sigma_{\mathrm{Tr}}^{-1/2}]\sigma_{\mathrm{Tr}}^{1/4}\right]$$

$$= \sigma_{\mathrm{Tr}}^{-1/4}E_{\mathcal{N}}\left[(X - E_{\mathcal{N}}[X])E_{\mathcal{N}}[\sigma_{\mathrm{Tr}}^{1/2}X\sigma_{\mathrm{Tr}}^{-1/2}]\right]\sigma_{\mathrm{Tr}}^{1/4}$$

$$= \sigma_{\mathrm{Tr}}^{-1/4}E_{\mathcal{N}}\left[(X - E_{\mathcal{N}}[X])\right]E_{\mathcal{N}}[\sigma_{\mathrm{Tr}}^{1/2}X\sigma_{\mathrm{Tr}}^{-1/2}]\sigma_{\mathrm{Tr}}^{1/4}$$

$$= 0\,,$$

and similarly for the second term. Consequently, since $(X - E_{\mathcal{N}}[X])_{\mathcal{N}} = (|X - E_{\mathcal{N}}[X]|_2)_{\mathcal{N}}$, and by a use of the almost convexity of the von Neumann entropy (see Theorem 11.10 of [Nielsen and Chuang, 2010]), the term between brackets in (8.25) can be upper bounded by $H((p, 1-p))$, where H denotes the binary Shannon entropy. This is itself upper bounded by $\ln 2$, from which we get the result. $\qquad\square$

We can now easily prove Theorem 8.4.8.

Proof of Theorem 8.4.8 : This is a simple corollary of Proposition 8.4.9. Indeed, $\mathrm{LSI}_{2,\mathcal{N}}(c, d)$ applied to $|X - E_{\mathcal{N}}[X]|_2$ gives

$$\mathrm{Ent}_{2,\mathcal{N}}(|X - E_{\mathcal{N}}[X]|_2) \le c\,\mathcal{E}_{2,\mathcal{L}}(|X - E_{\mathcal{N}}[X]|_2) + 2d\,\mathrm{Var}_{\mathcal{N}}(X)$$

$$\le c\,\mathcal{E}_{2,\mathcal{L}}(X) + 2d\,\mathrm{Var}_{\mathcal{N}}(X)\,,$$

where we used that $\mathcal{E}_{2,\mathcal{L}}(|X - E_{\mathcal{N}}[X]|_2) \le \mathcal{E}_{2,\mathcal{L}}(X - E_{\mathcal{N}}[X]) = \mathcal{E}_{2,\mathcal{L}}(X)$ (see Theorem 4.7 of [Cipriani, 1997]). Besides, the DF-Poincaré inequality $\mathrm{PI}_{\mathcal{N}}(\lambda)$ implies $\lambda(\mathcal{L})\,\mathrm{Var}_{\mathcal{N}}(X) \le \mathcal{E}_{2,\mathcal{L}}(X)$. Consequently, we get by (8.24):

$$\mathrm{Ent}_{2,\mathcal{N}}(X) \le \left(c + 2\frac{d+1}{\lambda(\mathcal{L})}\right)\mathcal{E}_{2,\mathcal{L}}(X) + 2\ln\sqrt{2}\,\|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}^2\,,$$

which is the desired result.

$$\square$$

8.4.3. Bounding log-Sobolev constants via interpolation

The idea to use interpolation in order to obtain estimates on the log-Sobolev constants goes back to [Gross, 1975a]. The strategy can be summarised as follows: assume a bound of the form $\|P_{t_p} : \mathbb{L}_2(\sigma) \to \mathbb{L}_p(\sigma)\| \le M$ is known for some fixed $t_p \ge 0$ and $p > 2$, with $M \ge 1$. Then can one show by extrapolation from this bound that hypercontravitity holds for all $t \ge 0$? The answer is yes and its proof uses the crucial fact that the $\mathbb{L}_p$ norms used for the definition of hypercontractivity form an interpolating family of norms.

Theorem 8.4.10. *Let $(\mathcal{P}_t)_{t\ge0}$ be a faithful* KMS-*symmetric QMS on $\mathcal{B}(\mathcal{H})$ and assume that for some $2 < p \le +\infty$, there exist $t_p, M_p > 0$ such that for all X positive semidefinite, $\|\mathcal{P}_{t_p}(X)\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} \le M_p\|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}$. Then $\mathrm{LSI}_{2,\mathcal{N}}\left(\frac{2p\,t_p}{p-2}, \frac{p}{p-2}\ln M_p\right)$ holds.*

Proof. The proof follows closely the analogous statement for classical Markov chains [Diaconis and Saloff-Coste, 1996a] and primitive QMS [Temme et al., 2014]. The complex time semigroup

$$\mathcal{P}_z := \mathrm{e}^{z\mathcal{L}} = \sum_{n=0}^{\infty}\frac{z^n}{n!}\mathcal{L}^n\,, \qquad\qquad z \in \mathbb{C}\,,$$

defines an analytic family of operators. Define the time dilated complex semigroup $\tilde{\mathcal{P}}_z := \mathcal{P}_{t_p z}$. Since $(\mathcal{P}_t)_{t \geq 0}$ is KMS-symmetric, its spectral radius does not change upon the replacement $x \mapsto ix$, and therefore, for any $a > 0$ and X positive semidefinite:

$$\|\tilde{\mathcal{P}}_{ia}(X)\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})} \leq \|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}.$$

Therefore,

$$\|\tilde{\mathcal{P}}_{1+ia}(X)\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} = \|\tilde{\mathcal{P}}_1 \circ \tilde{\mathcal{P}}_{ia}(X)\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} \leq M_p \|\tilde{\mathcal{P}}_{ia}(X)\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})} \leq M_p \|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}.$$

Hence, by Stein-Weiss' interpolation Theorem ([Bergh and Löfström, 2012, Stein and Weiss, 2016]), for all $0 \leq s \leq 1$, and any $X \in \mathcal{B}(\mathcal{H})$:

$$\|\tilde{\mathcal{P}}_s(X)\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_{p_s}(\sigma_{\mathrm{Tr}}))} \leq M_p^s \|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})},$$

for p_s such that

$$\frac{1}{p_s} = \frac{s}{p} + \frac{1-s}{2}.$$

Taking $t = s t_p$ and $p(t) := p_s$, we get

$$\|\mathcal{P}_t(X)\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))} \leq e^{\frac{t}{t_p} \ln M_p} \|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}, \tag{8.27}$$

with equality at $t = 0$, where

$$p(t) = \frac{2 p t_p}{(2-p)t + p t_p}.$$

Taking derivatives on both sides of (8.27) at 0,

$$-\frac{\ln M_p}{t_p} \|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})} + \frac{d}{dt} \|\mathcal{P}_t(X)\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))}\Big|_{t=0} \leq 0. \tag{8.28}$$

Using Corollary 8.3.4, with $p(0) = 2$ and $p'(0) = \frac{2(p-2)}{p t_p}$,

$$\frac{d}{dt} \|\mathcal{P}_t(X)\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))}\Big|_{t=0} = \frac{p-2}{p t_p \|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}} \left[\frac{1}{2} \mathrm{Ent}_{2, \mathcal{N}}(X) - \frac{p t_p}{p-2} \mathcal{E}_{2, \mathcal{L}}(X) \right].$$

Hence, (8.28) can be rewritten as

$$\frac{p-2}{2 p t_p} \mathrm{Ent}_{2, \mathcal{N}}(X) \leq \mathcal{E}_{2, \mathcal{L}}(X) + \frac{\ln M_p}{t_p} \|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}^2, \tag{8.29}$$

which leads to the desired result. $\qquad\square$

In the following corollary, we combine Theorem 8.2.5 and Theorem 8.4.10 to further provide upper bounds on the log-Sobolev constants in terms of the spectral gap of the QMS $(\mathcal{P}_t)_{t \geq 0}$. As such, it can be seen as an extension of Theorem 5 of [Temme et al., 2014] to the case of faithful KMS-symmetric QMS.

Corollary 8.4.11. *Given a* KMS-*symmetric QMS* $(\mathcal{P}_t)_{t \geq 0}$ *with spectral gap* $\lambda(\mathcal{L})$, $\mathrm{LSI}_{2, \mathcal{N}}(c, \ln\sqrt{2})$

holds, with

$$c \le \frac{\ln(\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty) + 2}{\lambda(\mathcal{L})}.$$

Proof. From (1.20), we get that for any $X \ge 0$,

$$\|X\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_4(\sigma_{\mathrm{Tr}}))} \le \|X\|_{\mathbb{L}_4(\sigma_{\mathrm{Tr}})} \le \|\sigma_{\mathrm{Tr}}^{-1}\|_\infty^{1/4} \|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})},$$

where the last inequality is a well-known property of $\mathbb{L}_p$ norms. Together with the contractivity of $(\mathcal{P}_t)_{t \ge 0}$ (cf. (i) of Proposition 8.3.1), we find

$$\|\mathcal{P}_t(X)\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_4(\sigma_{\mathrm{Tr}}))} \le \|X\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_4(\sigma_{\mathrm{Tr}}))} \le \|\sigma_{\mathrm{Tr}}^{-1}\|_\infty^{1/4} \|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}.$$

We conclude with successive applications of Theorem 8.4.10 and Theorem 8.2.5, taking the limit $t_4 \to 0$ and $M_4 = \|\sigma_{\mathrm{Tr}}^{-1}\|_\infty^{1/4}$. $\qquad\square$

8.5. Non-positivity of the strong LSI constant

In this section, we show that a strong DF-log-Sobolev inequality does not hold for a non-trivially faithful QMS, that is a QMS that is neither primitive nor unitary. We deduce from this that the amalgamated $\mathbb{L}_p$ norms do not satisfy uniform convexity for $1 \le p \le 2$ as soon as $\mathcal{N}$ is non-trivial.

By comparison of Dirichlet forms, it is enough to consider the case of the $\mathcal{N}$-decoherent QMS defined by $\mathcal{L}_\mathcal{N} := E_\mathcal{N} - \mathrm{id}$, where $\mathcal{N}$ is any $*$-subalgebra of $\mathcal{B}(\mathcal{H})$ and $E_\mathcal{N}$ is a conditional expectation on it. Indeed, if $(\mathcal{P}_t)_{t \ge 0}$ is any faithful QMS with DF-algebra $\mathcal{N}$ and the same conditional expectation $E_\mathcal{N}$, then the following inequality holds [Müller-Hermes et al., 2016][2]:

$$\lambda(\mathcal{L}) \, \mathcal{E}_{2, \mathcal{L}_\mathcal{N}}(X) \le \mathcal{E}_{2, \mathcal{L}}(X) \le \left\| \frac{\mathcal{L} + \hat{\mathcal{L}}}{2} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \right\| \mathcal{E}_{2, \mathcal{L}_\mathcal{N}}(X),$$

where $\hat{\mathcal{L}}$ is the conjugate of $\mathcal{L}$ with respect to $\langle .,. \rangle_{\sigma_{\mathrm{Tr}}}$, and $\lambda(\mathcal{L})$ is the spectral gap of $(\mathcal{P}_t)_{t \ge 0}$. From this inequality we directly obtain that if $\mathrm{LSI}_{2, \mathcal{N}}(c_\mathcal{N}, 0)$ holds for $\mathcal{L}_\mathcal{N}$, then $\mathrm{LSI}_{2, \mathcal{N}}(c, 0)$ holds for $\mathcal{L}$ with:

$$0 < \frac{\lambda(\mathcal{L})}{c_\mathcal{N}} \le \frac{1}{c} \le \left\| \frac{\mathcal{L} + \hat{\mathcal{L}}}{2} : \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \to \mathbb{L}_2(\sigma_{\mathrm{Tr}}) \right\| \frac{1}{c_\mathcal{N}}.$$

Our goal is thus to show that if $\mathcal{N}$ is non-trivial and $\mathrm{LSI}_{2, \mathcal{N}}(c_\mathcal{N}, 0)$ holds for $\mathcal{L}_\mathcal{N}$, then $c_\mathcal{N} = +\infty$.

Theorem 8.5.1. *Let $\mathcal{N}$ be any non-trivial $*$-subalgebra of $\mathcal{B}(\mathcal{H})$ (that is, $\mathcal{N} \ne \mathbb{C}\mathbb{1}$ and $\mathcal{N} \ne \mathcal{B}(\mathcal{H})$) and consider the Lindbladian $\mathcal{L}_\mathcal{N} := E_\mathcal{N} - \mathrm{id}$, where $E_\mathcal{N}$ is any conditional expectatin on $\mathcal{N}$. Define $\sigma_{\mathrm{Tr}} := E_{\mathcal{N}*}(d_\mathcal{H}^{-1} \mathbb{1}_\mathcal{H})$. Assume that there exists $\alpha \ge 0$ such that for all positive semi-definite $X \in \mathcal{B}(\mathcal{H})$,*

$$\alpha \, \mathrm{Ent}_{2, \mathcal{N}}(X) \le \mathcal{E}_{2, \mathcal{L}_\mathcal{N}}(X). \tag{8.30}$$

Then $\alpha = 0$.

Proof. Let $\alpha \ge 0$ be such that inequality (8.30) holds for all positive semi-definite $X \in \mathcal{B}(\mathcal{H})$. We

---

[2]A proof of the lower bound is provided in the proof of Corollary 10.1.8

shall construct a sequence $(Z_k)_{k\in\mathbb{N}}$ such that

$$\frac{\mathcal{E}_{2,\mathcal{L}_\mathcal{N}}(Z_k)}{\mathrm{Ent}_{2,\mathcal{N}}(Z_k)} \xrightarrow[k\to\infty]{} 0 \,,$$

which directly implies that $\alpha = 0$. More precisely, we shall construct a sequence of density matrices $(\rho_k)_{k\geq 1}$ such that $Z_k = \Gamma_{\sigma_{\mathrm{Tr}}}^{-\frac{1}{2}}(\sqrt{\rho_k})$ and

$$\frac{\mathcal{E}_{2,\mathcal{L}_\mathcal{N}}(\sigma_{\mathrm{Tr}}^{-1/4}\sqrt{\rho_k}\,\sigma_{\mathrm{Tr}}^{-1/4})}{D(\rho_k\|\rho_{\mathcal{N},k})} \xrightarrow[k\to\infty]{} 0 \,, \tag{8.31}$$

where $\rho_{\mathcal{N},k} := E_{\mathcal{N}*}(\rho_k)$. Now assume that $\mathcal{H}$ and $\mathcal{N}$ admit the decomposition given by Equation (0.8) and Equation (0.9). As $\mathcal{N}$ is non-trivial, we can assume that either there exists $i \in I$ such that $\dim \mathcal{H}_i > 1$ and $\dim \mathcal{K}_i > 1$, or $|I| > 1$. We shall construct a sequence $(\rho_k)_{k\geq 1}$ in each case and then treat them simultaneously to prove the limit in (8.31).

We start by considering the first case and, without loss of generality, we assume that $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ and that $\mathcal{N} = \mathcal{B}(\mathcal{H}_A) \otimes \mathbb{1}_{\mathcal{H}_B}$, with $\dim \mathcal{H}_B := d_B > 1$ and $\dim \mathcal{H}_A = 2$. One can recover the general case by adding zeros in the corresponding entries of ρ_k. Then, it means that there exists a density matrix $\tau \in \mathcal{D}(\mathcal{H}_B)$ such that for all $\omega \in \mathcal{S}_1(\mathcal{H})$,

$$E_{\mathcal{N}*}(\omega) = \mathrm{Tr}_{\mathcal{H}_B}(\omega) \otimes \tau \,.$$

We define, in an orthonormal basis in which τ is diagonal and in any orthonormal basis of $\mathcal{H}_A$,

$$\Delta := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & & \cdots & 0 \\ 0 & & \ddots & & \vdots \\ \vdots & & & & \\ 0 & & \cdots & & 0 \end{pmatrix}}_{d_B}, \qquad \rho_{\mathcal{N},k} = \begin{pmatrix} \frac{1}{k} & 0 \\ 0 & 1-\frac{1}{k} \end{pmatrix} \otimes \tau \,.$$

It is clear that $E_{\mathcal{N}*}(\Delta) = 0$. Next, define

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

so that $\langle e_i, \Delta e_j \rangle = 1 - \delta_{ij}$. We also define $\lambda_1 := k\langle e_1, \rho_{\mathcal{N},k}\, e_1 \rangle$ and $\lambda_2 := \frac{k}{k-1}\langle e_2, \rho_{\mathcal{N}_k}\, e_2 \rangle$, which clearly do not depend on k. We now set, for $\varepsilon \geq 0$,

$$\rho_{k,\varepsilon} := \rho_{\mathcal{N},k} + \varepsilon \Delta \,,$$

so that $E_{\mathcal{N}*}(\rho_{k,\varepsilon}) = \rho_{\mathcal{N},k}$. Since the $\rho_{\mathcal{N},k}$ are full-rank, the $\rho_{k,\varepsilon}$ are well-defined density matrices for ε small enough.

We now turn to the case where $|I| > 1$. Up to adding zero entries in the matrices defining ρ_k, we can assume that $|I| = 2$. Denote by P_i the orthogonal projection on $\mathcal{H}_i \otimes \mathcal{K}_i$ for $i \in I$, and consider $\eta_i = \frac{\mathbb{1}_{\mathcal{H}_i}}{\dim \mathcal{H}_i} \otimes \tau_i$. We also denote by $e_i \in \mathcal{H}_i \otimes \mathcal{K}_i$ an eigenvector of η_i of associated eigenvalue $\lambda_i > 0$.

We then set

$$\Delta = |e_1\rangle\langle e_2| + |e_2\rangle\langle e_1|, \qquad \rho_{\mathcal{N},k} = \frac{1}{k}\eta_1 + \left(1 - \frac{1}{k}\right)\eta_2,$$

so that again $E_{\mathcal{N}*}(\Delta) = 0$ and $\langle e_i, \Delta\, e_j\rangle = 1 - \delta_{ij}$. As before, we define $\rho_{k,\varepsilon} := \rho_{\mathcal{N},k} + \varepsilon\,\Delta$.

Remark that in both cases, we have $E_{\mathcal{N}*}(\Delta) = 0$ and

$$
\begin{aligned}
\Delta &= |e_1\rangle\langle e_2| + |e_2\rangle\langle e_1|, & \langle e_i, \Delta\, e_j\rangle &= 1 - \delta_{ij}, \\
\lambda_1 &:= k\,\langle e_1, \rho_{\mathcal{N},k}\, e_1\rangle, & \lambda_2 &:= \frac{k}{k-1}\langle e_2, \rho_{\mathcal{N}_k}\, e_2\rangle.
\end{aligned}
\tag{8.32}
$$

This will be enough to treat both cases simultaneously. We shall now prove that the limit in (8.31) holds with $\rho_k = \lim_{\varepsilon \to 0} \rho_{k,\varepsilon}$. The first step is to obtain a limit for a fixed $k \geq 1$ and $\varepsilon \to 0$, that is, to obtain a continuous extension of the quotient appearing in the limit at $\rho_{\mathcal{N},k}$. For this purpose, we compute the Taylor expansion of both the numerator and the denominator. A simple calculation using the integral representations of the logarithm and of the square root functions [Hiai et al., 2011] shows that (see also the proofs of Theorem 16 in [Kastoryano and Temme, 2013] and Lemma 3.5 in [Bardet, 2017]):

$$
\begin{aligned}
D(\rho_{k,\varepsilon} \| \rho_{\mathcal{N},k}) &= \varepsilon^2 \int_0^\infty \mathrm{Tr}\left[\Delta\,\frac{1}{t - \rho_{\mathcal{N},k}}\,\Delta\,\frac{1}{t - \rho_{\mathcal{N},k}}\right] dt + \mathcal{O}(\varepsilon^3) \\
\mathcal{E}_{2,\mathcal{N}}(\sigma_{\mathrm{Tr}}^{-1/4}\sqrt{\rho_{k,\varepsilon}}\,\sigma_{\mathrm{Tr}}^{-1/4}) &= \pi^2\varepsilon^2 \iint_{[0,\infty)^2} \sqrt{s\,t}\,\mathrm{Tr}\left[\frac{1}{t + \rho_{\mathcal{N},k}}\,\Delta\,\frac{1}{t + \rho_{\mathcal{N},k}}\,\Delta\right] \\
&\quad - \pi^2\varepsilon^2 \iint_{[0,\infty)^2} \sqrt{s\,t}\,\mathrm{Tr}\left[\frac{\sigma_{\mathrm{Tr}}^{1/4}}{t + \rho_{\mathcal{N},k}}\,\Delta\,\frac{\sigma_{\mathrm{Tr}}^{1/4}}{t + \rho_{\mathcal{N},k}}\,E_{\mathcal{N}}\left[\frac{\sigma_{\mathrm{Tr}}^{-1/4}}{s + \rho_{\mathcal{N},k}}\,\Delta\,\frac{\sigma_{\mathrm{Tr}}^{-1/4}}{s + \rho_{\mathcal{N},k}}\right]\right] + \mathcal{O}(\varepsilon^3).
\end{aligned}
$$

Using Equation (8.32) we can compute explicitly these integrals. For instance, the second integral in the second equation is null, since $E_{\mathcal{N}}\left[\frac{\sigma_{\mathrm{Tr}}^{-1/4}}{s+\rho_{\mathcal{N},k}}\,\Delta\,\frac{\sigma_{\mathrm{Tr}}^{-1/4}}{s+\rho_{\mathcal{N},k}}\right] = 0$. This can be checked directly using the fact that both e_1 and e_2 are eigenvectors of $\frac{\sigma_{\mathrm{Tr}}^{-1/4}}{s+\rho_{\mathcal{N},k}}$ and that $E_{\mathcal{N}}[|e_1\rangle\langle e_2|] = E_{\mathcal{N}}[|e_2\rangle\langle e_1|] = 0$. We thus obtain:

$$D(\rho_{k,\varepsilon} \| \rho_{\mathcal{N},k}) = \varepsilon^2\,g\left(\frac{1}{k}\lambda_1, \left(1 - \frac{1}{k}\right)\lambda_2\right) |\langle e_1|\Delta|e_2\rangle|^2 + \mathcal{O}(\varepsilon^3), \tag{8.33}$$

$$\mathcal{E}_{2,\mathcal{N}}(\sigma_{\mathrm{Tr}}^{-1/4}\sqrt{\rho_{k,\varepsilon}}\,\sigma_{\mathrm{Tr}}^{-1/4}) = 2\pi^2\,\varepsilon^2\,f\left(\frac{1}{k}\lambda_1, \left(1 - \frac{1}{k}\right)\lambda_2\right) |\langle e_1|\Delta|e_2\rangle|^2 + \mathcal{O}(\varepsilon^3), \tag{8.34}$$

where

$$
f(x,y) := \begin{cases} \dfrac{(\sqrt{x} - \sqrt{y})^2}{(x-y)^2} & \text{if } x \neq y \\[2mm] \dfrac{1}{4\,x} & \text{else} \end{cases}, \qquad
g(x,y) := \begin{cases} \dfrac{\ln(x) - \ln(y)}{x - y} & \text{if } x \neq y \\[2mm] \dfrac{1}{x} & \text{else}. \end{cases}
\tag{8.35}
$$

For a fixed $k \geq 1$, we thus obtain that

$$\frac{\mathcal{E}_{2,\mathcal{L}_{\mathcal{N}}}(\sigma_{\mathrm{Tr}}^{-1/4}\sqrt{\rho_{k,\varepsilon}}\,\sigma_{\mathrm{Tr}}^{-1/4})}{D(\rho_{k,\varepsilon} \| \rho_{\mathcal{N},k})} \xrightarrow[\varepsilon \to 0]{} 2\pi^2\,\frac{f\left(\frac{1}{k}\lambda_1, (1 - \frac{1}{k})\lambda_2\right)}{g\left(\frac{1}{k}\lambda_1, (1 - \frac{1}{k})\lambda_2\right)}.$$

We just have to take the limit $k \to +\infty$ to conclude. Indeed,

$$f\left(\frac{1}{k}\lambda_1, (1-\frac{1}{k})\lambda_2\right) \underset{k\to+\infty}{\longrightarrow} 1/\lambda_2\,,$$
$$g\left(\frac{1}{k}\lambda_1, (1-\frac{1}{k})\lambda_2\right) \underset{k\to+\infty}{\longrightarrow} +\infty\,.$$

$\square$

The above result implies the following straightforward corollary:

Corollary 8.5.2. *The $\mathbb{L}_2\left(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}})\right)$ spaces do not satisfy the uniform convexity property.*

8.6. Application to decoherence times

In this section, we apply the framework of DF-log-Sobolev inequalities in order to find bounds on the decoherence rates of a non-primitive quantum Markov semigroup. We recall that, for $0 < \varepsilon < 1$, the decoherence time of a KMS-symmetric QMS $(\mathcal{P}_t)_{t\geq 0}$ is defined as

$$\tau_{\mathrm{deco}}(\varepsilon) := \inf\left\{t \geq 0 : \ \|\mathcal{P}_{t*}(\rho - \rho_{\mathcal{N}})\|_1 \leq \varepsilon\right\}\,,$$

where $\rho_{\mathcal{N}} \equiv E_{\mathcal{N}*}(\rho)$. A classical technique to get rapid decoherence for all times comes from looking at the spectral gap of a KMS-symmetric QMS:

$$\|\mathcal{P}_t\left(X - E_{\mathcal{N}}[X]\right)\|_\infty \leq \|\sigma_{\mathrm{Tr}}^{-1}\|_\infty^{\frac{1}{2}} \|\mathcal{P}_t\left(X - E_{\mathcal{N}}[X]\right)\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}$$
$$\leq \|\sigma_{\mathrm{Tr}}^{-1}\|_\infty^{\frac{1}{2}} \, \mathrm{e}^{-\lambda(\mathcal{L})t} \|X - E_{\mathcal{N}}(X)\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})}\,,$$

where the second inequality follows from (8.3). In the dual Schrödinger picture, such a bound translates into

$$\|\mathcal{P}_{*t}(\rho - \rho_{\mathcal{N}})\|_1 \leq \|\sigma_{\mathrm{Tr}}^{-1}\|_\infty^{1/2} \, \mathrm{e}^{-\lambda(\mathcal{L})t}\,. \tag{8.36}$$

However, as discussed in Section 3.2, already in the classical case the spectral gap does not usually provide tight enough bounds on the decoherence time of a Markov semigroup [Diaconis and Saloff-Coste, 1996a]. Moreover, in practice, the coefficient $\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty^{\frac{1}{2}}$ explodes exponentially fast as the dimension of the system grows. If $\mathrm{LSI}_{2,\mathcal{N}}(c,0)$ held with $c < \infty$, the original techniques of [Temme et al., 2014] could be adapted to yield

$$\|\mathcal{P}_{*t}(\rho - \rho_{\mathcal{N}})\|_1 \leq (2\ln\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty)^{1/2} \, \mathrm{e}^{-\frac{4t}{c}}\,,$$

improving significantly the bound (8.36) derived from the spectral gap method. However, as discussed in the last section, a strong LSI never holds for non-primitive QMS. This motivates the search for a technique that would deal with the weak version of the log-Sobolev inequality. Fortunately, such a technique already exists in the classical literature [Zegarlinski, 1995, Martinelli, 1999, Diaconis and Saloff-Coste, 1996a]: it consists in combining hypercontractivity bounds at short times with the spectral gap at long times (see estimate (3.13)). Using such a method, we can prove the exponential convergence in terms of the ∞-norm.

Proposition 8.6.1. *Assume that a faithful QMS $(\mathcal{P}_t)_{t\geq 0}$ of corresponding decoherence-free algebra $\mathcal{N}$*

satisfies $\mathrm{HC}_{2,\mathcal{N}}(c,d)$, *and that* $\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty \geq \mathrm{e}$. *Then for* $t = \frac{c}{4}\ln\ln\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty + \frac{\kappa}{\lambda(\mathcal{L})}$, $\kappa > 0$, *and all* $X \in \mathcal{B}(\mathcal{H})$

$$\|\mathcal{P}_t (X - E_\mathcal{N}[X])\|_\infty \leq \left(\max_{i \in I} \sqrt{d_{\mathcal{H}_i}}\right) \mathrm{e}^{1+d-\kappa} \|X\|_\infty, \tag{8.37}$$

where $d_{\mathcal{H}_i}$ *denote the dimensions of the spaces* $\mathcal{H}_i$ *appearing in the decomposition of* $\mathcal{N}$ *given by* (0.9). *By duality, we get the following similar bound:*

$$\forall \rho \in \mathcal{D}(\mathcal{H}), \qquad \|\mathcal{P}_{*t}(\rho - E_{\mathcal{N}*}(\rho))\|_1 \leq \max_{i \in I} \sqrt{d_{\mathcal{H}_i}} \, \mathrm{e}^{1+d-\kappa}. \tag{8.38}$$

The above inequality provides a bound on the decoherence time of the QMS:

$$\tau_{\mathrm{deco}}(\varepsilon) \leq \frac{\ln\left(\max_{i \in I} \sqrt{d_{\mathcal{H}_i}}\, \varepsilon^{-1}\right) + 1 + d}{\lambda(\mathcal{L})} + \frac{c}{4}\ln\ln\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty.$$

Proof. Let $t, s > 0$. Then:

$$\begin{aligned}
\|\mathcal{P}_{t+s}(X - E_\mathcal{N}[X])\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_\infty(\sigma_{\mathrm{Tr}}))} &\leq \|{\sigma_{\mathrm{Tr}}}^{-1}\|_\infty^{\frac{1}{p}} \|\mathcal{P}_{t+s}(X - E_\mathcal{N}[X])\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} \\
&\leq \|{\sigma_{\mathrm{Tr}}}^{-1}\|_\infty^{\frac{1}{p}} \exp\left(2d\left(\frac{1}{2} - \frac{1}{p}\right)\right) \|\mathcal{P}_t(X - E_\mathcal{N}[X])\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})} \\
&\leq \|{\sigma_{\mathrm{Tr}}}^{-1}\|_\infty^{\frac{1}{p}} \exp\left(2d\left(\frac{1}{2} - \frac{1}{p}\right)\right) \|X - E_\mathcal{N}[X]\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})} \, \mathrm{e}^{-\lambda(\mathcal{L})t} \\
&\leq \mathrm{e}^d \, \|\sigma_{\mathrm{Tr}}^{-1}\|_\infty^{\frac{1}{p}} \|X - E_\mathcal{N}[X]\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})} \, \mathrm{e}^{-\lambda(\mathcal{L})t},
\end{aligned}$$

where the first inequality follows from (1.28) in Section 1.1.2 applied to $\mathcal{P}_{t+s}(X - E_\mathcal{N}[X])$, the second inequality from $\mathrm{HC}_{2,\mathcal{N}}(c,d)$, and the third one by definition of the spectral gap. Since $\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty \geq \mathrm{e}$, one can choose $s := \frac{c}{4}\ln\ln\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty$, and $p \equiv p(s) = 1 + \ln\|\sigma_{\mathrm{Tr}}^{-1}\|_\infty$, so that

$$\|\mathcal{P}_{t+s}(X - E_\mathcal{N}[X])\|_{\mathbb{L}_2(\mathcal{N}, \mathbb{L}_\infty(\sigma_{\mathrm{Tr}}))} \leq \|X\|_\infty \, \mathrm{e}^{1+d-\lambda(\mathcal{L})\,t},$$

where we use that $\|X - E_\mathcal{N}[X]\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})} \leq \|X\|_{\mathbb{L}_2(\sigma_{\mathrm{Tr}})} \leq \|X\|_\infty$. The result follows by applying the following norm estimate proved in Proposition 1.1.13:

$$\|\mathrm{id} : \mathbb{L}_2(\mathcal{N}, \mathbb{L}_\infty(\sigma_{\mathrm{Tr}})) \to \mathcal{B}(\mathcal{H})\| \leq \max_{i \in I} \sqrt{d_{\mathcal{H}_i}}.$$

By duality, we get,

$$\begin{aligned}
\left\|\mathcal{P}_{(t+s)*}(\rho - \rho_\mathcal{N})\right\|_1 &= \sup_{\|X\|_\infty \leq 1} \mathrm{Tr}\left(\mathcal{P}_{(t+s)*}(\rho - E_{\mathcal{N}*}(\rho)) X\right) \\
&= \sup_{\|X\|_\infty \leq 1} \mathrm{Tr}\left(\rho \, \mathcal{P}_{(t+s)}(X - E_\mathcal{N}[X])\right) \\
&\leq \max_{i \in I} \sqrt{d_{\mathcal{H}_i}} \, \mathrm{e}^{1+d-\lambda(\mathcal{L})\,t}.
\end{aligned}$$

$\square$

8.A. Proof of Theorem 8.3.3

Let $t \mapsto X(t) \in \mathcal{B}(\mathcal{H})$ be an operator-valued twice continuously differentiable function, where $X(t) > 0$ for all $t \in [-\eta, \eta]$, for some $\eta > 0$, as well as an increasing twice continuously differentiable function

$\mathbb{R} \ni t \mapsto p(t)$ with $p(0) = q \geq 1$. Define

$$s(t) := \frac{1}{q} - \frac{1}{p(t)} \,,$$

and for a positive definite operator $A \in \mathcal{N}$, such that $\|A\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1$,

$$M(t, A) := A^{-s(t)/2} X(t) A^{-s(t)/2} \,.$$

Thus $M(t, A)$ is positive definite for any $t \in [-\eta, \eta]$. Define moreover

$$\Phi(X(t), A, p(t)) := \|M(t, A)\|_{p(t)} \,. \tag{8.39}$$

The following proposition gathers straightforward generalization of results proved in [Beigi and King, 2016] which were used to prove the relation between hypercontractivity and the log-Sobolev inequality for the completely bounded norm. (cf. lemmas 8, 9 of [Beigi and King, 2016]). We recall that $\mathcal{S}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$ denotes the set of positive definite operators on the sphere of radius one in $\mathbb{L}_1(\sigma_{\mathrm{Tr}})$.

Proposition 8.A.1. *For a fixed $t \in (-\eta, \eta)$, $A \mapsto \Phi(X(t), A, p(t))^{p(t)}$ is convex for $1 \leq q \leq p(t) \leq 2q$ and concave for $1 \leq p(t) \leq q$. Moreover, the following assertions hold true:*

1 The function $(t, A) \mapsto \frac{\partial^2}{\partial t^2} \Phi(X(t), A, p(t))$ is continuous on $(-\eta, \eta) \times \mathcal{N}(\mathcal{P}) \cap \mathcal{S}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$.

2 The function $A \mapsto \Phi(X(t), A, p(t))$ is continuously differentiable for all $A \in \mathcal{N}(\mathcal{P}) \cap \mathcal{S}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$.

3 For all $A \in \mathcal{N}(\mathcal{P}) \cap \mathcal{S}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$ and $t \in (-\eta, \eta)$,

$$\begin{aligned}
\frac{\partial}{\partial t} \Phi(X(t), A, p(t)) =& \frac{p'(t) \Phi(X(t), A, p(t))}{p(t)^2 \operatorname{Tr}\left[M(t, A)^{p(t)}\right]} \Big(-\operatorname{Tr}\left[M(t, A)^{p(t)}\right] \ln \operatorname{Tr}\left[M(t, A)^{p(t)}\right] \\
&+ \operatorname{Tr}\left[M(t, A)^{p(t)} \ln M(t, A)^{p(t)}\right] - \operatorname{Tr}\left[M(t, A)^{p(t)} \ln A\right] \\
&+ \frac{p(t)^2}{p'(t)} \operatorname{Tr}\left[M(t, A)^{p(t)-1} A^{-s(t)/2} X'(t) A^{-s(t)/2}\right] \Big) \,.
\end{aligned} \tag{8.40}$$

In what follows, we fix a positive definite $Y \in \mathcal{B}(\mathcal{H})$ and set $X(t) = \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p(t)}}(Y(t))$, where $t \mapsto Y(t)$ is some twice continuously differentiable matrix-valued function with $Y(0) = Y$. Therefore,

$$\frac{d}{dt} X(t)\Big|_{t=0} = \frac{d}{dt} \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p(t)}}(Y(t))\Big|_{t=0} = -\frac{p'(0)}{2q^2} \left\{ \ln \sigma_{\mathrm{Tr}}, \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y(0)) \right\} + \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y'(0)) \,,$$

where we used that $p(0) = q$ and where $\{\cdot, \cdot\}$ is the anticommutator. Thus, using that $M(0, A) = \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y(0))$ and that $\Phi(X(0), A, q) = \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}$, Equation (8.40) reduces to

$$\begin{aligned}
\frac{\partial}{\partial t} \Phi(X(t), A, p(t))\Big|_{t=0} =& \frac{p'(0)}{q^2 \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}} \Big[-\|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q \ln \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q + \operatorname{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^q \ln \left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^q \right) \\
&- \operatorname{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^q \ln A \right) - \operatorname{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^q \ln \sigma_{\mathrm{Tr}} \right) + \frac{q^2}{p'(0)} \operatorname{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^{q-1} \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y'(0)) \right) \Big] \,.
\end{aligned} \tag{8.41}$$

In fact, in the case when $Y(t) = Y \in \mathcal{B}_{sa}(\mathcal{H})$, and $p(t) = q + t$, one can similarly show the following

$$\frac{\partial}{\partial p} \Phi(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(Y), A, p)\Big|_{p=q} = \frac{1}{q^2 \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}} \left[-\|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q \ln \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q + \mathrm{Tr}\left(\left| \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y) \right|^q \ln \left| \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y) \right|^q \right) \right.$$
$$\left. - \mathrm{Tr}\left(\left| \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y) \right|^q \ln A \right) - \mathrm{Tr}\left(\left| \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y) \right|^q \ln \sigma_{\mathrm{Tr}} \right) \right]. \tag{8.42}$$

Now, define $G(A)$ as the part in the parenthesis:

$$G(A) := -\|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q \ln \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q + \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y) \right]^q \ln \left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y) \right]^q \right)$$
$$- \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y) \right]^q \ln A \right) - \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y) \right]^q \ln \sigma_{\mathrm{Tr}} \right)$$
$$+ \frac{q^2}{p'(0)} \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y) \right]^{q-1} \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y'(0)) \right), \tag{8.43}$$

and let, for a given $Y \in \mathcal{B}_{sa}(\mathcal{H})$,

$$Y_{\mathcal{N}} := \frac{E_{\mathcal{N}}\left[\mathbb{I}_{1,q}(Y) \right]}{\|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q}. \tag{8.44}$$

Next, we derive a formula that will be useful in what follows.

Lemma 8.A.2. *With the above notations and for positive semidefinite $Y \in \mathcal{B}(\mathcal{H})$,*

$$G(A) - G(Y_{\mathcal{N}}) = \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q \ D(\Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}}) \| \Gamma_{\sigma_{\mathrm{Tr}}}(A)). \tag{8.45}$$

Remark that $G(A) - G(Y_{\mathcal{N}})$ does not depend on $Y'(0)$ and therefore one can check that the same result holds for $Y \in \mathcal{B}_{sa}(\mathcal{H})$.

Proof. First note that

$$G(A) - G(Y_{\mathcal{N}}) = \mathrm{Tr}\left(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)(\ln Y_{\mathcal{N}} - \ln A) \right).$$

As $Y_{\mathcal{N}}$ and A are in $\mathcal{N}$, they commute with σ_{Tr} and therefore we get

$$G(A) - G(Y_{\mathcal{N}}) = \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q \ \mathrm{Tr}\left(\frac{\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q}{\|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q} \left(\ln \Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}}) - \ln \Gamma_{\sigma_{\mathrm{Tr}}}(A) \right) \right).$$

Now, as again $Y_{\mathcal{N}}, A \in \mathcal{N}(\mathcal{P})$, $\ln Y_{\mathcal{N}}$ and $\ln A$ also belong to $\mathcal{N}(\mathcal{P})$ and we get

$$G(A) - G(Y_{\mathcal{N}}) = \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q \ \mathrm{Tr}\left(\frac{\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q}{\|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q} E_{\mathcal{N}}\left[\ln \Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}}) - \ln \Gamma_{\sigma_{\mathrm{Tr}}}(A) \right] \right)$$
$$= \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q \ \mathrm{Tr}\left(E_{\mathcal{N}*}\left(\frac{\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q}{\|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q} \right) \left(\ln \Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}}) - \ln \Gamma_{\sigma_{\mathrm{Tr}}}(A) \right) \right)$$
$$= \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q \ \mathrm{Tr}\left(\Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}}) \left(\ln \Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}}) - \ln \Gamma_{\sigma_{\mathrm{Tr}}}(A) \right) \right),$$

which is the desired result. $\qquad \square$

Theorem 8.3.3 follows from a direct adaptation of the proof of Theorem 7 of [Beigi and King, 2016]. In a nutshell, all the lemmas used in [Beigi and King, 2016] to prove it can be generalized to

our framework, when replacing the equation (25) of [Beigi and King, 2016] by Equation (8.45). In particular, one can prove that

$$\Delta(t) := \frac{1}{t}\left(\|Y(t)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))} - \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}\right) - \frac{G(Y_\mathcal{N})p'(0)}{q^2\|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}}.$$

converges to 0, which leads to the desired result. The details are provided for sake of completeness.

Lemma 8.A.3. *There exist $\kappa > 0$ and $K < \infty$, such that for all $t \in [-\eta/2, \eta/2]$ and $A \in \mathcal{S}(\kappa) :=$ $\mathcal{N}(\mathcal{P}) \cap \{B > 0, \ \|B\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1, \ \|B - Y_\mathcal{N}\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} \le \kappa\}$,*

$$\left|\Phi(X(t), A, p(t)) - \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} - t \ \frac{p'(0)G(A)}{q^2\|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}}\right| \le Kt^2 \,,$$

where $X(t) = \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p(t)}}(Y(t))$.

Proof. The proof is similar to the one of Lemma 10 of [Beigi and King, 2016]. Let $t \in [-\eta/2, \eta/2]$. Since the set $\mathcal{N}(\mathcal{P}) \cap \{B > 0, \ \|B\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1\}$ is open, there exists $\kappa > 0$ such that $\mathcal{S}(\kappa)$ is a compact and a subset of $\mathcal{N}(\mathcal{P}) \cap \{B > 0, \ \|B\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1\}$. By Proposition 8.A.1, the function $\frac{\partial^2}{\partial t^2}\Phi(X(t), A, p(t))$ is continuous on $(-\eta, \eta) \times \mathcal{N}(\mathcal{P}) \cap \mathcal{S}_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^+$. Hence, there exists $K > \infty$ such that

$$-2K \le \frac{\partial^2\Phi(X(t), A, p(t))}{\partial t^2} \le 2K \,,$$

for all $t \in [-\eta/2, \eta/2]$ and all $A \in \mathcal{S}(\kappa)$. Therefore, for any $t \in [-\eta/2, \eta/2]$ and $A \in \mathcal{S}(\kappa)$:

$$\left|\Phi(X(t), A, p(t)) - \Phi(X(0), A, q) - t \left.\frac{\partial\Phi(X(u), A, p(u))}{\partial u}\right|_{u=0}\right|$$
$$= \left|\int_0^t (t-v)\frac{\partial^2\Phi(X(v), A, p(v))}{\partial v^2}dv\right| \le Kt^2 \,.$$

Noting that $\Phi(X(0), A, q) = \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}$ and using the definition of $G(A)$, we find that

$$\left|\Phi(X(t), A, p(t)) - \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} - t \ \frac{p'(0)G(A)}{q^2\|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}}\right| \le Kt^2 \,,$$

for all $t \in [-\eta/2, \eta/2]$ and $A \in \mathcal{S}(\kappa)$. $\qquad\square$

Lemma 8.A.4. *With the notations of Lemma 8.A.3, for any $0 < \varepsilon \le \kappa$, there exists $\delta > 0$ such that for all $t \in [-\delta, \delta]$ there is $A(t) \in \mathcal{N}(\mathcal{P}) \cap \mathcal{S}_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^+$ satisfying*

$$\|Y(t)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))} = \Phi(X(t), A(t), p(t)), \qquad \|Y_\mathcal{N} - A(t)\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} \le \varepsilon.$$

Proof. The proof is similar to the one of Lemma 11 of [Beigi and King, 2016]. Given $\varepsilon \le \kappa$, choose $\delta' > 0$ satisfying

$$\delta' < \min\left\{\frac{\eta}{2}, \frac{\varepsilon^2 \, p'(0) \, \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}{4Kq^2}\right\}$$

where K is defined in Lemma 8.A.3. We have

$$\mathcal{S}(\varepsilon) \subset \mathcal{S}(\kappa) \subset \mathcal{N}(\mathcal{P}) \cap \mathcal{S}_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^+$$

and so the boundary of $\mathcal{S}(\varepsilon)$ is contained in $\mathcal{N}(\mathcal{P}) \cap \mathcal{S}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$. Suppose that A is on the boundary of $\mathcal{S}(\varepsilon)$, so that

$$\|Y_{\mathcal{N}} - A\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = \varepsilon.$$

By the quantum Pinsker inequality,

$$D(\Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}}) \| \Gamma_{\sigma_{\mathrm{Tr}}}(A)) \geq \frac{1}{2} \|Y_{\mathcal{N}} - A\|^2_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = \frac{\varepsilon^2}{2}.$$

From Equation (8.45) we deduce

$$G(A) \geq G(Y_{\mathcal{N}}) + \frac{\varepsilon^2 \|Y\|^q_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}{2}. \tag{8.46}$$

Let us first consider the case where $t \geq 0$. From Lemma 8.A.3, we deduce that

$$\Phi(X(t), A, p(t)) \geq \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} + t\, \frac{p'(0)\, G(A)}{q^2 \|Y\|^{q-1}_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}} - Kt^2. \tag{8.47}$$

Our choice of δ' implies that for all $0 \leq t \leq \delta'$,

$$t\, \frac{\varepsilon^2\, p'(0)\, \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}{2\, q^2} - Kt^2 > Kt^2,$$

and hence, combining this with 8.46 and 8.47,

$$\Phi(X(t), A, p(t)) > \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} + t\, \frac{p'(0)\, G(Y_{\mathcal{N}})}{q^2 \|Y\|^{q-1}_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}} + Kt^2. \tag{8.48}$$

Furthermore, from Lemma 8.A.3, we also deduce that

$$\Phi(X(t), Y_{\mathcal{N}}, p(t)) \leq \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} + t\, \frac{p'(0)\, G(Y_{\mathcal{N}})}{q^2 \|Y\|^{q-1}_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}} + Kt^2. \tag{8.49}$$

Combining 8.48 and 8.49, we find that

$$\Phi(X(t), Y_{\mathcal{N}}, p(t)) < \Phi(X(t), A, p(t)).$$

Since this inequality holds for any A on the boundary of $\mathcal{S}(\varepsilon)$, we conclude that for all $0 \leq t \leq \delta'$, the function $A \mapsto \Phi(X(t), A, p(t))$ has a local minimum $A(t)$ in the interior of $\mathcal{S}(\varepsilon)$. We now choose $0 < \delta_+ \leq \delta'$ so that $q \leq p(t) \leq 2q$ for all $0 \leq t \leq \delta_+$ (the existence of $\delta_+ > 0$ is guaranteed by the assumptions that $p(0) = q \geq 1$ and that $t \mapsto p(t)$ is increasing and differentiable). Applying Proposition 8.A.1, we conclude that, for all $0 \leq t \leq \delta_+$, the local minimum of the convex function $A \mapsto \Phi(X(t), A, p(t))^{p(t)}$ in the interior of $\mathcal{S}(\varepsilon)$ is in fact a global minimum. Since $A \mapsto \Phi(X(t), A, p(t))$ and $A \mapsto \Phi(X(t), A, p(t))^{p(t)}$ share the same minimum $A(t) \in \mathcal{S}(\varepsilon)$, we conclude that

$$\|Y\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))} = \Phi(X(t), A(t), p(t)), \qquad \|Y_{\mathcal{N}} - A(t)\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} \leq \varepsilon.$$

We consider now the case $t \le 0$. Using Lemma 8.A.3 as well as inequality 8.46, we deduce that

$$\Phi(X(t), A, p(t)) \le \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} + t \frac{p'(0)G(A)}{q^2 \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}} + Kt^2$$

$$\le \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} + t \frac{p'(0) G(Y_{\mathcal{N}})}{q^2 \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}} + t \frac{\varepsilon^2 p'(0) \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}{2 q^2} + Kt^2,$$

where the second inequality follows from the fact that $t \le 0$. Now, for $-\delta' \le t \le 0$,

$$t \frac{\varepsilon^2 p'(0) \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}{2 q^2} + Kt^2 < -Kt^2,$$

and thus

$$\Phi(X(t), A, p(t)) < \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} + t \frac{G(Y_{\mathcal{N}})p'(0)}{q^2 \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}} - Kt^2. \tag{8.50}$$

Combining with the lower bound for F obtained from Lemma 8.A.3 we deduce that

$$\Phi(X(t), Y_{\mathcal{N}}, p(t)) > \Phi(X(t), A, p(t)),$$

for all A on the boundary of $\mathcal{S}(\varepsilon)$. Thus, we conclude that for all $\delta' \le t \le 0$, the function $A \mapsto \Phi(X(t), A, p(t))$ has a local maximum in the interior of $\mathcal{S}(\varepsilon)$. Choose $0 < \delta_- \le \delta'$ so that $1 \le p(t) \le 2$ for all $-\delta_- \le t \le 0$. Applying Proposition 8.A.1, we conclude that the local maximum of the concave function $A \mapsto \Phi(X(t), A, p(t))^{p(t)}$ in the interior of $\mathcal{S}(\varepsilon)$ is in fact a global maximum for all $-\delta_- \le t \le 0$. Finally, take $\delta := \min\{\delta_+, \delta_-\}$ to deduce that for all $t \in [-\delta, \delta]$ there exists $A(t) \in \mathcal{N}(\mathcal{P}) \cap \mathcal{S}_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^+$ satisfying:

$$\|Y\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))} = \Phi(X(t), A(t), p(t)), \qquad \|Y_{\mathcal{N}} - A(t)\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} \le \varepsilon.$$

$$\square$$

We are finally ready to state and prove Theorem 8.3.3.

Proof of Theorem 8.3.3 : Recall from (8.43) that

$$G(Y_{\mathcal{N}}) = -\|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q \ln \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^q + \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^q \ln \left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^q\right) - \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^q \ln Y_{\mathcal{N}}\right)$$

$$- \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^q \ln \sigma_{\mathrm{Tr}}\right) + \frac{q^2}{p'(0)} \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^{q-1} \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y'(0))\right). \tag{8.51}$$

Using the expression (8.44) for $Y_{\mathcal{N}}$, (8.51) reduces to

$$G(Y_{\mathcal{N}}) = \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^q \ln \left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^q\right) - \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^q \ln E_{\mathcal{N}}\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{-1}\left(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right)^q\right]\right)$$

$$- \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^q \ln \sigma_{\mathrm{Tr}}\right) + \frac{q^2}{p'(0)} \mathrm{Tr}\left(\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)\right]^{q-1} \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y'(0))\right).$$

Define now

$$\Delta(t) := \frac{1}{t}\left(\|Y(t)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))} - \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}\right) - \frac{p'(0) G(Y_{\mathcal{N}})}{q^2 \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}}. \tag{8.52}$$

We next prove that $\Delta(t) \to 0$ as $t \to 0$. Let $\varepsilon > 0$ be such that

$$0 < \varepsilon < \min\left\{\kappa, \eta, \frac{\lambda_{\min}(\Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}}))}{2}\right\} \tag{8.53}$$

where κ is the parameter introduced in Lemma 8.A.3 and $\lambda_{\min}(\Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}}))$ is the minimum eigenvalue of $\Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}})$. According to Lemma 8.A.4, there exists $\delta > 0$ such that for every $0 < t < \delta$ there is an operator $A(t) \in \mathcal{N}(\mathcal{P}) \cap \mathcal{S}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$ such that

$$\|A(t) - Y_{\mathcal{N}}\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} \leq \varepsilon \leq \kappa, \qquad \|Y(t)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))} = \Phi(X(t), A(t), p(t)).$$

Then

$$\Delta(t) = \frac{1}{t}\left(\Phi(X(t), A(t), p(t)) - \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}\right) - \frac{p'(0) G(Y_{\mathcal{N}})}{q^2 \|Y\|^{q-1}_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}$$

$$= \frac{1}{t}\left(\Phi(X(t), A(t), p(t)) - \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} - t \frac{p'(0) G(A(t))}{q^2 \|Y\|^{q-1}_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}\right) + \frac{p'(0)(G(A(t)) - G(Y_{\mathcal{N}}))}{q^2 \|Y\|^{q-1}_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}.$$

Since $A(t) \in \mathcal{S}(\varepsilon)$, Lemma 8.A.3 implies that

$$\left|\Phi(X(t), A(t), p(t)) - \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} - t \frac{p'(0) G(A(t))}{q^2 \|Y\|^{q-1}_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}\right| \leq K t^2. \tag{8.54}$$

Furthermore, from (Equation (8.45)) and using Lemma 14 of [Beigi and King, 2016]:

$$|G(A(t)) - G(Y_{\mathcal{N}})| = \|Y\|^q_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} D(\Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}}) \| \Gamma_{\sigma_{\mathrm{Tr}}}(A(t))) \leq \frac{2 \|Y\|^q_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}{\lambda_{\min}(\Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}}))} \|Y_{\mathcal{N}} - A(t)\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$$

$$\leq \frac{2 \|Y\|^q_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}{\lambda_{\min}(\Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}}))} \varepsilon. \tag{8.55}$$

Using (8.54) and (8.55), we obtain the bound

$$|\Delta(t)| \leq K t + \frac{2 p'(0) \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}{\lambda_{\min}(\Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}})) q^2} \varepsilon, \tag{8.56}$$

for all ε satisfying (8.53) and all $0 < t < \delta$. Therefore,

$$\limsup_{t \to 0} |\Delta(t)| \leq \frac{2 p'(0) \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}}{\lambda_{\min}(\Gamma_{\sigma_{\mathrm{Tr}}}(Y_{\mathcal{N}})) q^2} \varepsilon, \tag{8.57}$$

and since ε may be arbitrarily small, we deduce that

$$\limsup_{t \to 0} |\Delta(t)| = \lim_{t \to 0} |\Delta(t)| = 0.$$

$\square$

8.B. Towards the proof of Theorem 8.4.1(ii)

In this appendix, we define and study the properties of an object that turns out to be useful in the derivation of Theorem 8.4.1(ii): first define the following norm on operators $A \in \mathcal{B}(\mathcal{H})$:

$$\|A\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} := |I| \max_{i \in I} \|P_i \, A \, P_i\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} \, .$$

In what follows, we also denote by $\tilde{\mathcal{S}}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$ the set of positive definite operators A of norm $\|A\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1$. Now, given a positive semidefinite operator X and $1 \le q < p \le \infty$, let

$$\|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} := \inf_{A \in \mathcal{N}(\mathcal{P}) \cap \tilde{\mathcal{S}}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}} \left\| A^{-1/2r} X A^{-1/2r} \right\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})} \, . \tag{8.58}$$

The following lemma is straightforward:

Lemma 8.B.1. *For all X positive semidefinite, and any $1 \le q < p \le \infty$, $\|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} \ge \|X\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))}$, and equality holds whenever $|I| = 1$. Moreover, the optimum in Equation (8.58) in attained on the subset of positive definite operators $A \in \mathcal{N}(\mathcal{P})$ such that $\|P_i A P_i\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = \frac{1}{|I|}$ for all $i \in I$.*

Proof. The second part of the lemma follows from the observation that for any two positive semidefinite operators $A \in \mathcal{N}(\mathcal{P})$ and $X \in \mathcal{B}(\mathcal{H})$,

$$\|A^{-1/2r} X A^{-1/2r}\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})} = \|A^{-1/2r} \Gamma^{1/p}_{\sigma_{\mathrm{Tr}}}(X) A^{-1/2r}\|_p = \|[\Gamma^{1/p}_{\sigma_{\mathrm{Tr}}}(X)]^{1/2} A^{-1/r} [\Gamma^{1/p}_{\sigma_{\mathrm{Tr}}}(X)]^{1/2}\|_p \, .$$

Since $\frac{1}{r} = \frac{1}{q} - \frac{1}{p} \le 1$, $x \mapsto x^{1/r}$ is operators monotone and therefore the optimization in Equation (8.58) occurs at the boundary of $\tilde{\mathcal{S}}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$, that is for $\|P_i A P_i\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = \frac{1}{|I|}$ for all $i \in I$. The first part follows directly form the latter fact, since it implies that $\|A\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1$. □

Theorem 8.4.1(ii) relies crucially on the below Lemmas 8.B.3 and 8.B.4, which respectively generalize Lemmas 12 and 13 of [Beigi and King, 2016] to the non unital case and for $|I| \ge 1$. In order to prove these results, we first need to extend Lemmas 8.A.3 and 8.A.4 to the quantity defined in Equation (8.58).

Proposition 8.B.2. *Let $q \ge 1$, $[0, \infty) \ni t \mapsto p(t)$ by a twice continuously differentiable increasing function with $p(0) = q$ and $[0, \infty) \ni Y(t)$ be a twice continuous differentiable positive semidefinite matrix-valued function with $Y(0) = Y$, and for any $\kappa > 0$, define $\tilde{\mathcal{S}}(\kappa) := \mathcal{N}(\mathcal{P}) \cap \{B > 0, \|B\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1, \|B - \tilde{Y}_{\mathcal{N}}\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} \le \kappa\}$, where*

$$\tilde{Y}_{\mathcal{N}} := \sum_{i \in I} \frac{P_i \, E_{\mathcal{N}}[\mathrm{I}_{1,q}(Y)] \, P_i}{|I| \, \mathrm{Tr}[P_i \, (\Gamma^{\frac{1}{q}}_{\sigma_{\mathrm{Tr}}}(Y))^q \, P_i]} \, .$$

Then, there exists $\tilde{\kappa} > 0$ and $\tilde{K} > 0$ such that for all $t \ge 0$ and $A \in \tilde{\mathcal{S}}(\tilde{\kappa})$,

$$\left| \Phi\big(\Gamma^{\frac{1}{p(t)}}_{\sigma_{\mathrm{Tr}}}(Y(t)), A, p(t)\big) - \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} - t \, \frac{p'(0) G(A)}{q^2 \|Y\|^{q-1}_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}} \right| \le \tilde{K} t^2 \, . \tag{8.59}$$

Moreover, for any $\tilde{\varepsilon} \le \tilde{\kappa}$, there exists $\tilde{\delta} > 0$ such that for all $t \in [0, \tilde{\delta}]$ there is $A(t) \in \mathcal{N}(\mathcal{P}) \cap \tilde{\mathcal{S}}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$ satisfying

$$\|Y(t)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))} = \Phi\big(\Gamma^{\frac{1}{p(t)}}_{\sigma_{\mathrm{Tr}}}(Y(t)), A(t), p(t)\big), \qquad \|\tilde{Y}_{\mathcal{N}} - A(t)\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} \le \tilde{\varepsilon} \, . \tag{8.60}$$

Proof. The proof of (8.59) follows the exact same lines as the proof of Lemma 8.A.3. Now, let $X(t) := \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p(t)}}(Y(t))^{p(t)}$ and given $\tilde{\varepsilon} \leq \tilde{\kappa}$, choose $\tilde{\delta}' > 0$ satisfying

$$\tilde{\delta}' < \frac{\tilde{\varepsilon}^2 \, \min_{j\in I} \mathrm{Tr}(P_j \, \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q \, P_j) \, p'(0)}{4 \, \tilde{K} \, q^2 \, \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}} \, .$$

Then, we have

$$\tilde{\mathcal{S}}(\tilde{\varepsilon}) \subset \tilde{\mathcal{S}}(\tilde{\kappa}) \subset \mathcal{N}(\mathcal{P}) \cap \tilde{\mathcal{S}}_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^+ \, .$$

Suppose that A belongs to the boundary of $\tilde{\mathcal{S}}(\tilde{\varepsilon})$, so that

$$\left\| \tilde{Y}_{\mathcal{N}} - A \right\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = \tilde{\varepsilon}.$$

Hence, as in the proof of Lemma 8.A.2, we can show that

$$G(A) - G(\tilde{Y}_{\mathcal{N}}) = \mathrm{Tr}\left(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q \, (\ln \tilde{Y}_{\mathcal{N}} - \ln A) \right)$$

$$= \sum_{i\in I} \mathrm{Tr}\left(E_{\mathcal{N}*}[P_i \, \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q \, P_i](\ln \tilde{Y}_{\mathcal{N}} - \ln A) \right).$$

Now, define for any $i \in I$ the states $\sigma_i := |I| \, P_i \, \sigma_{\mathrm{Tr}}^{1/2} A \sigma_{\mathrm{Tr}}^{1/2} P_i$ and $\eta_i := |I| \, P_i \, \sigma_{\mathrm{Tr}}^{1/2} \tilde{Y}_{\mathcal{N}} \sigma_{\mathrm{Tr}}^{1/2} P_i$, one can easily verify that $E_{\mathcal{N}*}[P_i \, \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q \, P_i] = \mathrm{Tr}(P_i \, \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q \, P_i) \, \eta_i$, so that

$$G(A) - G(\tilde{Y}_{\mathcal{N}}) = \sum_{i\in I} \mathrm{Tr}(E_{\mathcal{N}*}[P_i \, \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q \, P_i](\ln \eta_i - \ln \sigma_i))$$

$$= \sum_{i\in I} \mathrm{Tr}(P_i \, \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q \, P_i) D(\eta_i \| \sigma_i)$$

$$\geq \frac{1}{2} \sum_{i\in I} \mathrm{Tr}(P_i \, \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q \, P_i) \, \|\eta_i - \sigma_i\|_1^2$$

$$= \frac{|I|^2}{2} \sum_{i\in I} \mathrm{Tr}(P_i \, \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q \, P_i) \, \|P_i \, A \, P_i - P_i \, \tilde{Y}_{\mathcal{N}} \, P_i\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^2$$

$$\geq \frac{|I|^2}{2} \min_{j\in I} \mathrm{Tr}(P_j \, \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q \, P_j) \sum_{i\in I} \|P_i \, A \, P_i - P_i \, \tilde{Y}_{\mathcal{N}} \, P_i\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^2$$

$$\geq \frac{1}{2} \min_{j\in I} \mathrm{Tr}(P_j \, \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q \, P_j) \, \left\| A - \tilde{Y}_{\mathcal{N}} \right\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}^2$$

$$\geq \frac{\tilde{\varepsilon}^2}{2} \min_{j\in I} \mathrm{Tr}(P_j \, \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{q}}(Y)^q \, P_j),$$

where we used Pinsker's inequality on the third line above. Following the steps of the proof of (8.48), we can show from (8.59) that for all $0 \leq t \leq \tilde{\delta}'$,

$$\Phi(X(t), A, p(t)) > \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} + t \, \frac{p'(0) \, G(\tilde{Y}_{\mathcal{N}})}{q^2 \, \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}} + \tilde{K} \, t^2 \, .$$

This, together with another use ot (8.59) applied to $A = \tilde{Y}_{\mathcal{N}}$ implies that

$$\Phi(X(t), \tilde{Y}_{\mathcal{N}}, p(t)) < \Phi(X(t), A, p(t)).$$

The rest of the proof follows similarly to the proof of Lemma 8.A.4. $\qquad\square$

Lemma 8.B.3. *Let* $Y \in \mathcal{B}(\mathcal{H})$ *positive definite and for* $1 \leq q < p \leq \infty$, *let* $\frac{1}{r} := \frac{1}{q} - \frac{1}{p}$. *Then the function* $\Psi(Y, \cdot, p) : A \mapsto \|A^{-1/2r} Y A^{-1/2r}\|_{\mathbb{L}_p(\sigma_{\mathrm{Tr}})}$ *is strictly convex. Moreover, there exists a unique* $\tilde{A} \in \mathcal{N}(\mathcal{P}) \cap \tilde{\mathcal{S}}^+_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})}$ *such that*

$$\Psi(Y, \tilde{A}, p) = \|Y\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_p(\sigma_{\mathrm{Tr}}))}. \tag{8.61}$$

Moreover, the optimizer $\tilde{A}$ *of Equation (8.61) satisfies the following constraint*

$$P_i \tilde{A} P_i = \frac{P_i E_{\mathcal{N}}\left[\mathrm{I}_{1,p}(\tilde{A}^{-1/2r} Y \tilde{A}^{-1/2r})\right] P_i}{|I| \,\mathrm{Tr}\left[P_i \left(\Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(\tilde{A}^{-1/2r} Y \tilde{A}^{-1/2r})\right)^p P_i\right]}. \tag{8.62}$$

Proof. Following the exact same steps as in the proof of Lemma 12 of [Beigi and King, 2016], one can show that the function

$$\Phi(X, \cdot, p) : A \mapsto \|A^{-1/2r} X A^{-1/2r}\|_p$$

is strictly convex. Let $X = \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p}}(Y)$. The first point then follows from the observation that $[A, \sigma_{\mathrm{Tr}}] = 0$ for $A \in \mathcal{N}$, so that $\Psi(Y, A, p) = \Phi(X, A, p)$. The fact that the infimum is achieved at a unique point $\tilde{A}$ also follows from the same lemma. Now, we prove Equation (8.62). Let $A \in \mathcal{N}$ such that for all $i \in I$, $\mathrm{Tr}(P_i A P_i) = \frac{1}{|I|}$. Moreover, let $D \in \mathcal{N}(\mathcal{P})$ be a self-adjoint operator such that $\mathrm{Tr}(\sigma_{\mathrm{Tr}} P_i D P_i) = 0$ for all $i \in I$. Then, it follows that for any $x \in \mathbb{R}$ sufficiently small, $A(x) := A + x D$ satisfies the same constraints as A. Let $B(x) := X^{\frac{1}{2}} A(x)^{-s/2}$ and $C(x) := A(x)^{s/2} \frac{d}{dx} A(x)^{-s/2} \in \mathcal{N}(\mathcal{P})$, where $s = 1/r$. Then the minimum is achieved at A if for any such D,

$$\begin{aligned}
0 = \frac{d}{dx}\Big|_{x=0} \Phi(X, A(x), p)^p &= \frac{d}{dx}\Big|_{x=0} \mathrm{Tr}\left[(B(x)^* B(x))^p\right] \\
&= p\,\mathrm{Tr}\left[(B(0)^* B(0))^{p-1}(B(0)^* B(0) C(0) + C(0)^* B(0)^* B(0))\right] \\
&= p\,\langle \Gamma_{\sigma_{\mathrm{Tr}}}^{-1}\left((B(0)^* B(0))^p\right), \left(C(0) + C(0)^*\right)\rangle_{\sigma_{\mathrm{Tr}}} \\
&= p\,\langle E_{\mathcal{N}}\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{-1}\left((B(0)^* B(0))^p\right)\right], \left(C(0) + C(0)^*\right)\rangle_{\sigma_{\mathrm{Tr}}} \\
&= p\,\langle A^{-1/2} E_{\mathcal{N}}\left[\Gamma_{\sigma_{\mathrm{Tr}}}^{-1}\left((B(0)^* B(0))^p\right)\right] A^{-1/2}, \Lambda_A(D)\rangle_{\sigma_{\mathrm{Tr}}}, \tag{8.63}
\end{aligned}$$

where $D \mapsto \Lambda_A(D) := A^{\frac{1}{2}}(C(0) + C(0)^*) A^{\frac{1}{2}}$ maps the space of Hermitian operators D in $\mathcal{N}(\mathcal{P})$ such that $\mathrm{Tr}[\sigma_{\mathrm{Tr}} P_i D P_i] = 0$ for all block $i \in I$ onto itself. Indeed, for any such D,

$$\begin{aligned}
\mathrm{Tr}[\sigma_{\mathrm{Tr}} P_i \Lambda_A(D) P_i] &= 2\,\mathrm{Tr}\,\sigma_{\mathrm{Tr}} P_i A^{s/2+1} \frac{d}{dx}\Big|_{x=0} (A(x)^{-s/2}) P_i \\
&= 2\,\mathrm{Tr}\,\sigma_{\mathrm{Tr}} P_i A^{s/2+1}(-s/2) A^{-s/2-1} D P_i \\
&= -s\,\mathrm{Tr}[\sigma_{\mathrm{Tr}} P_i D P_i] = 0.
\end{aligned}$$

Moreover, the map $D \mapsto \Lambda_A(D)$ is onto. To show this, we extend the definition of this map to a linear operator $\tilde{\Lambda}_A$ on the whole space of self-adjoint operators in $\mathcal{N}(\mathcal{P})$ and prove that $\tilde{\Lambda}_A$ is onto. First, notice that $D \mapsto D^{-s/2}$ is one-to-one on the set of positive definite matrices in $\mathcal{N}(\mathcal{P})$, and hence its differential at A

$$D \mapsto \frac{d}{dx}\Big|_{x=0} (A + x D)^{-s/2} \tag{8.64}$$

is onto on $\mathcal{N}(\mathcal{P}) \cap \mathcal{B}_{sa}(\mathcal{H})$. This directly implies that $\tilde{\Lambda}_A$ is onto, since it derives from the map

defined in Equation (8.64) by multiplication with positive definite operators. Hence, Λ_A is onto, which together with Equation (8.63) implies that for any $D \in \mathcal{N}(\mathcal{P}) \cap \mathcal{B}_{sa}(\mathcal{H})$ satisfying $\mathrm{Tr}[\,\sigma_{\mathrm{Tr}} P_i D P_i\,] = 0$ for all $i \in I$,

$$\langle A^{-1/2} E_{\mathcal{N}}[\Gamma_{\sigma_{\mathrm{Tr}}}^{-1}((B(0)^* B(0))^p)] A^{-1/2}, D \rangle_{\sigma_{\mathrm{Tr}}} = 0. \tag{8.65}$$

Thus, in each block $i \in I$, $P_i A^{-1/2} E_{\mathcal{N}}[\Gamma_{\sigma_{\mathrm{Tr}}}^{-1}((B(0)^* B(0))^p)] A^{-1/2} P_i$ is a multiple of the identity:

$$P_i E_{\mathcal{N}}[\Gamma_{\sigma_{\mathrm{Tr}}}^{-1}((B(0)^* B(0))^p)] P_i = c_i \; P_i A P_i, \qquad c_i \in \mathbb{R}.$$

Replacing $B(0)$ by its definition we find

$$P_i E_{\mathcal{N}}[\mathrm{I}_{1,p}(\tilde{A}^{-1/2r} Y \tilde{A}^{-1/2r})] P_i = c_i \; P_i A P_i. \tag{8.66}$$

Finally, the multiplicative factors c_i are found after tracing Equation (8.66) against σ_{Tr}, using the fact that $\mathrm{Tr}(\sigma_{\mathrm{Tr}} P_i A P_i) = \frac{1}{|I|}$ for all $i \in I$, and Equation (8.62) follows after rearranging the terms in Equation (8.66). $\qquad\square$

Lemma 8.B.4. *Given $X \in \mathcal{B}(\mathcal{H})$ positive definite and $q \geq 1$, the function*

$$[0, \infty) \ni t \mapsto \varphi(t) := \|\mathcal{P}_t(Y)\|_{\mathbb{L}_q(\mathcal{N}, \mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))} \equiv \Phi(X(t), \tilde{A}(t), p(t))$$

is continuous on $[0, \infty)$, for $p(t) := 1 + (q-1)\,\mathrm{e}^{4t/c}$, where Φ is the map defined in Equation (8.39), $X(t) \equiv \Gamma_{\sigma_{\mathrm{Tr}}}^{\frac{1}{p(t)}}(\mathcal{P}_t(Y))$ and $\tilde{A}(t)$ is the optimizer obtained in Lemma 8.B.3.

Proof. From (8.59), there exist $\tilde{\kappa} > 0$ and $\tilde{K} < \infty$, such that for all $t \in [0, \infty)$ and $A \in \tilde{\mathcal{S}}(\tilde{\kappa})$,

$$\left| \Phi(X(t), A, p(t)) - \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})} \right| \leq t\, \frac{p'(0) G(A)}{q^2 \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}} + \tilde{K} t^2.$$

Moreover, from the second part of Proposition 8.B.2 we know that, for sufficiently small t, the optimizer $\tilde{A}(t)$ is in $\tilde{\mathcal{S}}(\tilde{\kappa})$. Since $\varphi(0) = \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}$, the above inequality implies

$$|\varphi(t) - \varphi(0)| \leq t\, \frac{p'(0) G(\tilde{A}(t))}{q^2 \|Y\|_{\mathbb{L}_q(\sigma_{\mathrm{Tr}})}^{q-1}} + K t^2.$$

By definition, the map $A \mapsto G(A)$ defined in Equation (8.43) is continuous, and hence uniformly bounded on $\tilde{\mathcal{S}}(\tilde{\kappa})$. Hence, the continuity of φ at 0 follows. We now prove the continuity of φ at $t_0 > 0$. For any $0 < a < t_0 < b$, $t \in [a, b]$ and $s(t) = \frac{1}{q} - \frac{1}{p(t)}$,

$$\begin{aligned}
\varphi(t) &= \Phi(X(t), \tilde{A}(t), p(t)) \\
&= \|\tilde{A}(t)^{-s(t)/2} X(t) \tilde{A}(t)^{-s(t)/2}\|_{p(t)} \\
&\geq \|\tilde{A}(t)^{-s(t)}\|_{p(t), \sigma_{\mathrm{Tr}}} \, \|\mathcal{P}_t(Y)^{-1/2}\|_{\infty}^{-2} \\
&\geq \lambda_{\min}(\sigma_{\mathrm{Tr}}) \lambda_{\min}(\tilde{A}(t))^{-s(t)} \, \|\mathcal{P}_t(Y)^{-1/2}\|_{\infty}^{-2},
\end{aligned}$$

where $\lambda_{\min}(\tilde{A}(t))$ is the minimum eigenvalue of $\tilde{A}(t)$. On the other hand,

$$\varphi(t) = \inf_A \Phi(X(t), A, p(t)) \leq \Phi(X(t), \mathbb{1}, p(t)) = \|X(t)\|_{p(t)} \leq \|\mathcal{P}_t(Y)\|_{p(t), \sigma_{\mathrm{Tr}}} \leq \|\mathcal{P}_t(Y)\|_{\infty}.$$

Together with the previous bound, we arrive at

$$\lambda_{\min}(\tilde{A}(t))^{-s(a)} \le \lambda_{\min}(\tilde{A}(t))^{-s(t)} \le \lambda_{\min}(\sigma_{\mathrm{Tr}}) \|\mathcal{P}_t(Y)^{-1/2}\|_\infty^2 \, \|\mathcal{P}_t(Y)\|_\infty \,.$$

Above, we used that $t \mapsto s(t)$ increases, as well as the fact that $\lambda_{\min}(\tilde{A}(t)) \le 1$, since $\|\tilde{A}(t)\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1$. By continuity of $t \mapsto \mathcal{P}_t(Y)$, the right hand side of the above chain of inequalities is uniformly bounded by some positive constant $C > 0$ over the interval $[a, b]$. Therefore, $\tilde{A}(t)$ belongs to the compact set $\mathcal{R} \coloneqq \mathcal{N}(\mathcal{P}) \cap \{B > 0, \ \|B\|_{\mathbb{L}_1(\sigma_{\mathrm{Tr}})} = 1, \ \lambda_{\min}(B) \ge C^{-1/s(a)}\}$. The function $(t, A) \mapsto \Phi(X(t), A, p(t))$ restricted to the compact set $[a, b] \times \mathcal{R}$ is uniformly continuous, which means that for any $\varepsilon > 0$, there exists $\delta > 0$ such that for all $t, t' \in [a, b]$ such that $|t - t'| \le \delta$, and any $A \in \mathcal{R}$,

$$|\Phi(X(t), A, p(t)) - \Phi(t', A, p(t'))| \le \varepsilon \,.$$

Therefore,

$$\varphi(t) = \Phi(X(t), \tilde{A}(t), p(t)) \le \Phi(X(t), \tilde{A}(t'), p(t)) \le \Phi(X(t'), \tilde{A}(t'), p(t')) + \varepsilon = \varphi(t') + \varepsilon \,.$$

Conversely, $\varphi(t') \le \varphi(t) + \varepsilon$. Thus, $|\varphi(t) - \varphi(t')| \le \varepsilon$ for all $|t - t'| \le \delta$. We established the continuity of φ on the interval $[a, b]$, and hence at the point $t = t_0 \in [a, b]$. $\qquad\square$

Chapter 9.

The transference method

In Chapter 8, we introduced an extension of the theory of quantum hypercontractivity and its associated notion of a logarithmic Sobolev inequality to quantum Markov semigroups that are not necessarily primitive. However, computing the exact constants turns out to be a difficult problem, even for classical evolutions [Saloff-Coste, 1994, Diaconis et al., 1993], and requires a good knowledge of entropic and/or functional analytic methods. In this chapter, we show how one can get estimates on decoherence times for a particular class of QMS from a simple transfer of already known classical mixing times. In substance we show that, given a QMS in this class, there exists a classical Markov semigroup whose contractivity properties, and hence decoherence times, control the ones of the original QMS. On the other hand, given a projective representation of a (discrete or compact Lie) group G on a Hilbert space $\mathcal{H}$, any classical right-invariant transition kernel $(k_t)_{t \geq 0}$ on G gives rise to a convolution QMS on $\mathcal{B}(\mathcal{H})$, the convergence properties of which it controls. The "output" QMS can be primitive or nonprimitive, irrespective of the ergodic properties of the "input" classical Markov chain. The method however is at the moment restricted to the case doubly stochastic QMS, that is evolutions for which the completely mixed state is invariant.

This so-called *transference method* was introduced in [Gao et al., 2018b] who were exclusively concerned with the transfer of classical diffusions. Here, we broaden their scope to incorporate evolutions on finite groups. As we will see, given a quantum Markov semigroup $(\mathcal{P}_t)_{t \geq 0}$, one can find different classical semigroups acting on either a finite or a compact Lie group, both giving rise to $(\mathcal{P}_t)_{t \geq 0}$. The question of the optimality of the decoherence times that one gets from these different choices is then posed. We will provide examples of such situations and compare the results found from these different inputs. One important class of QMS that fits the present discussion is the one of collective decoherence introduced in Section 6.4.2. In particular, we show that given any right-invariant transition kernel on G, the decoherence times we get on the transferred QMS do not depend on the representation of the group that we choose. In the case of collective decoherence, this means that the estimate that we get is independent of the number of qubits constituting the system being studied.

Before moving to a more thorough analysis, we sketch the transference method on a simple example for sake of illustration: consider the Lindblad generator of the weak collective decoherence semigroup on $\mathcal{B}((\mathbb{C}^2)^{\otimes m})$, already introduced in Section 6.4.2:

$$\mathcal{L}_n^{\mathrm{wcd}}(X) := \Sigma_z^n \, X \, \Sigma_z^n - \frac{1}{2}((\Sigma_z^n)^2 X + X(\Sigma_z^n)^2), \qquad \text{where} \qquad \Sigma_z^n := \sum_{i=1}^n \mathbb{1}_{\mathbb{C}^2}^{\otimes i-1} \otimes \sigma_z \otimes \mathbb{1}_{\mathbb{C}^2}^{n-i}.$$

In this example, since the corresponding non-primitive semigroup $(\mathcal{P}_t^{\mathrm{wcd}})_{t \geq 0}$ is KMS-symmetric with respect to the completely mixed state, the decoherence-free subalgebra $\mathcal{N}(\mathcal{P}^{\mathrm{wcd}})$ coincides with the

fixed point algebra

$$\mathcal{N}(\mathcal{P}^{\mathrm{wcd}}) \equiv \mathcal{F}(\mathcal{P}^{\mathrm{wcd}}) = \{X : \forall t \geq 0, \, \mathcal{P}_t^{\mathrm{wcd}}(X) = X\},$$

and is given by the commutator of Σ_n^{wcd}. As seen in Chapter 6, the asymptotic behavior of the QMS is encoded into a conditional expectation $E_{\mathcal{F}} = E_{\mathcal{F}*}$:

$$E_{\mathcal{F}}(\rho) = \lim_{t \to \infty} \mathcal{P}_t^{\mathrm{wcd}}(\rho),$$

which is both a completely positive trace preserving map and the orthogonal projection onto the fixpoint algebra with respect to $\langle .,. \rangle_{\mathbb{1}/d_{\mathcal{H}}}$. Here, the main quantity of interest is the decoherence time (cf. Section 8.6):

$$\tau_{\mathrm{deco}}(\varepsilon) = \inf \{t \geq 0 \, ; \, \|\mathcal{P}_t(\rho) - E_{\mathcal{F}}(\rho)\|_1 \leq \varepsilon \quad \forall \rho\}.$$

Layout of the chapter: In Section 9.1, we explain the connection of QMS to classical diffusions and jump processes on groups via the so-called *transference technique*. In Section 9.2, we explain the technical tools that allow us to bound various norm estimates of a quantum Markov semigroup in terms of the kernel of an associated classical process. Section 9.3 consists of applications of the techniques developed in the previous sections to the derivation of bounds on decoherence times of some simple QMS.

9.1. Quantum Markov semigroups via group transference

In this chapter, we are exclusively concerned with quantum Markov semigroups that are self adjoint with respect to the Hilbert Schmidt inner product (or, equivalently, that are KMS symmetric with respect to the completely mixed state) on the algebra $\mathcal{B}(\mathcal{H})$ of linear operators on a finite dimensional Hilbert space $\mathcal{H}$. Here, we recall a general construction from [Gao et al., 2018b] to obtain quantum Markov semigroups (QMS) from classical ones, and apply it to form two classes of examples: QMS of diffusive and jump type. This construction can be interpreted as a simplified version of the characterization of quantum convolution semigroups of [Kümmerer and Maassen, 1987] provided in Section 5.5.3.

The starting point is a compact group G, either Lie or finite, with Haar measure μ_G. Let $(P_t)_{t\geq 0}$ be a Markov semigroup on the algebra $\mathbb{L}_\infty(G)$ of bounded, measurable functions on G. We will always assume that $(P_t)_{t\geq 0}$ admits a kernel representation (see Section 2.1):

$$P_t(f)(g) = \int_G k_t(g,h) \, f(h) \, d\mu_G(h). \tag{9.1}$$

We also assume that $(P_t)_{t\geq 0}$ is right-invariant, which means that the probability to visit h from g only depends on gh^{-1}. This implies that μ_G is an invariant probability distribution and that $k_t(g,h) = k_t(gh^{-1},e)$ where e is the neutral element of the group. We keep the same notation $k_t(g)$ for $k_t(g,e)$. Let $g \mapsto U_g$ be a projective representation of G on some finite dimensional Hilbert space $\mathcal{H}$. We define the following convolution QMS on $\mathcal{B}(\mathcal{H})$ which we call a *transferred QMS*:

$$\mathcal{P}_t(X) = \int_G k_t(g^{-1}) \, U_g^* \, X \, U_g \, d\mu_G(g). \tag{9.2}$$

At the root of the transference technique that we study in this chapter is the factorization property between $(P_t)_{t\geq 0}$ and $(\mathcal{P}_t)_{t\geq 0}$, involving the standard co-representation

$$\pi : \mathcal{B}(\mathcal{H}) \to \mathbb{L}_\infty(G, \mathcal{B}(\mathcal{H})), \quad \pi(X)(g) = U_g^* \, X \, U_g.$$

Indeed, it can be checked that the following relation holds for all $t \geq 0$ (Lemma 4.6 in [Gao et al., 2018b]):

$$\pi \circ \mathcal{P}_t = (P_t \otimes \mathrm{id}_{\mathcal{B}(\mathcal{H})}) \circ \pi. \tag{9.3}$$

Indeed, for any $X \in \mathcal{B}(\mathcal{H})$,

$$\begin{aligned}
\pi \circ \mathcal{P}_t(X)(g) &= U_g^* \int_G k_t(h^{-1}) U_{h^{-1}} X U_h \, d\mu_G(h) \, U_g \\
&= \int_G k_t(gg^{-1}h^{-1}) U_{(hg)^{-1}} X U_{hg} \, d\mu_G(h) \\
&= \int_G k_t(gh^{-1}) U_{h^{-1}} X U_h \, d\mu_G(h) \\
&= (P_t \otimes \mathrm{id}_{\mathcal{B}(\mathcal{H})})(\pi(X))(g).
\end{aligned}$$

From the invariance of μ_G, one can easily verify that any QMS $(\mathcal{P}_t)_{t \geq 0}$ transferred from $(P_t)_{t \geq 0}$ is doubly stochastic: $\mathcal{P}_{t*}(d_{\mathcal{H}}^{-1}\mathbb{1}) = d_{\mathcal{H}}^{-1}\mathbb{1}$ for any $t \geq 0$. In practice, we only consider situations where the classical Markov semigroup $(P_t)_{t \geq 0}$ is *primitive*, that is, μ_G is the unique invariant distribution and furthermore

$$P_t f \underset{t \to +\infty}{\longrightarrow} \mathbb{E}_{\mu_G}[f] = \int_G f(g) \, d\mu_G(g).$$

This does not imply that the transferred QMS $(\mathcal{P}_t)_{t \geq 0}$ is also primitive. On the other hand, the reversibility of $(P_t)_{t \geq 0}$ is transferred to the QMS $(\mathcal{P}_t)_{t \geq 0}$.

Lemma 9.1.1. *Assume that the Markov semigroup $(P_t)_{t \geq 0}$ is reversible, or equivalently that $k_t(g) = k_t(g^{-1})$ for any $g \in G$. Then any QMS $(\mathcal{P}_t)_{t \geq 0}$ transferred from $(P_t)_{t \geq 0}$ is KMS-symmetric with respect to $d_{\mathcal{H}}^{-1} \mathbb{1}_{\mathcal{H}}$.*

Proof. The result follows from the simple calculation:

$$\begin{aligned}
\langle X, \mathcal{P}_t(Y) \rangle_{\mathrm{HS}} &= \int_G k_t(g^{-1}) \, \mathrm{Tr}\left(X^* U_g^* Y U_g \right) d\mu_G(g) \\
&= \int_G k_t(g^{-1}) \, \mathrm{Tr}\left((U_g X U_g^*)^* Y \right) d\mu_G(g) \\
&= \int_G k_t(g^{-1}) \, \mathrm{Tr}\left((U_g^* X U_g)^* Y \right) d\mu_G(g) \\
&= \langle \mathcal{P}_t(X), Y \rangle_{\mathrm{HS}},
\end{aligned}$$

where the third line follows from the identity $k_t(g) = k_t(g^{-1})$ for all $g \in G$. $\qquad\square$

Since $d_{\mathcal{H}}^{-1} \mathbb{1}_{\mathcal{H}}$ is an invariant state of $(\mathcal{P}_t)_{t \geq 0}$, Theorem 6.1.2 applies and the set of fixed points $\mathcal{F}$ is an algebra. Assuming moreover that $(P_t)_{t \geq 0}$ is self-adjoint, then so is $(\mathcal{P}_t)_{t \geq 0}$, and the decoherence-free algebra $\mathcal{N}$ is equal to $\mathcal{F}$.

Moreover, the fixed-point algebra $\mathcal{F}$ is then characterized as the commutant of the projective representation (see Theorem 6.13 of [Wolf, 2012]):

$$\mathcal{F} = \{ U_g \, ; \, g \in G \}'. \tag{9.4}$$

By definition, it is also the algebra of fixed points of the $*$-automorphisms $X \mapsto U_g^* X U_g$, $g \in G$. This implies that the following commuting diagram also holds:

$$
\begin{array}{ccc}
\mathcal{B}(\mathcal{H}) & \xrightarrow{\ E_{\mathcal{F}}\ } & \mathcal{F} \\
\downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\
\mathbb{L}_\infty(G, \mathcal{B}(\mathcal{H})) & \xrightarrow{\ \mathbb{E}_{\mu_G}\ } & \mathcal{B}(\mathcal{H}).
\end{array}
$$

We now turn our attention to two special cases of such construction. In both cases, we explicitly construct the Linblad generator of the QMS.

9.1.1. Diffusion

Let $(P_t)_{t\geq 0}$ be the semigroup defined in Section 3.7.3 corresponding to the Hörmander system $V = \{V_1, ..., V_m\}$ on a given Lie group G, of corresponding generator

$$L_V = \sum_j V_j^2 \,. \tag{9.5}$$

The generator L_V generates a semigroup $P_t = e^{tL_V}$ on $\mathbb{L}_\infty(G)$. Since the semigroup commutes with the right action of the group, it is implemented by a convolution kernel as in Equation (9.1):

$$P_t(f)(g) = \int k_t(gh^{-1}) f(h) \, d\mu_G(h) \,.$$

Next, considering a projective representation $g \mapsto U_g$ of G on some finite dimensional Hilbert space $\mathcal{H}$, we want to find the Lindblad generator of the QMS defined by Equation (9.2). We first observe that, given the geodesic g_j associated to the vector field V_j, $U_{g_j(t)}$ is a one parameter family of unitaries and hence

$$\frac{d}{dt} U_{g_j(t)} \Big|_{t=0} = iA_j \tag{9.6}$$

is given by a self-adjoint matrix in finite dimension. This implies that, for any $X \in \mathcal{B}(\mathcal{H})$,

$$
\begin{aligned}
(V_j \otimes \mathrm{id}_{\mathcal{B}(\mathcal{H})}) \circ \pi(X)(g) &= \frac{d}{dt} \pi(X)(g_j(t)g)\Big|_{t=0} \\
&= \frac{d}{dt} U_g^* U_{g_j(t)}^* \, X \, U_{g_j(t)} U_g \Big|_{t=0} \\
&= -i\pi([A_j, X])(g) \,.
\end{aligned}
$$

Therefore we get

$$(L_V \otimes \mathrm{id}_{\mathcal{B}(\mathcal{H})}) \circ \pi(X) = \sum_j i^2 \pi([A_j, [A_j, X]]) = \pi\Big(\sum_j 2A_j X A_j - A_j^2 X - X A_j^2 \Big) \,.$$

By Equation (9.3), it means that the Linblad generator of the transferred QMS is given by

$$\mathcal{L}_V(X) := \sum_j 2A_j X A_j - A_j^2 X - X A_j^2 \,.$$

Remark that conversely, if a Lindblad generator of a QMS on $\mathcal{B}(\mathcal{H})$ has the form given by the previous equation for some self-adjoint elements $A_j \in \mathcal{B}_{sa}(\mathcal{H})$, then we can consider the A_j as tangent elements of the Lie group $\mathcal{U}(\mathcal{H})$ at the identity $\mathbb{1}_\mathcal{H}$. Therefore they generate a Hörmander system. Furthermore they span the whole tangent space of the Lie-subgroup of $\mathcal{U}(\mathcal{H})$ induced by this Hörmander system, so that the corresponding generator L_V defined through Equation (9.5) is the one of a primitive Markov semigroup $(P_t)_{t\geq 0}$. We summarize this discussion in the next theorem.

Theorem 9.1.2. *Let $g \mapsto U_g$ be a projective representation of G on some finite dimensional Hilbert space $\mathcal{H}$. Then the Linblad generator of the transferred QMS $(\mathcal{P}_t = e^{t\mathcal{L}_V})_{t\geq 0}$ as defined by Equation (9.2) is given by*

$$\mathcal{L}_V(X) = \sum_j 2A_j \, X \, A_j - A_j^2 \, X - X \, A_j^2 \,, \tag{9.7}$$

where the A_j's are defined by Equation (9.6). Conversely, let $\mathcal{L}$ be the Lindblad generator of a QMS on $\mathcal{B}(\mathcal{H})$ which takes the form (9.7) for some self-adjoint elements $A_j \in \mathcal{B}_{sa}(\mathcal{H})$. Then there exists a compact Lie group G, a continuous projective representation $U : G \to \mathcal{U}(H)$ and a Hörmander system $V = \{V_1, ..., V_m\}$ in the Lie algebra of G such that $\pi : X \mapsto (g \mapsto U_g^* X U_g)$ satisfies

$$\pi(\mathcal{L}_V(X)) = (L_V \otimes \mathrm{id}_{\mathcal{B}(\mathcal{H})})(\pi(X)).$$

9.1.2. Jumps

Let now G be a finite group and let $(k_t(g,h))_{g,h \in G}$ be a right-invariant density kernel on G (cf. Section 2.4). We write $(g_t)_{t \geq 0}$ the stochastic process on G induced by this kernel. The corresponding Markov semigroup admits a transition matrix L such that $P_t = e^{tL}$ for all $t \geq 0$. Writing $\sigma_h = L(h^{-1}, e) > 0$ for all $h \neq e$, we then have by right invariance that for all $f \in \mathbb{L}_\infty(G)$,

$$L(f)(g) = \sum_{h \in G} \sigma_h \left(f(hg) - f(g) \right),$$

where we used that $\sum_h \sigma_h = 0$. Thanks to the right-invariance, we can then define a family of independent Poisson processes $\left((\tilde{N}_t^h)_{t \geq 0} \right)_{h \in G}$ with intensity σ_h such that for any function $f \in \mathbb{L}_\infty(G)$:

$$f(g_t) - f(g_{t-}) = \sum_{h \in G} \left(f(hg_t) - f(g_{t-}) \right) \left(\tilde{N}_t^h - \tilde{N}_{t-}^h \right).$$

Define the compensated Poison process with intensity σ_h and jumps $1/\sqrt{\sigma_h}$:

$$N_t^h = \frac{1}{\sqrt{\sigma_h}} \left(\tilde{N}_t^h - \sigma_h t \right).$$

Writing $df(g_t) := f(g_t) - f(g_{t-})$ and $dN_t^h := N_t^h - N_{t-}^h$, we can rewrite the previous equation as the stochastic differential equation:

$$df(g_t) = \sum_{h \in G} \sigma_h \left(f(hg_t) - f(g_{t-}) \right) dt + \sum_{h \in G} \sqrt{\sigma_h} \left(f(hg_t) - f(g_{t-}) \right) dN_t^h. \tag{9.8}$$

We are now ready to build a QMS from this Markov chain. Let $g \mapsto U_g$ be a projective representation of G on some finite dimensional Hilbert space $\mathcal{H}$. We want to find a stochastic differential equation for $(U_{g_t})_{t \geq 0}$. To this end, take $Y \in \mathcal{B}(\mathcal{H})$ and define $f_Y : h \in G \mapsto \mathrm{Tr}[Y U_h]$. Applying Equation (9.8) to f_Y we find

$$df_Y(g_t) = \sum_{h \in G} \sigma_h \, \mathrm{Tr}\left[Y \left(U_{hg_{t-}} - U_{g_{t-}} \right) \right] dt + \sum_{h \in G} \sqrt{\sigma_h} \, \mathrm{Tr}\left[Y \left(U_{hg_{t-}} - U_{g_{t-}} \right) \right] dN_t^h.$$

From this we deduce

$$dU_{g_t} = \sum_{h \in G} \sigma_h \left(U_g - \mathbb{1}_{\mathcal{H}} \right) U_{g_{t-}} dt + \sum_{h \in G} \sqrt{\sigma_h} \left(U_h - \mathbb{1}_{\mathcal{H}} \right) U_{g_{t-}} dN_t^h. \tag{9.9}$$

This equation is well-known in the theory of quantum stochastic calculus, see [Hudson and Parthasarathy, 1984, Meyer, 1993].

Theorem 9.1.3. *Let $(P_t = e^{tL})_{t \geq 0}$ be a Markov semigroup on a finite group G with right-invariant Markov kernel. Write $\sigma_g = L(g^{-1}, e)$. Then the generator of the QMS $(\mathcal{P}_t)_{t \geq 0}$ defined by Equation (9.2)*

is given for all $X \in \mathcal{B}(\mathcal{H})$ by

$$\mathcal{L}(X) = \sum_{g \in G} \sigma_g \left(U_g^* X U_g - X \right). \tag{9.10}$$

Furthermore, $\frac{\mathbb{1}_{\mathcal{H}}}{d_{\mathcal{H}}}$ is an invariant density matrix and if $(P_t)_{t \geq 0}$ is self-adjoint, then so is $(\mathcal{P}_t)_{t \geq 0}$. Conversely, let $\mathcal{L}$ be a Lindblad generator on $\mathcal{B}(\mathcal{H})$ of the form

$$\mathcal{L}(X) = \sum_{k=1}^{m} \sigma_k (U_k^* X U_k - X),$$

for some unitary operators $U_k \in \mathcal{U}(\mathcal{H})$ and some positive constants σ_k. Assume that the group G generated by $U_1, ..., U_m$ is finite and define

$$L(f)(g) = \sum_{k=1}^{m} \sigma_k \left(f(U_k\, g) - f(g) \right).$$

Then L is the generator of a primitive Markov semigroup $(P_t = e^{tL})_{t \geq 0}$ on the oriented graph (G, E), where

$$E = \{ (g, U_k g) \mid k = 1, ..., m \,; g \in G \}.$$

Furthermore, the map $\pi : \mathcal{B}(\mathcal{H}) \to \mathbb{L}_\infty(G, \mathcal{B}(\mathcal{H}))$ satisfying the defining property

$$\pi(X)(U_k) = U_k^* X U_k$$

extend to a $*$-representation of $\mathcal{B}(\mathcal{H})$ on $\mathbb{L}_\infty(G, \mathcal{B}(\mathcal{H}))$ such that $(L \otimes \mathrm{id}_{\mathcal{B}(\mathcal{H})}) \circ \pi(X) = \pi(\mathcal{L}(X))$ for all $X \in \mathcal{B}(\mathcal{H})$.

Proof. We begin by proving Equation (9.10). By definition, we have for all $X \in \mathcal{B}(\mathcal{H})$

$$\mathcal{L}(X) = \frac{d}{dt} \mathcal{P}_t(X) \Big|_{t=0} = \frac{d}{dt} \mathbb{E}\left[U_{g_t}^* X U_{g_t} \right] \Big|_{t=0}.$$

Equation (9.10) follows from an application of the Îto formula for compensated Poisson processes. The fact that $\frac{\mathbb{1}_{\mathcal{H}}}{d_{\mathcal{H}}}$ is an invariant density matrix is straightforward as clearly

$$\mathcal{L}_* \left(\frac{\mathbb{1}_{\mathcal{H}}}{d_{\mathcal{H}}} \right) = 0.$$

Now, if $(P_t)_{t \geq 0}$ is self-adjoint, then $\sigma_g = \sigma_{g^{-1}}$ for all $g \in G$. We then have for all $x \in \mathcal{B}(\mathcal{H})$:

$$\begin{aligned}
\mathcal{L}_*(X) &= \sum_{g \in G} \sigma_g \left(U_g X U_g^* - X \right) \\
&= \sum_{g \in G} \sigma_{g^{-1}} \left(U_{g^{-1}} X (U_{g^{-1}})^* - X \right) \\
&= \sum_{g \in G} \sigma_g \left(U_g^* X U_g - X \right) \\
&= \mathcal{L}(X).
\end{aligned}$$

The second part of the proof is straightforward. $\qquad\square$

9.1.3. The general situation

The two cases explored above are particular instances of the convolution QMS defined by [Kossakowski, 1972] and fully characterized (in the finite dimensional case) by [Kümmerer and Maassen, 1987] (cf.

Section 5.5.3). In particular, the generators of the form given by Equation (5.51) are the sum of three parts:

(i) The first part corresponds to a unitary evolution, with generator given by $\mathcal{B}(\mathcal{H}) \ni x \mapsto i\,[H, X]$ where H is self-adjoint;

(ii) A diffusion part, given by

$$\mathcal{B}(\mathcal{H}) \ni X \mapsto \sum_{j=1}^{k} 2A_j X A_j - A_j^2 X - X A_j^2,$$

where the A_j are self-adjoint operators. Any such family $\{A_j\}$ is a Hörmander system for the sub-Lie algebra that they generate, as elements of the unitary group $\mathcal{U}(\mathcal{H})$ of $\mathcal{H}$. Consequently the result of Section 9.1.1 applies.

(iii) A jump part, given by

$$\mathcal{B}(\mathcal{H}) \ni X \mapsto \sum_{i=1}^{l} c_i \left(U_i^* X U_i - X \right),$$

where the U_i's are unitary operators on $\mathcal{H}$. Compared to previously, this class is larger than the one presented in Section 9.1.2. Indeed, the family $\{U_i\}$ spans a subgroup of the unitary group $\mathcal{U}(\mathcal{H})$, however in general it will not be a finite group.

Remark 9.1.4. Starting with a Linblad generator, there may be an ambuiguity around the choice of the underlying group and classical Markov semigroup leading to it. Indeed, in the jump scenario when the QMS is self-adjoint, it is always possible to write the Linblad generator as in the diffusive case. Then either the group is large, i.e. the commutator is $\mathbb{C}\,\mathbb{1}_{\mathcal{H}}$, and we can treat it as an Hörmander system, or the group is small (for us finite) and we can treat it as a Markov semigroup with jumps on the Cayley graph of the group. In both cases, estimates on the decoherence time of the corresponding QMS can be found.

9.2. Norm transference and decoherence times

Now, we show how the machinery described above provides a control on the norms and entropies of the output of the quantum Markov convolution semigroups of Section 9.1 in terms of the kernel of their associated classical semigroup. We start by recalling the notations of Section 9.1: $(P_t)_{t\geq 0}$ is a Markov semigroup on the compact group G (either Lie or finite), with right-invariant kernel $(k_t)_{t\geq 0}$. Let the QMS $(\mathcal{P}_t = \mathrm{e}^{t\mathcal{L}})_{t\geq 0}$ be the transferred QMS on $\mathcal{B}(\mathcal{H})$ defined by Equation (9.2) through the projective representation $g \mapsto U_g$ of G on a finite dimensional Hilbert space $\mathcal{H}$.

Next theorem regroups all the transference techniques which will be frequently used in the remaining of this chapter (see [Gao et al., 2018b]). We recall that the spectral gap $\lambda(\mathcal{L})$ of the symmetric Lindblad generator $\mathcal{L}$ of $(\mathcal{P}_t)_{t\geq 0}$ (resp. the spectral gap $\lambda(L)$ of the generator L of $(P_t)_{t\geq 0}$) is the smallest non-zero eigenvalue of $-\mathcal{L}$ (resp. of $-L$).

Theorem 9.2.1 (Transference)**.** *Let $\mathcal{P}_t = \mathrm{e}^{t\mathcal{L}}$ and $P_t = \mathrm{e}^{tL}$ as in Section 9.1. Then*

(i) $\lambda(\mathcal{L}) \geq \lambda(L)$;

(ii) $\|\mathcal{P}_t : \mathbb{L}_s^p(\mathcal{F} \subset \mathcal{B}(\mathcal{H})) \to \mathbb{L}_s^q(\mathcal{F} \subset \mathcal{B}(\mathcal{H}))\|_{\mathrm{cb}} \leq \|P_t : \mathbb{L}_p(\mu_G) \to \mathbb{L}_q(\mu_G)\|_{\mathrm{cb}}$;

(iii) $\|\mathcal{P}_t - E_{\mathcal{F}} : \mathbb{L}_s^p(\mathcal{F} \subset \mathcal{B}(\mathcal{H})) \to \mathbb{L}_s^q(\mathcal{F} \subset \mathcal{B}(\mathcal{H}))\|_{\mathrm{cb}} \leq \|P_t - \mathbb{E}_{\mu_G}[.]\,\mathbb{1}_G : \mathbb{L}_p(\mu_G) \to \mathbb{L}_q(\mu_G)\|_{\mathrm{cb}}$.

The above theorem is very powerful, since it directly implies the following bound on the decoherence time of the QMS $(\mathcal{P}_t)_{t\geq 0}$ in terms of the mixing time of the corresponding classical Markov chain $(P_t)_{t\geq 0}$.

Corollary 9.2.2. *Let $(\mathcal{P}_t)_{t\geq 0}$ a self-adjoint QMS of the form of Equation (9.2), of associated self-adjoint classical Markov chain $(P_t)_{t\geq 0}$. Then, for any $\varepsilon > 0$:[1]*

$$\tau_{\mathrm{deco}}(\varepsilon) < \tau_{\mathrm{mix}}(\varepsilon/2) := \inf \left\{ t \geq 0 : \left\| P_t \left(\frac{d\nu}{d\mu_G} \right) - 1 \right\|_{\mathbb{L}_1(\mu_G)} \leq \varepsilon \quad \forall \nu << \mu_G \right\}. \tag{9.11}$$

Proof. This is a direct consequence of point (iii) of Theorem 9.2.1 together with Lemma 1.1.14. $\quad\square$

As seen in Section 3.4, the semigroup $(P_t)_{t\geq\infty}$ is called ultracontractive when it is bounded from $\mathbb{L}_1(\mu_G)$ (or $\mathbb{L}_2(\mu_G)$) to $\mathbb{L}_\infty(G)$. One can recast this property in terms of the density kernel as follows:

$$\| P_t : \mathbb{L}_1(\mu_G) \to \mathbb{L}_\infty(G) \|_{\mathrm{cb}} = \| P_t : \mathbb{L}_1(\mu_G) \to \mathbb{L}_\infty(G) \| = \sup_{g \in G} |k_t(g)|, \tag{9.12}$$

where the first identity follows from Lemma 1.1.14, whereas the second identity follows by duality. Similarly (see e.g. [Saloff-Coste, 1997]):

$$\| P_t - \mathbb{E}_{\mu_G}[.]\mathbb{1}_G : \mathbb{L}_1(\mu_G) \to \mathbb{L}_2(\mu_G) \| = \left(\int_G |k_t(g) - 1|^2 d\mu_G(g) \right)^{1/2}. \tag{9.13}$$

This together with Theorem 9.2.1 implies that any estimate on the kernel directly translates into an estimate on the transferred QMS. Such estimates can be found e.g. in [Saloff-Coste, 1994, Diaconis and Saloff-Coste, 1996a]. In other words, ultracontractivity of a classical semigroup translates to ultracontractivity of its transferred QMS.

9.3. Examples

Here, we illustrate the method developed in the previous sections by listing easy examples of known quantum convolution semigroups, as well as the constants that one gets from the transference technique described in Section 9.2.

9.3.1. The depolarizing QMS

Perhaps the simplest QMS that one can think about is the depolarizing semigroup on $\mathcal{B}(\mathbb{C}^n)$:

$$\mathcal{L}^{\mathrm{depol}}(\rho) = \frac{\mathbb{1}_{\mathbb{C}^n}}{n} - \rho, \qquad \mathcal{P}_t^{\mathrm{depol}}(\rho) = \mathrm{e}^{-t}\rho + (1 - \mathrm{e}^{-t})\frac{\mathbb{1}_{\mathbb{C}^n}}{n}.$$

Using the representation of $\mathbb{Z}_n \times \mathbb{Z}_n$ via the discrete Weyl matrices $\{U_{i,j}\}_{i,j\in[n]}$ (see e.g. [Wolf, 2012]) together with Equation (9.10), we directly find that the QMS transferred in this representation is

$$\mathcal{L}(\rho) = \frac{1}{n^2} \sum_{i,j=1}^{n} \left(U_{i,j}\, \rho\, U_{i,j}^* - \rho \right) \tag{9.14}$$

$$= \mathrm{Tr}(\rho) \frac{\mathbb{1}_{\mathbb{C}^n}}{n} - \rho \tag{9.15}$$

$$= \mathcal{L}^{\mathrm{depol}}(\rho), \tag{9.16}$$

---

[1]The factor $\varepsilon/2$ comes from the definition of the mixing time for classical Markov semigroups in terms of the total variation distance.

where we took $\sigma_{i,j} := \frac{1}{n^2}$ for all i, j. This choice implies that the uniform random walk on the complete graph with n^2 vertices transfers to $(\mathcal{P}_t^{\mathrm{depol}})_{t \geq 0}$. Recalling the bound (3.13) and the constant for the complete graph given in (3.30), we find the following upper bound on the mixing time of $(\mathcal{P}_t^{\mathrm{depol}})_{t \geq 0}$:

$$\tau_{\mathrm{deco}}(\varepsilon) \leq \tau_{\mathrm{mix}}(\varepsilon/2) \leq n^2 \frac{1 - \ln \varepsilon}{n^2 - 1} + \frac{n^2 \ln(n^2 - 1)}{2(n^2 - 2)} \ln \ln n^2 \underset{n \to \infty}{\sim} \frac{\ln(n) \ln \ln(n)}{2}.$$

This can be compared with the tighter bound that one can get from the modified logarithmic Sobolev constant $\alpha_1(\mathcal{L}^{\mathrm{depol}}) \geq \frac{1}{4}$:

$$\|\rho_t - n^{-1} \mathbb{1}_{\mathbb{C}^n}\|_1 \leq \sqrt{2 \ln n} \, e^{-\frac{t}{2}} \quad \Rightarrow \quad \tau_{\mathrm{deco}}(\varepsilon) \leq 2 \ln \frac{\sqrt{2 \ln n}}{\varepsilon} \sim \ln \ln n.$$

9.3.2. The dephasing QMS

We recall that the *dephasing quantum Markov semigroup* (also called *decoherent QMS* in [Bardet, 2017]) on $\mathcal{B}(\mathbb{C}^n)$, $n \geq 3$, is given by

$$\mathcal{L}^{\mathrm{deph}}(X) = E_{\mathrm{diag}}[X] - X, \qquad \mathcal{P}_t^{\mathrm{deph}}(X) = e^{-t} X + (1 - e^{-t}) E_{\mathrm{diag}}[X],$$

where E_{diag} denotes the projection on the space of matrices that are diagonal in some prefixed eigenbasis. Here, we show how simple representations of the discrete and continuous torus both lead to the *dephasing quantum Markov semigroup.*

Dephasing from the discrete torus   Choose the uniform random walk of kernel $K(j, k) = 1/n$ for any $j, k \in \mathbb{Z}_n$. A simple unitary representation of $\mathbb{Z}_n$ is given by taking $\mathcal{H} = \mathbb{C}^n$ and

$$U_j := U^j, \quad j \in \mathbb{Z}_n,$$

where U denotes the Weyl unitary operator given by $U = \mathrm{diag}\left(1, e^{\frac{2i\pi}{n}}, ..., e^{\frac{2i(n-1)\pi}{n}}\right)$ on $\mathcal{B}(\mathbb{C}^n)$, where the diagonal is chosen to be the one corresponding to E_{diag}. One can easily verify from Equation (3.29) and Equation (9.2) that the QMS $(\mathcal{P}_t^{\mathrm{deph}})_{t \geq 0}$ coincides with the generator of the transferred QMS corresponding to the uniform kernel on $\mathbb{Z}_n$, since by a direct calculation $E_{\mathrm{diag}}[X] = \frac{1}{n} \sum_{j \in \mathbb{Z}_n} U^{-j} X U^j$.

Now, it results from Corollary 9.2.2, (3.13) and (3.30) that

$$\tau_{\mathrm{deco}}(\varepsilon) \leq n \frac{1 - \ln \varepsilon}{n - 1} + \frac{n \ln(n - 1)}{2(n - 2)} \ln \ln n \underset{n \to \infty}{\sim} \frac{\ln(n) \ln \ln(n)}{2}.$$

Dephasing from the n-dimensional torus   Take the representation of the n-dimensional torus that consists of diagonal unitary matrices:

$$\mathbb{T}^n \ni (t_1, ..., t_n) \mapsto \begin{bmatrix} e^{2it_1\pi} & & 0 \\ & \ddots & \\ 0 & & e^{2it_n\pi} \end{bmatrix}.$$

The QMS associated to the heat semigroup and the above representation corresponds to $(\mathcal{P}_t^{\mathrm{deco}})_{t \geq 0}$. This simply follows from Equation (9.7) by taking the generators $A_j := |j\rangle\langle j|$ of $\mathbb{T}^n$, so that the

generator of the transferred QMS is equal to

$$\mathcal{L}(X) = \frac{1}{2} \sum_{j=1}^{n} 2 |j\rangle\langle j| X |j\rangle\langle j| - |j\rangle\langle j| X - X |j\rangle\langle j|$$

$$= E_{\mathrm{diag}}[X] - X.$$

Then the estimation (3.32) leads to the following bound on the decoherence time of these QMS:

$$\tau_{\mathrm{deco}}(\varepsilon) \leq \frac{1}{2} \ln\left(\frac{1}{2} n \ln n\right) + 6 - \ln \varepsilon \underset{n \to \infty}{\sim} \frac{\ln(n)}{2}.$$

Hence, the estimate found on the decoherence time from the continuous torus turns out to be sharper than the one found from the discrete torus. Moreover, these two bounds can be compared with the one found via decoherence-free modified logarithmic Sobolev inequality in [Bardet, 2017]. There, the DF-MLSI constant $\alpha_1(\mathcal{L}^{\mathrm{deph}})$ of the dephasing semigroup was found to be equal to $1/4$. Therefore, a simple use of Pinsker inequality provides:

$$\tau_{\mathrm{deco}}(\varepsilon) \leq 2 \ln \frac{\sqrt{2 \ln n}}{\varepsilon} \underset{n \to \infty}{\sim} \ln \ln n.$$

9.3.3. Collective decoherence

The bounds provided by the transference method for the examples studied in the last two sections, namely the depolarizing and the dephazing semigroups, are worse than the already known ones derived from the modified logarithmic Sobolev inequality. In this section, on the other hand, we show that our method provides an easy way of deriving estimates for collective decoherence on lattice spin systems (cf. Section 6.4.2). The power of the method lies in the fact that the constants derived are independent of the representation chosen. In particular, we get estimates that are independent of the lattice size by choosing tensor product representations.

More precisely, let G be a group and $U : G \to \mathcal{B}(\mathcal{H})$ a projective representation of G on some finite dimensional Hilbert space $\mathcal{H}$. For all $n \geq 1$, this representation induces a new representation on $\mathcal{H}^{\otimes n}$ given by:

$$g \mapsto U_g^{\otimes n}.$$

Let $(P_t)_{t \geq 0}$, $(\mathcal{P}_t)_{t \geq 0}$ be defined as in Equations (9.1) and (9.2) using the representation $g \mapsto U_g$. We write $(\mathcal{P}_t^{(n)})_{t \geq 0}$ the corresponding QMS on $\mathcal{H}^{\otimes n}$ for the representation $U^{\otimes n}$ and $\mathcal{L}_n$ its generator.

Diffusive case: In the diffusive case presented in Section 9.1.1, the generator $\mathcal{L}$ of $(\mathcal{P}_t)_{t \geq 0}$ has the following form:

$$\mathcal{L}(x) = \sum_k 2 A_k X A_k - A_k^2 X - X A_k^2, \tag{9.17}$$

where the A_k's are selfajdoint operators on $\mathcal{H}$. Then the generator $\mathcal{L}_n$ takes the form

$$\mathcal{L}_n(x) = \sum_k 2 A_k(n) X A_k(n) - A_k(n)^2 X - X A_k(n)^2, \quad A_k(n) = \sum_{j=1}^{n} \mathbb{1}_{\mathcal{H}}^{\otimes j-1} \otimes A_k \otimes \mathbb{1}_{\mathcal{H}}^{n-j},$$

where in the jth term of the above sum, A_k acts on the jth copy of $\mathcal{H}$. If we assume that $\mathcal{L}$ is ergodic, then the A_k belong to some unknown compact Lie group, hence the family satisfies the transferrence principle and the previous results of this chapter can be applied. As a consequence, we obtain dimension free bounds.

Jump case: In the jump case presented in Section 9.1.2, the generator $\mathcal{L}$ of $(\mathcal{P}_t)_{t\geq0}$ has the following form:

$$\mathcal{L}(X) = \sum_{k=1}^{m} \sigma_k \left(X - U_k^* \, X \, U_k \right), \tag{9.18}$$

where the U_k's are unitary operators on $\mathcal{H}$. Then the generator $\mathcal{L}_n$ takes the form

$$\mathcal{L}_n(X) = \sum_{k=1}^{m} \sigma_k \left(X - V_k^* \, X \, V_k \right),$$

where $V_k = U_k^{\otimes n}$ for all k. If the unitary operators U_k generate a finite group G then thanks to Theorem 9.1.3 we can find a Markov semigroup on G and all the estimates we find on this semigroup can be transferred to $(\mathcal{P}_t^{(n)})_{t\geq0}$ for all n.

Remark 9.3.1. Unfortunately, it is not the decoherence time or any other interesting quantity for $\mathcal{L}$ itself which transfers to all the $\mathcal{L}_n$, but the underlying group which gives the corresponding estimates. Thus, the choice of the group and the classical Markov semigroup on it are particularly important.

We now discuss two particular examples of collective decoherence already introduced in Section 6.4.2, namely the weak and the strong collective decoherences.

Weak collective decoherence We recall the generator of the weak collective decoherence on n qubits:

$$\mathcal{L}_n^{\mathrm{wcd}}(X) := \Sigma_z^n \, X \, \Sigma_z^n - \frac{1}{2}\big((\Sigma_z^n)^2 X + X(\Sigma_z^n)^2\big), \qquad \text{where} \qquad \Sigma_z^n := \sum_{i=1}^{n} \mathbb{1}_{\mathbb{C}^2}^{\otimes i-1} \otimes \sigma_z \otimes \mathbb{1}_{\mathbb{C}^2}^{n-i}. \tag{9.19}$$

One can easily show that the completely mixed state $2^{-n}\mathbb{1}_{(\mathbb{C}^2)^{\otimes n}}$ is invariant, since $\mathcal{L}_{n*}^{\mathrm{wcd}}(\mathbb{1}_{(\mathbb{C}^2)^{\otimes n}}) = 0$. Moreover, since Σ_z^n is self-adjoint, $\mathcal{L}_n^{\mathrm{wcd}}$ is KMS-symmetric with respect to that state.

Proposition 9.3.2. *For any $n \geq 2$, $\lambda(\mathcal{L}_n^{\mathrm{wcd}}) = 2$.*

Proof. Finding the spectral gap of the generator $\mathcal{L}_n^{\mathrm{wcd}}$ is equivalent to finding the spectral gap of its matrix representation $\tilde{\mathcal{L}}_n^{\mathrm{wcd}}$ (see e.g. [Wolf, 2012]): denote by $\{|i\rangle\}_{i=0,1}$ the eigenbasis of σ_z. Then,

$$\tilde{\mathcal{L}}_n^{\mathrm{wcd}} := \Sigma_z^n \otimes \Sigma_z^n - \frac{1}{2}\left((\Sigma_z^n)^2 \otimes \mathbb{1} + \mathbb{1} \otimes (\Sigma_z^n)^2\right),$$

where $\Sigma_z^n|i_1,...,i_n\rangle = \sum_{j=1}^{n}(-1)^{i_j}|i_1...i_n\rangle$ for any $(i_1,...,i_n) \in \{0,1\}^n$, so that

$$\tilde{\mathcal{L}}_n^{\mathrm{wcd}}|i_1...i_n\rangle \otimes |j_1...j_n\rangle = \left[\sum_{k=1}^{n}(-1)^{i_k}\sum_{k=1}^{n}(-1)^{j_k} - \frac{1}{2}\sum_{k,l=1}^{n}(-1)^{i_k+i_l} - \frac{1}{2}\sum_{k,l=1}^{n}(-1)^{j_k+j_l}\right]|i_1...i_n\rangle \otimes |j_1...j_n\rangle$$

$$= -2(|\mathbf{i}| - |\mathbf{j}|)^2|i_1...i_n\rangle \otimes |j_1...j_n\rangle,$$

where $|\mathbf{i}|$, resp. $|\mathbf{j}|$, denotes the number of 1's in the string $(i_1,...,i_n)$, resp. $(j_1,...,j_n)$. Therefore, the spectral gap of $\mathcal{L}_n^{\mathrm{wcd}}$ is equal to 2. $\qquad\square$

From this theorem and the universal upper bound found in Corollary 8.4.11, we find that the weak collective decoherence satisfies $\mathrm{HC}(c, \ln\sqrt{2})$ with

$$c \leq \frac{n\ln 2 + 2}{2}. \tag{9.20}$$

Next, let us first consider the heat diffusion on the one dimensional torus $\mathbb{T}^1$, which we represent on $(\mathbb{C}^2)^{\otimes n}$ as follows:

$$\mathbb{T}^1 \ni \theta \mapsto (e^{i\theta\sigma_z})^{\otimes n}.$$

One can easily verify that the QMS transferred via the above representation is the weak collective decoherence semigroup (up to a rescaling of the Lindblad operators by a factor of $\sqrt{2}$) as a direct consequence of Equation (9.17). Then, by Corollary 9.2.2 and the estimate of Section 3.7.2, we find for all $t \geq 0$, and any $\rho \in \mathcal{D}(\mathcal{H})$:

$$\|\mathcal{P}_t^{\mathrm{wcd},n}(\rho - E_{\mathcal{F}*}(\rho))\|_1 \leq \sqrt{2 + \sqrt{\pi/t}}\, e^{-\frac{t}{2}},$$

which represents a fast convergence independent of the size n of the system.

Strong collective decoherence

We recall the generator of the strong collective decoherence on n qubits:

$$\mathcal{L}_n^{\mathrm{scd}}(X) := \sum_{i \in \{x,y,z\}} \Sigma_i^n \, X \, \Sigma_i^n - \frac{1}{2}((\Sigma_i^n)^2 X + X(\Sigma_i^n)^2), \qquad \text{where} \qquad \Sigma_i^n := \sum_{k=1}^n \mathbb{1}_{\mathbb{C}^2}^{\otimes k-1} \otimes \sigma_i \otimes \mathbb{1}_{\mathbb{C}^2}^{n-k},$$

where the difference with Equation (9.19) arises from the consideration of all three Lindblad operator Σ_x^n, Σ_y^n and Σ_z^n.

We consider the three-dimensional simple Lie group $SU(2)$ of associated generators σ_x, σ_y, and σ_z spanning the Lie algebra $\mathfrak{su}(2)$, as well as the n-fold representation $SU(2) \ni g \mapsto U_g^{\otimes n}$, where U denotes the defining spin 1/2 representation of $SU(2)$: for any $\psi \in \mathbb{C}^2$, and $g \in SU(2)$,

$$U_g \psi = g \psi.$$

Just like previously, an easy use of Equation (9.17) shows that the semigroup transferred from the heat semigroup on $SU(2)$ via the above tensor product representation coincides with the strong decoherence (up to a rescaling of the Lindblad operators by a factor $\sqrt{2}$).

An easy application of Corollary 9.2.2 and the estimate of Theorem 3.7.1 for $n = 3$ provides the following dimension-independent bound for the decoherence time of the strong collective decoherence:

$$\tau_{\mathrm{deco}}(\varepsilon) \leq \frac{64}{3} - \frac{32}{3}\ln \varepsilon + 4\ln\left(1 + \frac{3}{2}\ln \frac{3}{4}\right).$$

Chapter 10.

Tensorization

The great advantage of classical logarithmic Sobolev inequalities over other methods resides in their tensorization property: the strong log-Sobolev constant of the product of independent Markovian evolutions is equal to the maximum over the set of strong log-Sobolev constants of the individual evolutions. However, this claim is strongly believed to be false in the non-commutative setting.

More precisely, given a set $\{\mathcal{L}_k\}_{k \in \{1, \cdots, N\}}$ of Lindblad generators acting on the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on a finite dimensional Hilbert space $\mathcal{H}$, define

$$\widehat{\mathcal{L}}_k := \mathrm{id}^{\otimes(k-1)} \otimes \mathcal{L}_k \otimes \mathrm{id}^{\otimes(N-k)} \tag{10.1}$$

as an operator acting on $\mathcal{B}(\mathcal{H}^{\otimes N})$. We also set

$$\mathcal{K}_N := \sum_{k=1}^{N} \widehat{\mathcal{L}}_k. \tag{10.2}$$

Observe that if each $\mathcal{L}_k$ is primitive/satisfies σ_k-DBC/is KMS-symmetric with respect to a state σ_k, then $\mathcal{K}_N$ is primitive/satisfies σ-DBC/is KMS-symmetric with respect to $\sigma^{(N)} := \bigotimes_{k=1}^{N} \sigma_k$. Moreover, the $\widehat{\mathcal{L}}_k$'s commute with each other and

$$\left(\mathrm{e}^{t\,\mathcal{K}_N} \right)_{t \geq 0} = \left(\bigotimes_{k=1}^{N} \mathrm{e}^{t\,\mathcal{L}_k} \right)_{t \geq 0}.$$

Now we can ask how the logarithmic Sobolev constants of $(\mathrm{e}^{t\,\mathcal{K}_N})_{t \geq 0}$ are related to those for each individual $(\mathrm{e}^{t\,\mathcal{L}_k})_{t \geq 0}$. In the commutative case the answer is easy: for instance, $\alpha_2(\mathcal{K}_N)$ equals $\min_{k \in \{1, \cdots, N\}} \alpha_2(\mathcal{L}_k)$ for all N. In particular, if each local semigroup satisfies a (reverse) hypercontractivity inequality, then so does $(\mathrm{e}^{t\,\mathcal{K}_N})_{t \geq 0}$. One easy way to understand this is by noticing that, in the commutative case, operator norms are multiplicative, or equivalently, the entropy function satisfies a certain subadditivity property (see e.g., [Mossel et al., 2013]). The aforementioned property that, in the classical case, $\alpha_p(\mathcal{K}_N)$ is independent of N, is usually called the *tensorization property*. The latter turns out to be useful even when allowing the subsystems to interact, e.g. for the Glauber dynamics [Martinelli, 1999, Cesi, 2001, Dai Pra et al., 2002, Caputo et al., 2015, Marton, 2015].

Tensorization property of log-Sobolev constants of quantum Lindblad generators, unlike its classical counterpart, is highly non-trivial. Thus proving (reverse) hypercontractivity inequalities that are independent of N is a difficult problem in the non-commutative setting: [Montanaro and Osborne, 2010] proved such hypercontractivity inequalities for the qubit depolarizing channel (see also [Kastoryano and Temme, 2013]). King [King, 2014] generalized this result for all unital qubit QMS. [Cubitt et al., 2015] developed the theory of quantum reverse hypercontractivity inequalities

in the unital case in and proved some tensorization-type results. Also, [Cubitt et al., 2015, Temme et al., 2014] developed some techniques for proving bounds on log-Sobolev constants $\alpha_p(\mathcal{K}_N)$ that are independent of N. For example, in Theorem 9 of [Temme et al., 2014], a weaker result was shown to hold in finite dimensions, the proof of which relies on interpolation properties of $\mathbb{L}_p$ spaces:

Theorem 10.0.1. *Let $N \in \mathbb{N}^*$ and, for any $k \in \{1, \cdots, N\}$, let $(\mathcal{P}_t^k)_{t \geq 0}$ be a primitive reversible quantum Markov semigroup on $\mathcal{B}(\mathbb{C}^d)$ with respective stationary state σ_k and spectral gap λ_k. Then, the log-Sobolev constant $\alpha_2(\mathcal{K}_N)$ of the product QMS $(\otimes_{i=1}^N \mathcal{P}_t^k)_{t \geq 0}$ satisfies*

$$\frac{\min_k \lambda_k}{\ln(d_{\mathcal{H}}^4 \max_k \|\sigma_k^{-1}\|_\infty) + 11} \leq \alpha_2(\mathcal{K}_N) \leq \frac{\min_k \lambda_k}{2} \, . \tag{10.3}$$

The importance of this result lies in the independence of the lower bound of (10.3) in the number N of systems. However, this bound highly depends on the local dimensions d. In particular, it is not well-suited to the study of quantum diffusions.

From there, two approaches can be pursued. A more conservative one consists in showing that tensorization holds in specific examples. Another novel approach, initiated in the unital case by [Beigi and King, 2016], is to replace the use of non-commutative $\mathbb{L}_p$ norms in the definition of hypercontractivity by their completely bounded versions, which are known to be multiplicative [Devetak et al., 2006] (cf. Section 1.1.2). However, since these norms follow from specific amalgamated $\mathbb{L}_p$ norms associated to algebras of the form $\mathcal{N}_k = \mathcal{B}(\mathbb{C}^k) \otimes \mathbb{1}$, we know from Section 8.5 that there is no hope to get the cancellation of the weak hypercontractivity constant d. Since this constant increases linearly with the number of systems under consideration, the problem of the independence of the strong constant c with N is only shifted.

Layout of the chapter: This chapter is divided into two main parts. In Section 10.1, we show the tensorization property of α_1 and α_2 for tensor products of the generalized depolarizing semigroup introduced in Section 5.5.1. Then, we introduce the related concepts of a complete logarithmic Sobolev inequality and of CB hypercontractivity in Section 10.2 for non-unital primitive QMS, which extends the framework of [Beigi and King, 2016] and Chapter 9.

10.1. Tensorizations for the generalized depolarizing semigroup

10.1.1. Quasi-tensorization of α_1

Theorem 10.1.1 provides a uniform bound on α_1 for the *generalized depolarizing semigroups* and their tensor powers. The proof of this result is a generalization of the proof of a similar result in the classical case [Mossel et al., 2013]. This tensorization result together with the quantum Stroock-Varopoulos inequality (Theorem 5.4.2) provides the tensorization of reverse hypercontractivity inequality which we use in Chapter 13.

Theorem 10.1.1. *Let $\sigma_1, \ldots, \sigma_N \in \mathcal{D}_+(\mathbb{C}^d)$. Let $\mathcal{L}_{\sigma_k}(X) = \mathrm{Tr}(\sigma_k X)\mathbb{1}_{\mathbb{C}^d} - X$ be the generator of the generalized depolarizing semigroup associated to the state σ_k, let $\widehat{\mathcal{L}}_{\sigma_k} := \mathrm{id}^{\otimes(k-1)} \otimes \mathcal{L}_{\sigma_k} \otimes \mathrm{id}^{\otimes(N-k)}$ and define $\mathcal{K}_N$ by (10.2). Then $\mathrm{MLSI}(\alpha_1)$ holds for $\alpha_1 \geq \frac{1}{4}$. In particular, this bound is independent of N and d.*

Proof. Let $\mathcal{H}_{A^N} = (\mathbb{C}^d)^{\otimes N}$ and $\mathcal{H}_{A_k} \simeq \mathbb{C}^d$ be the k-th local system. We need to show that for all $\rho_{A^N} \in \mathcal{D}_+(\mathcal{H}_{A^N})$:

$$D(\rho_{A^N} \| \sigma_{A^N}) \leq \mathrm{EP}_{\mathcal{K}_N}(\rho_{A^N}) \, ,$$

where $\sigma_{A_k} = \sigma_k$ and $\sigma_{A^N} = \sigma_1 \otimes \cdots \otimes \sigma_N$. Using the expression for the generator $\mathcal{K}_N$, we thus need to show that

$$D(\rho_{A^N}\|\sigma_{A^N}) \leq -\sum_{k=1}^N \text{Tr}\Big[\widehat{\mathcal{L}}_{\sigma_k*}(\rho_{A^N}) \cdot \big(\ln\rho_{A^N} - \ln(\sigma_{A^N})\big)\Big] \tag{10.4}$$

$$= \sum_{k=1}^N \text{Tr}\Big[\big(\rho_{A^N} - \rho_{A^{\sim k}} \otimes \sigma_{A_k}\big) \cdot \big(\ln\rho_{A^N} - \ln(\sigma_{A^N})\big)\Big] \tag{10.5}$$

$$= \sum_{k=1}^N \Big[D(\rho_{A^N}\|\sigma_{A^N}) + D(\rho_{A^{\sim k}} \otimes \sigma_{A_k}\|\rho_{A^N}) - D(\rho_{A^{\sim k}} \otimes \sigma_{A_k}\|\sigma_{A^N})\Big]. \tag{10.6}$$

where $A^{\sim k} = (A_1, \cdots, A_{k-1}, A_{k+1}, \cdots, A_N)$ and $\rho_{A^{\sim k}} = \text{Tr}_{A_k}(\rho_{A^N})$ is the partial trace of ρ_{A^N} with respect to the k-th subsystem. Now since $D(\rho_{A^{\sim k}} \otimes \sigma_{A_k}\|\rho_{A^N}) \geq 0$, it suffices to show that

$$D(\rho_{A^N}\|\sigma_{A^N}) \leq \sum_{k=1}^N \Big[D(\rho_{A^N}\|\sigma_{A^N}) - D(\rho_{A^{\sim k}} \otimes \sigma_{A_k}\|\sigma_{A^N})\Big], \tag{10.7}$$

We note that $D(\xi_B\|\tau_B) = -S(B)_\xi - \text{Tr}(\xi\ln\tau)$ where $S(B)_\xi = -\text{Tr}(\xi\ln\xi)$ is the von Neumann entropy. Moreover, $\ln(\xi \otimes \tau) = \ln\xi \otimes \mathbb{1} + \mathbb{1} \otimes \ln\tau$. Therefore, (10.7) is equivalent to

$$-S(A^N)_\rho - \sum_{k=1}^N \text{Tr}(\rho_{A_k}\ln\sigma_k) \leq \sum_{k=1}^N \Big[-S(A^N)_\rho - \sum_{j=1}^N \text{Tr}(\rho_{A_j}\ln\sigma_j) + S(A^{\sim k})_\rho + \sum_{j\neq k}\text{Tr}(\rho_{A_j}\ln\sigma_j)\Big]$$

$$= \sum_{k=1}^N \Big[-S(A^N)_\rho - \text{Tr}(\rho_{A_k}\ln\sigma_k) + S(A^{\sim k})_\rho\Big]$$

$$= \sum_{k=1}^N \Big[-S(A_k|A^{\sim k})_\rho - \text{Tr}(\rho_{A_k}\ln\sigma_k)\Big].$$

This is itself is equivalent to

$$S(A^N)_\rho \geq \sum_{k=1}^N S(A_k|A^{\sim k})_\rho,$$

which is an immediate consequence of the data processing inequality for the partial trace, i.e., $S(B|C)_\xi \geq S(B|CD)_\xi$ (cf. DPI), once we use the chain rule

$$S(A^N)_\rho = S(A_1)_\rho + \sum_{k=2}^N S(A_k|A_1, \cdots, A_{k-1})_\rho.$$

This concludes the proof. $\qquad\square$

Remark 10.1.2. Letting the σ_k's to be identical in the above theorem, we obtain the promised tensorization-type result for the modified logarithmic Sobolev constant. Note that this result was independently obtained recently in [Capel et al., 2018] by introducing the notion of a conditional logarithmic Sobolev constant and finding a uniform lower bound on the latter. Moreover, a special case of the above theorem corresponding to σ being the completely mixed state was already proved in [Müller-Hermes et al., 2016].

We can now use Corollary 7.4.10 and the fact that tensor products of generalized depolarizing semigroups satisfy ω-DBC to conclude the following.

Corollary 10.1.3. *Let* $\sigma_1, \ldots, \sigma_N$ *be arbitrary positive definite density matrices. Let* $\mathcal{L}_{\sigma_k}(X) = \text{Tr}(\sigma_k X)\mathbb{1} - X$ *be the generator associated to the generalized depolarizing semigroup* $\mathcal{P}_{k,t}^{\text{depol}}(X) = e^{-t}X + (1-e^{-t})\text{Tr}(\sigma_k X)\mathbb{1}$. *Define* $\sigma^{(N)} = \sigma_1 \otimes \cdots \otimes \sigma_N$ *and* $\mathcal{P}_t^{\text{depol},(N)} = \mathcal{P}_{1,t}^{\text{depol}} \otimes \cdots \otimes \mathcal{P}_{N,t}^{\text{depol}}$. *Then for*

$p \leq q < 1$ *and* $t \geq \ln \frac{p-1}{q-1}$, *we have for all* $N \in \mathbb{N}$ *and any positive definite operator* $X \in \mathcal{B}(\mathcal{H}^{\otimes N})$:

$$\left\| \mathcal{P}_t^{\mathrm{depol},(N)}(X) \right\|_{\mathbb{L}_p(\sigma^{(N)})} \geq \| X \|_{\mathbb{L}_q(\sigma^{(N)})}.$$

10.1.2. Tensorization of α_2

The second tensorization result, Theorem 10.1.4, shows that α_2 is independent of N in the case of qubits. In Corollary 10.1.8 we use these results to establish a uniform bound on the 2-logarithmic Sobolev constant of *any* qubit quantum Markov semigroup and its tensor powers. We note that the latter bound improves over the bounds provided in [Temme et al., 2014].

Before stating the tensorization results of this section, let us briefly explain the ideas behind their proof. Previously, Theorem 10.1.4 was known in the doubly stochastic case (the usual depolarizing semigroup), the proof of which was based on an inequality on the norms of a 2×2 block matrix and its submatrices from [King, 2003]. Our proof of Theorem 10.1.4 is based on the same inequality. First in Lemma 10.1.5 we derive an infinitesimal version of that inequality in terms of the entropies of a 2×2 block matrix and its submatrices, and then use it to prove Theorem 10.1.4. Finally, Corollary 10.1.8 is a quantum generalization of a classical result from [Diaconis and Saloff-Coste, 1996a] with an essentially similar proof except that tensorization is taken care of separately.

Theorem 10.1.4. *Let* $\dim \mathcal{H} = 2$ *and* $\mathcal{L}_\sigma(X) = \mathrm{Tr}(\sigma X)\mathbb{1} - X$ *for some positive definite density matrix* $\sigma \in \mathcal{D}_+(\mathcal{H})$. *Then, for any* $N \in \mathbb{N}$,

$$\alpha_2(\mathcal{K}_N) = \alpha_2(\mathcal{L}_\sigma),$$

where $\mathcal{K}_N$ *is defined in* (10.2) *when taking* $\mathcal{L}_k = \mathcal{L}_\sigma$ *for all* k.

In order to prove this theorem, our main tool is the following entropic inequality that is of independent interest and can be useful elsewhere.

Lemma 10.1.5. *Let* $\mathcal{H}$ *and* $\mathcal{H}'$ *be Hilbert spaces with* $\dim \mathcal{H} = 2$. *Let* $X \in \mathcal{B}_{sa}(\mathcal{H})^+$ *be a positive semidefinite matrix with the block form*

$$X = \begin{pmatrix} A & C \\ C^* & B \end{pmatrix}, \tag{10.8}$$

where $A, B, C \in \mathcal{B}(\mathcal{H}')$. *For a full-rank density matrix* ρ *on* $\mathcal{H}'$, *the matrix* M *defined as*

$$M = \begin{pmatrix} \|A\|_{\mathbb{L}_2(\rho)} & \|C\|_{\mathbb{L}_2(\rho)} \\ \|C^*\|_{\mathbb{L}_2(\rho)} & \|B\|_{\mathbb{L}_2(\rho)} \end{pmatrix} \tag{10.9}$$

is positive semidefinite. Moreover, let σ *be a full-rank density matrix on* $\mathcal{H}$ *of the form*

$$\sigma = \begin{pmatrix} \theta & 0 \\ 0 & 1 - \theta \end{pmatrix}, \tag{10.10}$$

where $\theta \in (0, 1)$. *Then we have*

$$\begin{aligned} \mathrm{Ent}_{2,\sigma \otimes \rho}(X) \leq{} & \mathrm{Ent}_{2,\sigma}(M) + \theta \, \mathrm{Ent}_{2,\rho}(A) + (1 - \theta) \, \mathrm{Ent}_{2,\rho}(B) \\ & + \sqrt{\theta(1-\theta)} \, \mathrm{Ent}_{2,\rho}(\mathrm{I}_{2,2}(C)) + \sqrt{\theta(1-\theta)} \, \mathrm{Ent}_{2,\rho}(\mathrm{I}_{2,2}(C^*)). \end{aligned} \tag{10.11}$$

Proof. For any $p \geq 2$ define

$$M_p := \begin{pmatrix} \|A\|_{\mathbb{L}_p(\rho)} & \|C\|_{\mathbb{L}_p(\rho)} \\ \|C^*\|_{\mathbb{L}_p(\rho)} & \|B\|_{\mathbb{L}_p(\rho)} \end{pmatrix},$$

so that $M_2 = M$. Since $X \geq 0$, both A and B are positive semidefinite. Moreover, we have

$$\Gamma_{\mathbb{1}\otimes\rho}^{\frac{1}{p}}(X) = \begin{pmatrix} \Gamma_\rho^{\frac{1}{p}}(A) & \Gamma_\rho^{\frac{1}{p}}(C) \\ \Gamma_\rho^{\frac{1}{p}}(C^*) & \Gamma_\rho^{\frac{1}{p}}(B) \end{pmatrix} \geq 0.$$

As a result, according to Theorem IX.5.9 of [Bhatia, 2015] there exists a *contraction* $R \in \mathcal{B}(\mathcal{H}')$ such that $\Gamma_\rho^{\frac{1}{p}}(C) = \left(\Gamma_\rho^{\frac{1}{p}}(A)\right)^{\frac{1}{2}} R \left(\Gamma_\rho^{\frac{1}{p}}(B)\right)^{\frac{1}{2}}$. Therefore, by Hölder's inequality we have

$$\begin{aligned}
\left\|\Gamma_\rho^{\frac{1}{p}}(C)\right\|_p &= \left\|\left(\Gamma_\rho^{\frac{1}{p}}(A)\right)^{\frac{1}{2}} R \left(\Gamma_\rho^{\frac{1}{p}}(B)\right)^{\frac{1}{2}}\right\|_p \\
&\leq \left\|\left(\Gamma_\rho^{\frac{1}{p}}(A)\right)^{\frac{1}{2}}\right\|_{2p} \cdot \|R\|_\infty \cdot \left\|\left(\Gamma_\rho^{\frac{1}{p}}(B)\right)^{\frac{1}{2}}\right\|_{2p} \\
&\leq \left\|\left(\Gamma_\rho^{\frac{1}{p}}(A)\right)^{\frac{1}{2}}\right\|_{2p} \cdot \left\|\left(\Gamma_\rho^{\frac{1}{p}}(B)\right)^{\frac{1}{2}}\right\|_{2p} \\
&= \left\|\Gamma_\rho^{\frac{1}{p}}(A)\right\|_p^{\frac{1}{2}} \cdot \left\|\Gamma_\rho^{\frac{1}{p}}(B)\right\|_p^{\frac{1}{2}}.
\end{aligned}$$

Then using $\|Y\|_{\mathbb{L}_p(\rho)} = \|\Gamma_\rho^{\frac{1}{p}}(Y)\|_p$, we find that

$$\|C\|_{\mathbb{L}_p(\rho)} \leq \|A\|_{\mathbb{L}_p(\rho)}^{\frac{1}{2}} \cdot \|B\|_{\mathbb{L}_p(\rho)}^{\frac{1}{2}},$$

and hence $M_p \geq 0$. In particular, $M_2 = M \geq 0$ and $\mathrm{Ent}_{2,\rho}(M)$ is well-defined.

Define $\psi(p) := \|M_p\|_{\mathbb{L}_p(\sigma)} - \|X\|_{\mathbb{L}_p(\sigma\otimes\rho)}$. It is shown by King [King, 2003] that $\psi(p) \geq 0$ for all $p \geq 2$. Indeed, this inequality is proven in [King, 2003] in the special case where σ and ρ are the identity operators on the relevant spaces. Nevertheless, we have

$$\|X\|_{\mathbb{L}_p(\sigma\otimes\rho)} = \left\|\begin{pmatrix} \theta^{\frac{1}{p}}\Gamma_\rho^{\frac{1}{p}}(A) & (\theta(1-\theta))^{\frac{1}{2p}}\Gamma_\rho^{\frac{1}{p}}(C) \\ (\theta(1-\theta))^{\frac{1}{2p}}\Gamma_\rho^{\frac{1}{p}}(C^*) & (1-\theta)^{\frac{1}{p}}\Gamma_\rho^{\frac{1}{p}}(B) \end{pmatrix}\right\|_p,$$

and

$$\|M_p\|_{\mathbb{L}_p(\sigma)} = \left\|\begin{pmatrix} \theta^{\frac{1}{p}}\left\|\Gamma_\rho^{\frac{1}{p}}(A)\right\|_p & (\theta(1-\theta))^{\frac{1}{2p}}\left\|\Gamma_\rho^{\frac{1}{p}}(C)\right\|_p \\ (\theta(1-\theta))^{\frac{1}{2p}}\left\|\Gamma_\rho^{\frac{1}{p}}(C^*)\right\|_p & (1-\theta)^{\frac{1}{p}}\left\|\Gamma_\rho^{\frac{1}{p}}(B)\right\|_p \end{pmatrix}\right\|_p,$$

Thus, King's result holds for arbitrary ρ and diagonal σ as well, and we have $\psi(p) \geq 0$ for all $p \geq 2$. On the other hand, a straightforward computation verifies that $\psi(2) = 0$. This means that $\psi'(2) \geq 0$, i.e.,

$$\frac{d}{dp}\left(\|M_p\|_{\mathbb{L}_p(\sigma)} - \|X\|_{\mathbb{L}_p(\sigma\otimes\rho)}\right)\bigg|_{p=2} \geq 0.$$

The derivatives can be computed using Proposition 7.4.7. We have

$$\frac{d}{dp}\|X\|_{\mathbb{L}_p(\sigma\otimes\rho)}\bigg|_{p=2} = \frac{1}{4}\|X\|_{\mathbb{L}_2(\sigma\otimes\rho)}^{-1} \cdot \mathrm{Ent}_{2,\sigma\otimes\rho}(X), \tag{10.12}$$

and

$$\frac{d}{dp}\|M_p\|_{\mathbb{L}_p(\sigma)}\bigg|_{p=2} = \frac{1}{4}\|M\|_{\mathbb{L}_2(\sigma)}^{-1} \cdot \left(\mathrm{Ent}_{2,\sigma}(M) + 4\,\mathrm{Tr}\left[\Gamma_\sigma^{\frac{1}{2}}(M_2') \cdot \Gamma_\sigma^{\frac{1}{2}}(M)\right]\right),$$

where

$$M_2' = \frac{d}{dp} M_p \Big|_{p=2} = \frac{1}{4} \begin{pmatrix} \|A\|_{\mathbb{L}_2(\rho)}^{-1} \cdot \mathrm{Ent}_{2,\rho}(A) & w \\ w & \|B\|_{\mathbb{L}_2(\rho)}^{-1} \cdot \mathrm{Ent}_{2,\rho}(B) \end{pmatrix},$$

and $w = \|C\|_{\mathbb{L}_2(\rho)}^{-1} \cdot \left(\mathrm{Ent}_{2,\rho}\left(\mathrm{I}_{2,2}(C) \right) + \mathrm{Ent}_{2,\rho}\left(\mathrm{I}_{2,2}(C^*) \right) \right)$. We conclude that

$$\frac{d}{dp} \|M_p\|_{\mathbb{L}_p(\sigma)} \Big|_{p=2} = \frac{1}{4} \|M\|_{\mathbb{L}_2(\sigma)}^{-1} \cdot \Big(\mathrm{Ent}_{2,\sigma}(M) + \theta \, \mathrm{Ent}_{2,\rho}(A) + (1-\theta) \, \mathrm{Ent}_{2,\rho}(B)$$
$$+ \sqrt{\theta(1-\theta)} \, \mathrm{Ent}_{2,\rho}\left(\mathrm{I}_{2,2}(C) \right) + \sqrt{\theta(1-\theta)} \, \mathrm{Ent}_{2,\rho}\left(\mathrm{I}_{2,2}(C^*) \right) \Big).$$

Comparing to (10.12) and using $\|M\|_{\mathbb{L}_2(\sigma)} = \|X\|_{\mathbb{L}_2(\sigma \otimes \rho)}$ the desired inequality follows.

$\square$

We need yet another lemma to prove Theorem 10.1.4.

Lemma 10.1.6. *For any Lindblad generator $\mathcal{K}$ that is KMS-symmetric with respect to some positive definite density matrix ρ we have, for any matrix C,*

$$\mathcal{E}_{2,\mathcal{K}}\left(\mathrm{I}_{2,2}(C) \right) + \mathcal{E}_{2,\mathcal{K}}\left(\mathrm{I}_{2,2}(C^*) \right) \leq -\langle C, \mathcal{K}(C) \rangle_\rho - \langle C^*, \mathcal{K}(C^*) \rangle_\rho.$$

Proof. Define $D := \Gamma_\rho^{\frac{1}{2}}(C)$. Then for $j \in \{0,1\}$ [Bhatia, 2015]:

$$Y_j := \begin{pmatrix} |D| & (-1)^j D^* \\ (-1)^j D & |D^*| \end{pmatrix} \geq 0.$$

Since $\Gamma_\rho^{-1/2}$ is completely positive we have

$$Z_j := \mathrm{id} \otimes \Gamma_\rho^{-1/2}(Y_j) = \begin{pmatrix} \mathrm{I}_{2,2}(C) & (-1)^j C^* \\ (-1)^j C & \mathrm{I}_{2,2}(C^*) \end{pmatrix} \geq 0.$$

On the other hand, $\mathcal{P}_t = e^{t\mathcal{K}}$ is completely positive. Therefore,

$$\mathrm{id} \otimes \mathcal{P}_t(Z_0) = \begin{pmatrix} \mathcal{P}_t(\mathrm{I}_{2,2}(C)) & \mathcal{P}_t(C^*) \\ \mathcal{P}_t(C) & \mathcal{P}_t(\mathrm{I}_{2,2}(C^*)) \end{pmatrix} \geq 0.$$

Putting these observations together we find that for any $t \geq 0$:

$$g(t) := \langle Z_1, \mathrm{id} \otimes \mathcal{P}_t(Z_0) \rangle_{\mathbb{1} \otimes \rho} \geq 0.$$

We note that

$$g(t) = \left\langle \mathrm{I}_{2,2}(C), \mathcal{P}_t(\mathrm{I}_{2,2}(C)) \right\rangle_\rho + \left\langle \mathrm{I}_{2,2}(C^*), \mathcal{P}_t(\mathrm{I}_{2,2}(C^*)) \right\rangle_\rho - \left\langle C, \mathcal{P}_t(C) \right\rangle_\rho - \left\langle C^*, \mathcal{P}_t(C^*) \right\rangle_\rho.$$

From this expression it is clear that

$$g(0) = \|\mathrm{I}_{2,2}(C)\|_{\mathbb{L}_2(\rho)}^2 + \|\mathrm{I}_{2,2}(C^*)\|_{\mathbb{L}_2(\rho)}^2 - \|C\|_{\mathbb{L}_2(\rho)}^2 - \|C^*\|_{\mathbb{L}_2(\rho)}^2 = 0.$$

Therefore, we must have $g'(0) \geq 0$ which is equivalent to the desired inequality. $\square$

Now we have all the required tools for proving Theorem 10.1.4. We actually prove a stronger statement from which Theorem 10.1.4 is implied by a simple induction procedure:

Theorem 10.1.7. *Let $\dim \mathcal{H} = 2$ and $\mathcal{L}_\sigma(X) = \mathrm{Tr}(\sigma X)\mathbb{1} - X$ for some positive definite density matrix*

σ. Also let $\mathcal{K}$ be a Lindblad generator associated to a primitive QMS that is reversible with respect to some positive definite state $\rho \in \mathcal{D}(\mathcal{H}')$. Then we have

$$\alpha_2(\mathcal{L}_\sigma \otimes \mathrm{id}' + \mathrm{id} \otimes \mathcal{K}) = \min\{\alpha_2(\mathcal{L}_\sigma), \alpha_2(\mathcal{K})\},$$

where id and id' denote the identity superoperators acting on $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}(\mathcal{H}')$ respectively.

Proof. Let $\alpha = \min\{\alpha_2(\mathcal{L}_\sigma), \alpha_2(\mathcal{K})\}$. By restricting X in the log-Sobolev inequality of order 2 to be of the tensor product form and using

$$\mathrm{Ent}_{2,\sigma\otimes\rho}(Y \otimes Y') = \mathrm{Ent}_{2,\sigma}(Y) + \mathrm{Ent}_{2,\rho}(Y'),$$

we conclude that $\alpha_2(\mathcal{L}_\sigma \otimes \mathrm{id}' + \mathrm{id} \otimes \mathcal{K}) \le \alpha$. To prove the inequality in the other direction we need to show that for any $X \in \mathcal{B}_{sa}(\mathcal{H} \otimes \mathcal{H}')^+$ we have

$$\alpha \, \mathrm{Ent}_{2,\sigma\otimes\rho}(X) \le \mathcal{E}_{2,\mathcal{L}_\sigma\otimes\mathrm{id}'+\mathrm{id}\otimes\mathcal{K}}(X). \tag{10.13}$$

Assume, without loss of generality, that σ is diagonal of the form (10.10), and that $X \in \mathcal{B}_{sa}(\mathcal{H} \otimes \mathcal{H}')^+$ has the block form (10.8). Define M by (10.9). Then by Lemma 10.1.5 we have

$$\begin{aligned}
\mathrm{Ent}_{2,\sigma\otimes\rho}(X) \le \ & \mathrm{Ent}_{2,\sigma}(M) + \theta\,\mathrm{Ent}_{2,\rho}(A) + (1-\theta)\,\mathrm{Ent}_{2,\rho}(B) \\
& + \sqrt{\theta(1-\theta)}\,\mathrm{Ent}_{2,\rho}(\mathrm{I}_{2,2}(C)) + \sqrt{\theta(1-\theta)}\,\mathrm{Ent}_{2,\rho}(\mathrm{I}_{2,2}(C^*)).
\end{aligned}$$

On the other hand by the definition of α we have

$$\alpha \, \mathrm{Ent}_{2,\sigma}(M) \le \mathcal{E}_{2,\mathcal{L}_\sigma}(M),$$

and

$$\alpha \, \mathrm{Ent}_{2,\rho}(Y) \le \mathcal{E}_{2,\mathcal{K}}(Y),$$

for all $Y \in \{A, B, \mathrm{I}_{2,2}(C), \mathrm{I}_{2,2}(C^*)\}$. Therefore, we have

$$\begin{aligned}
\alpha \, \mathrm{Ent}_{2,\sigma\otimes\rho}(X) \le \ & \mathcal{E}_{2,\mathcal{L}_\sigma}(M) + \theta\mathcal{E}_{2,\mathcal{K}}(A) + (1-\theta)\mathcal{E}_{2,\mathcal{K}}(B) \\
& + \sqrt{\theta(1-\theta)}\,\mathcal{E}_{2,\mathcal{K}}(\mathrm{I}_{2,2}(C)) + \sqrt{\theta(1-\theta)}\,\mathcal{E}_{2,\mathcal{K}}(\mathrm{I}_{2,2}(C^*)) \\
\le \ & \mathcal{E}_{2,\mathcal{L}_\sigma}(M) + \theta\mathcal{E}_{2,\mathcal{K}}(A) + (1-\theta)\mathcal{E}_{2,\mathcal{K}}(B) \\
& - \sqrt{\theta(1-\theta)}\,\langle C, \mathcal{K}(C)\rangle_\rho - \sqrt{\theta(1-\theta)}\,\langle C^*, \mathcal{K}(C^*)\rangle_\rho, \tag{10.14}
\end{aligned}$$

where in the second inequality we use Lemma 10.1.6. We now have

$$\begin{aligned}
\mathcal{E}_{2,\mathcal{L}_\sigma\otimes\mathrm{id}'+\mathrm{id}\otimes\mathcal{K}}(X) &= -\langle X, (\mathcal{L}_\sigma \otimes \mathrm{id}' + \mathrm{id} \otimes \mathcal{K})(X)\rangle_{\sigma\otimes\rho} \\
&= -\langle X, \mathcal{L}_\sigma \otimes \mathrm{id}'(X)\rangle_{\sigma\otimes\rho} - \left\langle \begin{pmatrix} A & C \\ C^* & B \end{pmatrix}, \begin{pmatrix} \mathcal{K}(A) & \mathcal{K}(C) \\ \mathcal{K}(C^*) & \mathcal{K}(B) \end{pmatrix} \right\rangle_{\sigma\otimes\rho}.
\end{aligned}$$

We compute each term in the above sum separately.

$$-\langle X, (\mathcal{L}_\sigma \otimes \mathrm{id}')(X)\rangle_{\sigma \otimes \rho}$$

$$= \left\langle \begin{pmatrix} A & C \\ C^* & B \end{pmatrix}, \begin{pmatrix} (1-\theta)(A-B) & C \\ C^* & \theta(B-A) \end{pmatrix} \right\rangle_{\sigma \otimes \rho}$$

$$= \theta(1-\theta)\langle A, A-B\rangle_\rho + \theta(1-\theta)\langle B, B-A\rangle_\rho + 2\sqrt{\theta(1-\theta)}\langle C, C\rangle_\rho$$

$$= \theta(1-\theta)\|A\|^2_{\mathbb{L}_2(\rho)} + \theta(1-\theta)\|B\|^2_{\mathbb{L}_2(\rho)} - 2\theta(1-\theta)\langle A, B\rangle_\rho + 2\sqrt{\theta(1-\theta)}\|C\|^2_{\mathbb{L}_2(\rho)}$$

$$\geq \theta(1-\theta)\|A\|^2_{\mathbb{L}_2(\rho)} + \theta(1-\theta)\|B\|^2_{\mathbb{L}_2(\rho)} - 2\theta(1-\theta)\|A\|_{\mathbb{L}_2(\rho)} \cdot \|B\|_{\mathbb{L}_2(\rho)} + 2\sqrt{\theta(1-\theta)}\|C\|^2_{\mathbb{L}_2(\rho)}$$

$$= -\langle M, \mathcal{L}_\sigma(M)\rangle_\sigma$$

$$= \mathcal{E}_{2,\mathcal{L}_\sigma}(M).$$

For the second term we compute

$$-\left\langle \begin{pmatrix} A & C \\ C^* & B \end{pmatrix}, \begin{pmatrix} \mathcal{K}(A) & \mathcal{K}(C) \\ \mathcal{K}(C^*) & \mathcal{K}(B) \end{pmatrix} \right\rangle_{\sigma \otimes \rho} = -\theta\langle A, \mathcal{K}(A)\rangle_\rho - (1-\theta)\langle B, \mathcal{K}(B)\rangle_\rho$$

$$- \sqrt{\theta(1-\theta)}\langle C, \mathcal{K}(C)\rangle_\rho - \sqrt{\theta(1-\theta)}\langle C^*, \mathcal{K}(C^*)\rangle_\rho$$

$$= \theta\mathcal{E}_{2,\mathcal{K}}(A) + (1-\theta)\mathcal{E}_{2,\mathcal{K}}(B)$$

$$- \sqrt{\theta(1-\theta)}\langle C, \mathcal{K}(C)\rangle_\rho - \sqrt{\theta(1-\theta)}\langle C^*, \mathcal{K}(C^*)\rangle_\rho.$$

Therefore, we have

$$\mathcal{E}_{2,\mathcal{L}_\sigma \otimes \mathrm{id}' + \mathrm{id} \otimes \mathcal{K}}(X) \geq \mathcal{E}_{2,\mathcal{L}_\sigma}(M) + \theta\mathcal{E}_{2,\mathcal{K}}(A) + (1-\theta)\mathcal{E}_{2,\mathcal{K}}(B)$$

$$- \sqrt{\theta(1-\theta)}\langle C, \mathcal{K}(C)\rangle_\rho - \sqrt{\theta(1-\theta)}\langle C^*, \mathcal{K}(C^*)\rangle_\rho.$$

Comparing this to (10.14) we arrive at the desired inequality (10.13). □

We can now derive a tensorization-type result for a wide class of Lindblad generators. Let $\mathcal{L}$ be a σ-reversible and primitive Lindblad generator. Recall that the *spectral gap* of $\mathcal{L}$ is defined by

$$\lambda(\mathcal{L}) = \inf_X \frac{\mathcal{E}_{2,\mathcal{L}}(X)}{\mathrm{Var}_\sigma(X)},$$

where $\mathrm{Var}_\sigma(X) = \langle X, X\rangle_\sigma - \langle X, \mathbb{1}\rangle^2_\sigma = \|X\|^2_{\mathbb{L}_2(\sigma)} - \langle X, \mathbb{1}\rangle^2_\sigma$. Observe that $\mathrm{Var}_\sigma(X)$ is the squared length of the projection of X onto the subspace orthogonal to $\mathbb{1} \in \mathcal{B}(\mathcal{H})$ with respect to the inner product $\langle \cdot, \cdot \rangle_\sigma$. On the other hand, $\mathbb{1}$ is the sole[1] 0-eigenvector of $\mathcal{L}$ which is self-adjoint with respect to this inner product. Therefore, $-\lambda(\mathcal{L})$ is the minimum non-zero eigenvalue of $\mathcal{L}$.

The spectral gap satisfies the tensorization property, as shown below. Observe that

$$\mathcal{K}_N = \sum_{k=1}^N \widehat{\mathcal{L}}_k$$

is a sum of mutually commuting operators. Then the eigenvalues of $\mathcal{K}_N$ are summations of eigenvalues of individual $\widehat{\mathcal{L}}_k$'s. Since each $\widehat{\mathcal{L}}_k$ is a tensor product of $\mathcal{L}$ with some identity superoperator, the set of its eigenvalues is the same as that of $\mathcal{L}$. Using these we conclude that for all $N \in \mathbb{N}$,

$$\lambda(\mathcal{K}_N) = \lambda(\mathcal{L}). \tag{10.15}$$

---

[1]This 0-eigenvector is unique since $\mathcal{L}$ is assumed to be primitive.

It is well-known that $\lambda(\mathcal{L}) \geq 2\alpha_2(\mathcal{L})$ (see [Carbone and Martinelli, 2015]). The following corollary gives a lower bound on $\alpha_2(\mathcal{L})$ in terms of $\lambda(\mathcal{L})$.

Corollary 10.1.8. *Let* $\dim \mathcal{H} = 2$ *and* $\sigma \in \mathcal{D}(\mathcal{H})$ *be full-rank. For any KMS-symmetric primitive Lindblad generator* $\mathcal{L}$ *with respect to* σ *we have*

$$\frac{1 - 2\,\|\sigma^{-1}\|_\infty^{-1}}{\ln\left(\|\sigma^{-1}\|_\infty - 1\right)}\,\lambda(\mathcal{L}) \leq \alpha_2(\mathcal{K}_n)\,.$$

Proof. Let $\mathcal{L}_\sigma$ be the generalized depolarizing Lindblad generator that is σ-reversible, and let $X \in \mathcal{B}_{sa}(\mathcal{H}^{\otimes N})^+$ be arbitrary. Then by Theorem 10.1.4 and Theorem 7.2.4 we have

$$\frac{1 - 2\|\sigma^{-1}\|_\infty^{-1}}{\ln\left(\|\sigma^{-1}\|_\infty - 1\right)}\,\mathrm{Ent}_{2,\sigma^{\otimes N}}(X) \leq -\sum_{k=1}^{N}\left\langle X, \widehat{\mathcal{L}}_{\sigma_k}(X)\right\rangle_{\sigma^{\otimes n}}, \tag{10.16}$$

where $\sigma_k = \sigma$ for all k. Next, let $\mathcal{W}_k \subset \mathcal{B}(\mathcal{H}^{\otimes N})$ be the subspace spanned by operators of the form $A_1 \otimes \cdots \otimes A_N \in \mathcal{B}(\mathcal{H}^{\otimes N})$ with $A_k = \mathbb{1} \in \mathcal{B}(\mathcal{H})$. In other words, $\mathcal{W}_k = \ker(\widehat{\mathcal{L}}_{\sigma_k})$. Then $-\left\langle X, \widehat{\mathcal{L}}_{\sigma_k}(X)\right\rangle_{\sigma^{\otimes N}}$ equals the squared length of the projection of X onto $\mathcal{W}_k^\perp$. On the other hand, since $\mathcal{L}$ is primitive and σ-reversible, we also have $\mathcal{W}_k = \ker \widehat{\mathcal{L}}_k$ and $\mathcal{W}_k^\perp$ is invariant under $\widehat{\mathcal{L}}_k$. Moreover, by definition $\lambda(\widehat{\mathcal{L}}_k)$ is the minimum eigenvalue of $\widehat{\mathcal{L}}_k$ restricted to $\mathcal{W}_k^\perp$ (i.e., the maximum non-zero eigenvalue). We conclude that

$$-\lambda(\widehat{\mathcal{L}}_k)\left\langle X, \widehat{\mathcal{L}}_{\sigma_k}(X)\right\rangle_{\sigma^{\otimes N}} \leq -\left\langle X, \widehat{\mathcal{L}}_k(X)\right\rangle_{\sigma^{\otimes N}}.$$

On the other hand since $\widehat{\mathcal{L}}_k$ equals the tensor product of $\mathcal{L}$ with some identity superoperators, $\lambda(\widehat{\mathcal{L}}_k) = \lambda(\mathcal{L})$. Therefore[2],

$$-\lambda(\mathcal{L})\left\langle X, \widehat{\mathcal{L}}_{\sigma_k}(X)\right\rangle_{\sigma^{\otimes N}} \leq -\left\langle X, \widehat{\mathcal{L}}_k(X)\right\rangle_{\sigma^{\otimes N}}.$$

Using this in (10.16) we arrive at

$$\lambda(\mathcal{L})\,\frac{1 - 2\|\sigma^{-1}\|_\infty^{-1}}{\ln\left(\|\sigma^{-1}\|_\infty - 1\right)}\,\mathrm{Ent}_{2,\sigma^{\otimes N}}(X) \leq -\sum_{k=1}^{N}\left\langle X, \widehat{\mathcal{L}}_k(X)\right\rangle_{\sigma^{\otimes N}} = -\left\langle X, \mathcal{K}_N(X)\right\rangle_{\sigma^{\otimes N}}.$$

This gives the desired bound on $\alpha_2(\mathcal{K}_N)$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 10.1.9. This corollary is a non-commutative version of Corollary A.4 of [Diaconis and Saloff-Coste, 1996a] and gives a stronger bound compared to Corollary 6 and Theorem 9 of [Temme et al., 2014]. It would be interesting to compare this corollary with the result of [King, 2014] who generalized the hypercontractivity inequalities of [Montanaro and Osborne, 2010] for the qubit depolarizing channel to all doubly stochastic qubit quantum Markov semigroups. Here, having a bound on the 2-logarithmic Sobolev constant of the KMS-symmetric generalized qubit depolarizing channel (and its tensorization property), we derive a bound on the 2-logarithmic Sobolev constant of all qubit σ-reversible QMS.

Corollary 10.1.10. *Let* $\dim \mathcal{H} = 2$ *and* $\sigma \in \mathcal{D}(\mathcal{H})$ *be full-rank. Let* $\mathcal{L}$ *be the generator of a KMS-symmetric QMS* $(\mathcal{P}_t)_{t\geq 0}$*. Then for any* $1 \leq q \leq p$ *and* $t \geq 0$ *satisfying*

$$t \geq \frac{\ln\left(\|\sigma^{-1}\|_\infty - 1\right)}{4\lambda(\mathcal{L})\left(1 - 2\|\sigma^{-1}\|_\infty^{-1}\right)}\,\ln\frac{p-1}{q-1}\,,$$

---

[2]This comparison of Dirichlet forms was already used in Section 8.5 in order to show the non-positivity of the strong decoherence free log-Sobolev constant by reducing the analysis to the case of the $\mathcal{N}$-decoherent semigroup.

we have $\|\mathcal{P}_t^{\otimes N}(X)\|_{\mathbb{L}_p(\sigma)} \leq \|X\|_{q,\sigma}$ *for all $X > 0$ and any $N \in \mathbb{N}$.*

10.2. CB hypercontractivity and complete log-Sobolev inequalities

As already discussed in the introduction of this chapter, in the classical case, log-Sobolev inequalities satisfy the very useful *tensorization property*, that is, given N primitive Markov semigroups $(P_t^{(i)})_{t \geq 0}$ with generators L_i, $i = 1, \cdots, N$, if for each i, the semigroup $(P_t^{(i)})_{t \geq 0}$ satisfies the log-Sobolev inequality $\text{LSI}_2(c_i, d_i)$, then the product semigroup $(P_t)_{t \geq 0}$ with $P_t = P_t^1 \otimes \cdots \otimes P_t^n$, satisfies the log-Sobolev inequality $\text{LSI}_2(\max_i c_i, \sum_i d_i)$. This can be seen as a consequence of the multiplicativity of the classical weighted $\mathbb{L}_p$ norms. It is strongly believed that this latter property no longer holds true in the quantum case, since quantum weighted $\mathbb{L}_p$ norms are not multiplicative. In [Beigi and King, 2016], the authors proposed to define the hypercontractivity property with respect to the completely bounded norm, which is known to be multiplicative even in the noncommutative framework, and proved that it is equivalent to the notion of a complete logarithmic Sobolev inequality for primitive doubly stochastic QMS. This provides a way to recover the tensorization property in the noncommutative framework. Here, we generalize the framework of [Beigi and King, 2016] to the case of any primitive (not necessarily doubly stochastic) QMS in finite dimensions. Recall that, given an operator $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, its *weighted completely bounded norm* is defined as follows (cf. Section 1.1.2):

$$\|\Phi : \mathbb{L}_q(\sigma) \to \mathbb{L}_p(\sigma)\|_{\text{cb}} := \sup_{n \in \mathbb{N}} \| \text{id}_{\mathbb{M}_n(\mathbb{C})} \otimes \Phi : \mathbb{L}_q(n^{-1} \mathbb{1}_{\mathbb{C}^n}, \mathbb{L}_q(\sigma)) \to \mathbb{L}_q(n^{-1} \mathbb{1}_{\mathbb{C}^n}, \mathbb{L}_p(\sigma))\|, \quad (10.17)$$

where the supremum in 10.17 is over all dimensions n. These norms are known to be multiplicative, as proved in [Devetak et al., 2006]. As a result, in order to define a notion of hypercontractivity and its associated log-Sobolev inequality that satisfy the tensorization property, we embed a primitive QMS $(\mathcal{P}_t)_{t \geq 0}$ on $\mathcal{B}(\mathcal{H})$ into the QMS $(\text{id}_k \otimes \mathcal{P}_t)_{t \geq 0}$ on $\mathcal{B}(\mathbb{C}^k \otimes \mathcal{H})$, and study the latter's DF-hypercontractivity properties, for each integer $k \geq 1$. Let σ be the unique invariant state of $(\mathcal{P}_t)_{t \geq 0}$. Then $\mathcal{N}_k := \mathcal{B}(\mathbb{C}^k) \otimes \mathbb{1}_{\mathcal{H}}$ and $\sigma_{\text{Tr}} = \frac{\mathbb{1}_k}{k} \otimes \sigma$. We are lead to the following definitions.

Definition 10.2.1. We say that $(\mathcal{P}_t)_{t \geq 0}$:

(i) satisfies a *q-complete logarithmic Sobolev inequality* with positive *strong logarithmic Sobolev constant* $c > 0$ and *weak logarithmic Sobolev constant* $d \geq 0$, which we denote by $\text{cLSI}_q(c, d)$, if for all integer $k \geq 1$, $\text{LSI}_{q, \mathcal{N}_k}(c, d)$ holds.

(ii) is *q-CB-hypercontractive* for positive constants $c > 0$ and $d \geq 0$, condition denoted by $\text{cHC}_q(c, d)$, if for all $t \geq 0$,

$$\|\mathcal{P}_t : \mathbb{L}_q(\sigma) \to \mathbb{L}_{p(t)}(\sigma)\|_{\text{cb}} \leq \exp\left(2d\left(\frac{1}{q} - \frac{1}{p(t)}\right)\right),$$

for any function $p : [0, +\infty) \to \mathbb{R}$ such that for any $t \geq 0$, $q \leq p(t) \leq 1 + (q-1) e^{\frac{4t}{c}}$. In the case $q = 2$, we simply denote the previous property by $\text{cHC}(c, d)$.

The above definitions extend the ones in [Beigi and King, 2016] to non-unital primitive QMS and to weak LSI and weak HC. In the next theorem, we establish the equivalence between the complete logarithmic Sobolev inequality and CB-hypercontractivity, hence extending Theorem 4 of [Beigi and King, 2016] to the cases mentioned above.

Theorem 10.2.2. *Let $(\mathcal{P}_t)_{t\geq 0}$ be a primitive QMS on $\mathcal{B}(\mathcal{H})$ with associated generator $\mathcal{L}$, and let $q \geq 1$, $d \geq 0$ and $p(t) = 1 + (q-1)\,\mathrm{e}^{\frac{4t}{c}}$ for some constant $c > 0$. Then*

(i) *If $\mathrm{cHC}_q(c,d)$ holds, then $\mathrm{cLSI}_q(c,d)$ holds.*

(ii) *If $\mathrm{cLSI}_{p(t)}(c,d)$ holds for all $t \geq 0$, then $\mathrm{cHC}_q(c,d)$ holds.*

Proof. We first prove (i). If $\mathrm{cHC}_q(c,d)$ holds, then for any k and any $X \in \mathcal{B}(\mathbb{C}^k \otimes \mathcal{H})$,

$$\|\,\mathrm{id}_k \otimes \mathcal{P}_t(X)\|_{\mathbb{L}_q(\mathcal{N}_k,\,\mathbb{L}_{p(t)}(\sigma_{\mathrm{Tr}}))} \leq \exp\left(2d\left(\frac{1}{q} - \frac{1}{p(t)}\right)\right) \|X\|_{\mathbb{L}_q\left(\frac{1_k}{k}\otimes\sigma\right)},$$

that is $\mathrm{HC}_{q,\mathcal{N}}(c,d)$ holds for the QMS $(\mathrm{id}_k \otimes \mathcal{P}_t)_{t\geq 0}$ of corresponding decoherence-free algebra $\mathcal{N}_k = \mathcal{B}(\mathbb{C}^k) \otimes \mathbb{1}_{\mathcal{H}}$ and $\sigma_{\mathrm{Tr}} = \frac{1_k}{k} \otimes \sigma$. The result then follows from a direct application of 8.2.2(i). (ii) follows similarly from 8.2.2(ii). $\qquad\square$

A direct application of the definitions for $\mathbb{L}_p$ regularity of Dirichlet forms then leads to the following:

Theorem 10.2.3. *Assume that $\mathrm{cLSI}_2(c,d)$ holds. Then, if the generator $\mathcal{L}$ is $\mathbb{L}_p$-regular for some $d_0 \geq 0$, then $\mathrm{cLSI}_q(c,d + c\,d_0/2)$ holds for all $q \geq 1$, so that $\mathrm{cHC}_2(c,d + c\,d_0/2)$ holds.*

As in the decoherence-free case, an application of Proposition 5.2 of [Olkiewicz and Zegarlinski, 1999] together with Theorem 4 of [Watrous, 2005] leads to the following corollary:

Corollary 10.2.4. *Assume that $\mathcal{L}$ is the generator of a primitive QMS with unique invariant state σ, and that $\mathrm{cLSI}_2(c,d)$ holds.*

(i) *If $\mathcal{L}$ is KMS-symmetric, then $\mathrm{cLSI}_q(c,d + c\,(\|\mathcal{L}:\mathbb{L}_2(\sigma) \to \mathbb{L}_2(\sigma)\| + 1$ holds for all $q \geq 1$ and consequently $\mathrm{cHC}_2(c,d + c\,(\|\mathcal{L}:\mathbb{L}_2(\sigma) \to \mathbb{L}_2(\sigma)\| + 1))$ holds.*

(ii) *If $\mathcal{L}$ satisfies σ-DBC, then $\mathrm{cLSI}_q(c,d)$ holds for all $q \geq 1$ and consequently $\mathrm{cHC}_2(c,d)$ holds.*

Proof. The result follows directly from the fact that reversibility of $\mathcal{L}$ w.r.t. σ implies reversibility of $\mathrm{id}_k \otimes \mathcal{L}$ w.r.t. σ_{Tr}, for any $k \in \mathbb{N}$, so that Corollary 8.2.4 applies. We conclude by noticing that for any $k \in \mathbb{N}$,

$$\begin{aligned}
\|\,\mathrm{id}_k \otimes \mathcal{L} : \mathbb{L}_2(\sigma) \to \mathbb{L}_2(k^{-1}\mathbb{1}_k \otimes \sigma)\| &= \|\Gamma_{\frac{1_k}{k}\otimes\sigma} \circ (\mathrm{id}_k \otimes \mathcal{L}) \circ \Gamma^{-1}_{\frac{1_k}{k}\otimes\sigma} : \mathcal{T}_2(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})\| \\
&= \|\Gamma_\sigma \circ \mathcal{L} \circ \Gamma^{-1}_\sigma : \mathcal{T}_2(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})\| \qquad (10.18) \\
&= \|\mathcal{L} : \mathbb{L}_2(\sigma) \to \mathbb{L}_2(\sigma)\|,
\end{aligned}$$

where we used Theorem 4 of [Watrous, 2005] in (10.18). The second part follows similarly by the second part of Corollary 8.2.4. In both cases, hypercontractivity follows from Theorem 10.2.2. $\qquad\square$

Moreover, we derive universal bounds on the complete-log-Sobolev constants:

Theorem 10.2.5 (Universal bounds on the complete logarithmic Sobolev constants)**.** *Let $(\mathcal{P}_t)_{t\geq 0}$ be a primitive reversible QMS, with unique invariant state σ and spectral gap $\lambda(\mathcal{L})$. Then, $\mathrm{cLSI}_2(c,\ln\sqrt{2})$ holds, with*

$$c \leq \frac{\ln\|\sigma^{-1}\|_\infty + 2}{\lambda(\mathcal{L})}. \qquad (10.19)$$

Proof. First notice that for all $k \in \mathbb{N}$, and any $X \in \mathcal{B}(\mathbb{C}^k \otimes \mathcal{H})$,

$$\|\,\mathrm{id}_k \otimes \mathcal{P}_t(X)\|_{\mathbb{L}_2(\mathcal{N}_k,\,\mathbb{L}_4(\sigma_{\mathrm{Tr}}))} \leq \|X\|_{\mathbb{L}_2(\mathcal{N}_k,\,\mathbb{L}_4(\sigma_{\mathrm{Tr}}))} \leq \|\,\mathrm{id}:\mathbb{L}_2(\sigma) \to \mathbb{L}_4(\sigma)\|_{\mathrm{cb}}\,\|X\|_{\mathbb{L}_2(k^{-1}\mathbb{1}_k\otimes\sigma)} \qquad (10.20)$$

where the first inequality follows from (i) of Proposition 8.3.1. Then, for any $1 \leq q \leq p \leq \infty$:

$$\| \operatorname{id} : \mathbb{L}_q(\sigma) \to \mathbb{L}_p(\sigma) \|_{\mathrm{cb}} = \| \Gamma_\sigma^{\frac{1}{p}-\frac{1}{q}} : \mathcal{T}_q(\mathcal{H}) \to \mathcal{T}_p(\mathcal{H}) \|_{\mathrm{cb}} = \| \sigma^{-1} \|_\infty^{\frac{1}{q}-\frac{1}{p}} . \tag{10.21}$$

Now, an application of Theorem 8.4.10 to the QMS $(\operatorname{id}_k \otimes \mathcal{P}_t)_{t \geq 0}$ together with the fact that $\lambda(\mathcal{L}) = \lambda(\operatorname{id}_k \otimes \mathcal{L})$ for any $k \in \mathbb{N}$ allow us to conclude. $\qquad\square$

Using the multiplicativity of CB norms [Devetak et al., 2006], we directly get the tensorization property of the complete logarithmic Sobolev inequality, hence extending Theorem 6 of [Beigi and King, 2016] to any primitive QMS.

Theorem 10.2.6. *Suppose that for all $i = 1, \cdots, N$ the primitive QMS $(\mathcal{P}^{(i)})_{t \geq 0}$ on $\mathcal{B}(\mathcal{H}_i)$ generated by $\mathcal{L}_i$ with invariant state σ_i satisfies $\mathrm{cLSI}_q(c_i, d_i)$. Then the QMS $(\mathcal{P}_t)_{t \geq 0}$ on $\mathcal{B}(\otimes_{i=1}^N \mathcal{H}_i)$ generated by $\mathcal{L}^{(N)} := \sum_{i=1}^N \otimes_{k=1}^{i-1} \operatorname{id}_{\mathcal{B}(\mathcal{H}_k)} \otimes \mathcal{L}_i \otimes \otimes_{k=i+1}^N \operatorname{id}_{\mathcal{B}(\mathcal{H}_k)}$ with invariant state $\otimes_{i=1}^N \sigma_i$ satisfies $\mathrm{cLSI}_q(c, d)$ with $c := \max_i c_i$ and $d = \sum_{i=1}^N d_i$.*

The additivity of the weak complete logarithmic Sobolev constant prevents one from obtaining relevant estimates for a large number of tensorized primitive QMS. In particular, estimating both constants separately as in Theorem 10.2.6 leads to weaker bounds than the ones found in [Temme et al., 2014, Müller-Hermes et al., 2016]. This however does not exclude the possibility of better controlling both constants simultaneously when considering tensor products of QMS. One way to achieve this is via the transference method introduced in Chapter 9. Before moving to the next section, we simply mention that a similar notion of a complete modified logarithmic Sobolev inequality was introduced in [Bardet, 2017] (see also [Gao et al., 2018b]). Contrary to the 2-complete logarithmic Sobolev inequality studied here, there exist examples of semigroups for which the "weak MLSI constant" vanishes. The question of the validity of this claim for any quantum Markov semigroup is still open. Another interesting problem is the one of finding the range of parameters q for which the weak complete logarithmic Sobolev constant becomes non-zero.

Chapter 11.

Quantum geometric and information theoretic inequalities

In Section 4.5, we recalled an information theoretical proof of the logarithmic Sobolev inequality for the Ornstein Uhlenbeck semigroup. At the core of the proof lies the entropy power inequality (cf. Theorem 4.5.2): for any two independent random variables X and Y on $\mathbb{R}^n$,

$$N(X + Y) \geq N(X) + N(Y), \qquad \text{where} \qquad N(X) := \mathrm{e}^{2S(X)/n} \qquad (11.1)$$

is the entropy power of signal X. Equivalently, the following entropy convex combination inequality holds: $S(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) \geq \lambda S(X) + (1-\lambda)S(Y)$. In words, the latter inequality expresses the fact that the entropy of the sum of two signals is greater than the sum of the entropies of the signals taken individually. The entropy power inequality is also closely related to the idea of isoperimetry and has important applications in channel coding theory [Shannon, 1948].

Recently, the entropy power inequality (EPI) has been a subject of focus in the quantum community [Koenig and Smith, 2014, Audenaert et al., 2016, Carlen et al., 2016]. In this chapter, we aim at extending the results mentioned in Section 4.5 to the quantum phase space setting. In this chapter, we consider a classical-quantum entropy power inequality, where the sum between the two independent random variables X and Y in Theorem 4.5.2 is replaced by the classical-quantum convolution between a quantum state ρ and a classical probability distribution function f as introduced by [Werner, 1984]. We show how this inequality provides convergence times for the quantum heat semigroup defined in Section 5.5.2. Some of the results stated and proved in this section were found independently in [Huber et al., 2017, Datta et al., 2017].

However, a generalization of the proof of the logarithmic Sobolev inequality via EPI for the quantum Ornstein Uhlenbeck semigroup (see Section 5.5.2) seems to require a fully quantum-quantum convolution as introduced in [Koenig and Smith, 2014][1], which is a channel that models the physical process corresponding to photons coming out of a beamsplitter. The proof was recently found in [De Palma and Trevisan, 2018] and relies on a series of papers [G. De Palma et al., 2014, De Palma et al., 2015b, Koenig, 2015, De Palma and Trevisan, 2018]. We summarize and compare their main results to the ones of Section 11.1 in Section 11.3.

---

[1]We also mention that a finite dimensional version of the quantum-quantum convolution was later introduced and a quantum-quantum entropy power inequality for that convolution was derived in [Audenaert et al., 2016], see also [Carlen et al., 2016] for a more recent proof and [Jeong et al., 2018] for a generalization in the presence of quantum memory.

Layout of the chapter: In Section 11.1, we state and prove some classical-quantum versions of the information theoretic inequalities of Section 4.5. These results, and in particular the isoperimetric inequality of Theorem 11.1.7 are used to prove some convergence bounds on the quantum heat semigroup in Section 11.2. Connection with the fully quantum entropy power inequality of [Koenig and Smith, 2014] is established, and the link to the entropy decay for the quantum Ornstein Uhlenbeck semigroup found in [De Palma and Huber, 2018] is recalled in Section 11.3.

11.1. A quantum-classical entropy power inequality

Given a random variable Z. its entropy power $N(Z)$ only depends on its probability density function f_Z. Therefore, (11.1) only depends on f_X, f_Y and their convolution

$$f_{X+Y}(z) = f_X * f_Y(z) = \int_{\mathbb{R}^n} f_X(x) f_Y(z-x) dx.$$

Any quantum extension of the (11.1) requires a non-commutative notion of convolution. In [Werner, 1984], the classical-quantum convolution between a quantum state $\rho \in \mathcal{D}(\mathcal{H})$, $\mathcal{H} = \mathbb{L}_2(\mathbb{R}^n)$, and a probability density function f on $\mathbb{R}^{2n}$ was defined as follows:

$$\rho * f := \int_{\mathbb{R}^{2n}} f(z) \, W_z \, \rho \, W_{-z} \, dz,$$

where the integral is defined in the Bochner sense. Here, the Weyl displacement operators, introduced in Section 0.2, play the role of translations in phase space. Next, we define the *quantum entropy power* of a state $\rho \in \mathcal{D}(\mathcal{H})$ as follows

$$N(\rho) := e^{S(\rho)/n} . \tag{11.2}$$

Once again, the difference in the normalization of the exponent in comparison with the classical entropy power (11.1) comes from the fact that the quantum phase space is $2n$ dimensional. The following classical-quantum entropy power inequality was first derived in [Huber et al., 2017]:

Theorem 11.1.1. *For any $\rho \in \mathcal{D}(\mathcal{H})$ and any probability density function f on $\mathbb{R}^{2n}$,*

$$N(\rho * f) \geq N(\rho) + N(f) . \tag{cq-EPI}$$

In this section, we provide a proof of Theorem 11.1.1 using tools developed independently in [Datta et al., 2017]. Along the way, we derive other entropic inequalities on the quantum phase space such as a quantum Blachman-Stam inequality, the concavity of the entropy power along the quantum heat semigroup, as well as an entropic isoperimetric inequality, that may be of independent interest.

11.1.1. Fisher informations on phase space

Divergence based quantum Fisher information Classically, the proof of the entropy power inequality relies on the Fisher information inequality of Lemma 4.5.4. The crucial step is the so-called de Bruijn identity (Lemma 1.3.2) relating the Fisher informations to the derivative of the Shannon entropy along the heat semigroup, which allows to interpret it as its corresponding entropy production. In the quantum setting, one is then tempted to define a quantum Fisher information that can similarly be related to the von Neumann entropy.

We first briefly recall the classical setting: given a smooth family $(f_\theta)_{\theta \in \Theta}$ of probability density functions on $\mathbb{R}^{2n}$, its classical Fisher information is defined by Equation (1.70) as the negative of the Laplacian of the function $(\alpha, \beta) \mapsto D(f_\alpha \| f_\beta)$, or equivalently:

$$I(\theta) = \Delta_\alpha D(f_\theta \| f_\alpha)|_{\alpha = \theta}. \tag{11.3}$$

In the case when the parameter θ simply consists of shifts from a fixed distribution f, the expression can be simplified as in Equation (1.71):

$$I(f) := I(0) = - \int f(x) \, \Delta(\ln f)(x) \, dx. \tag{11.4}$$

Up to regularity issues, this is nothing but the entropy production associated to the heat semigroup (cf. Section 3.3).

In the quantum case, we saw in Section 1.3 that, given a smooth family $(\rho_\theta)_{\theta \in \Theta}$ of quantum states on a finite dimensional Hilbert space, their Fisher information was defined, similarly, as the Laplacian of $(\alpha, \beta) \mapsto D(\rho_\alpha \| \rho_\beta)$. In analogy with (11.4), we are interested in the case when $(\rho_\theta)_{\theta \in \mathbb{R}^{2n}}$ can be interpreted as a family of shifts in the quantum phase space. This can naturally be done via the Weyl operators introduced in Section 0.2.

We assume from now on that $\mathcal{H} = \mathbb{L}_2(\mathbb{R}^n)$. Then, given a state $\rho \in \mathcal{D}(\mathcal{H})$, define the shifted operator along the direction of R_j as

$$\rho_{\theta, R_j} := e^{i\theta R_j} \rho \, e^{-i\theta R_j} := W_{z_{j,\theta}} \rho W_{-z_{j,\theta}}, \tag{11.5}$$

where $z_{j,\theta}$ is the vector in $\mathbb{R}^{2n}$ with entries

$$(z_{j,\theta})_r := \begin{cases} \theta & \text{if } r = j, \; j \text{ even}, \\ -\theta & \text{if } r = j, \; j \text{ odd}, \\ 0 & \text{otherwise}. \end{cases} \tag{11.6}$$

Assuming that $\rho \in \mathcal{D}_+(\mathcal{H})$, its *divergence-based quantum Fisher information* is defined in analogy with (11.3) as [Koenig and Smith, 2014, Koenig, 2015]:

$$J(\rho) := \sum_{i=1}^{2n} \frac{d^2}{d\theta^2} D(\rho \| \rho_{\theta, R_j}) \bigg|_{\theta = 0}, \tag{11.7}$$

whenever the functions $\theta \mapsto D(\rho \| \rho_{\theta, R_j})$ are twice differentiable. Formally [Koenig and Smith, 2014], this quantity can be shown to be equal to the entropy production of the quantum heat semigroup (cf. Equation (5.44)):

$$\frac{d}{dt} S(\mathcal{P}_{*t}^{\text{heat}}(\rho)) \bigg|_{t=0} = - \operatorname{Tr}(\mathcal{L}_*^{\text{heat}}(\rho) \ln \rho) = \frac{1}{4} J(\rho). \tag{11.8}$$

This statement was rigorously proved for Gaussian thermal states in [Koenig, 2015]. The second identity can be made more precise when ρ is a Schwartz operator (cf. Section 0.2):

Lemma 11.1.2. *Let ρ be a Schwartz state such that $\ln \rho$ is polynomially bounded. Then for any $j \in \{1, 2, \ldots, 2n\}$, the function $\theta \mapsto D(\rho \| \rho_{\theta, R_j})$ is twice differentiable at 0. Moreover, the operators $\rho [R_j, [R_j, \ln \rho]]$ are trace class, and the divergence-based quantum Fisher information defined through*

Equation (11.7) *is given by*

$$J(\rho) = \sum_{i=1}^{2n} \mathrm{Tr}\left(\rho\left[R_j, [R_j, \ln\rho]\right]\right). \tag{11.9}$$

In addition, $\frac{d}{d\theta} D(\rho\|\rho_{\theta, R_j})\big|_{\theta=0} = 0$ *for all* j.

Proof. We start from the following relative entropy $D(\rho\|\rho_{\theta, R_j}) = \mathrm{Tr}(\rho\left(\ln\rho - \ln\rho_{\theta, R_j}\right))$. Since its first term is constant, we can focus on the second term:

$$\begin{aligned}
\mathrm{Tr}\,\rho\ln\rho_{\theta, R_j} &= \mathrm{Tr}(\rho\,\mathrm{e}^{i\theta R_j}\ln\rho\,\mathrm{e}^{-i\theta R_j}) \\
&= \sum_{k=1}^{\infty}\langle\psi_k, \mathrm{e}^{-i\theta R_j}\rho\,\mathrm{e}^{i\theta R_j}\ln\rho\,\psi_k\rangle,
\end{aligned} \tag{11.10}$$

where $\{\psi_k\}_{k=1}^{\infty}$ is an orthonormal basis of Schwartz functions of $\mathcal{H}$ (take e.g. the Hermite polynomials). Recall that $\mathrm{e}^{i\theta R_j}$ is equal to $W_{z_{j,\theta}}$, with $z_{j,\theta}$ defined in (11.6). Let's denote $\varphi_k := \ln\rho\,\psi_k$. It is a Schwartz function, since $\ln\rho$ is polynomially bounded (see Proposition 3.21 of [Keyl et al., 2016]). Therefore, each term of Equation (11.10) is of the form

$$\langle\psi_k, W_{-z_{j,\theta}}\rho\,W_{z_{j,\theta}}\varphi_k\rangle, \qquad \psi_k, \varphi_k \in \mathcal{S}(\mathcal{H}). \tag{11.11}$$

We conclude by a direct use of Proposition 3.25 of [Keyl et al., 2016]. $\qquad\square$

The divergence-based Fisher information satisfies the following quantum version of the Blachman-Stam inequality. The result was independently found by [Huber et al., 2017] and [Datta et al., 2017].

Theorem 11.1.3 (Quantum Blachman-Stam inequality). *For any* $\alpha, \beta > 0$, $t > 0$, *for any state* ρ *such that* $\theta \mapsto D(\rho\|\rho_{\theta, R_j})$ *is twice differentiable at 0 for all* $j = 1, ..., 2n$, *we have*

$$(\alpha + \beta)^2 J(\rho * f) \leq \alpha^2 J(\rho) + \beta^2\,\mathrm{I}(f). \tag{11.12}$$

In order to prove the quantum Blachman-Stam inequality, we need the following technical lemma:

Lemma 11.1.4. *For any two quantum states* $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$ *and probability density functions* $f_1, f_2 \in \mathbb{L}_1(\mathbb{R}^{2n})$,

$$D(\rho_1 * f_1\|\rho_2 * f_2) \leq D(\rho_1\|\rho_2) + D(f_1\|f_2) \tag{11.13}$$

Proof. The proof involves a simple application of Uhlmann's monotonicity of the relative entropy (1.56) under Schwarz mapping in the context of general von Neumann algebras. Indeed, the result would follow if we can prove that for any $\rho \in \mathcal{D}(\mathcal{H})$ and any positive $g \in \mathbb{L}_1(\mathbb{R}^{2n})$, with $\|g\|_{\mathbb{L}^1(\mathbb{R}^{2n})} = 1$,

$$\omega_{\rho * g} = (\omega_\rho \otimes \omega_g) \circ \Phi, \tag{11.14}$$

where ω_ρ, resp. ω_f, is the state corresponding to the density operator ρ, resp. the density function f, and Φ is a unital, completely positive map from $\mathcal{B}(\mathcal{H}) \otimes \mathbb{L}_\infty(\mathbb{R}^{2n})$ to $\mathcal{B}(\mathcal{H})$. Indeed, $\omega_\rho \otimes \omega_g$ acts on tensor product elements of $\mathcal{B}(\mathcal{H}) \otimes \mathbb{L}_\infty(\mathbb{R}^{2n})$ by

$$(\omega_\rho \otimes \omega_g)(A \otimes f) = \mathrm{Tr}(\rho A) \times \int_{\mathbb{R}^{2n}} g(z)f(z)dz, \quad A \in \mathcal{B}(\mathcal{H}), f \in \mathbb{L}_\infty(\mathbb{R}^{2n}) \tag{11.15}$$

and extends by continuity [Kadison and Ringrose, 1983]. But $\mathcal{B}(\mathcal{H}) \otimes \mathbb{L}_\infty(\mathbb{R}^{2n})$ can be identified with the set $\mathbb{L}_\infty(\mathbb{R}^{2n}, \mathcal{B}(\mathcal{H}))$ of measurable functions from $\mathbb{R}^{2n}$ to $\mathcal{B}(\mathcal{H})$, in which case an element of it is a map $z \mapsto A_z$ and $\omega_\rho \otimes \omega_g$ acts as

$$(\omega_\rho \otimes \omega_g)(z \mapsto A_z) = \int_{\mathbb{R}^{2n}} \mathrm{Tr}(\rho A_z) \, g(z) \, \mathrm{d}z \qquad (11.16)$$

(which of course is consistent with the above if $A_z = f(z)A$). On the other hand, $\omega_{\rho \ast g}$ acts as

$$\omega_{\rho \ast g}(A) = \int_{\mathbb{R}^{2n}} g(z) \, \mathrm{Tr}(\rho \, W_{-z} A W_z) \, \mathrm{d}z, \qquad (11.17)$$

so to prove our statement it is enough to show that the map Φ from $\mathcal{B}(\mathcal{H})$ to $\mathbb{L}_\infty(\mathbb{R}^{2n}, \mathcal{B}(\mathcal{H}))$ defined as

$$\Phi : A \mapsto (z \mapsto W_{-z} A W_z) \qquad (11.18)$$

is unital and completely positive. The first requirement is obvious. To show complete positivity, consider a positive matrix $(A_{i,j})_{i,j=1}^N$ of elements of $\mathcal{B}(\mathcal{H})$, in the sense that for any $(\varphi_i)_{i=1}^N$ with each $\varphi_i \in \mathcal{H}$, one has

$$\sum_{i,j} \langle \varphi_i, A_{i,j} \varphi_j \rangle \geq 0.$$

We then prove that $(\Phi(A_{i,j}))_{i,j=1}^N$ is positive when acting on $\mathbb{L}_2(\mathbb{R}^{2n}, \mathcal{H})$: consider $\psi_i = (z \mapsto \psi_i(z)) \in \mathbb{L}_2(\mathbb{R}^{2n}, \mathcal{H})$, therefore

$$\sum_{i,j} \langle \psi_i, \Phi(A_{i,j}) \psi_j \rangle = \sum_{i,j} \int \langle \psi_i(z), W_z^* A_{i,j} W_z \, \psi_j(z) \rangle \, \mathrm{d}z = \int \sum_{i,j} \langle W_z \psi_i(z), A_{i,j} W_z \psi_j(z) \rangle \, \mathrm{d}z \geq 0.$$

This together with the additivity (1.54) of the relative entropy completes the proof of (11.13). $\qquad \square$

We are now ready to prove Theorem 11.1.3. The proof we give can be seen as a quantum analogue of the proof given by Stam in [a.J. Stam, 1959].

Proof of Theorem 11.1.3 : In the following, we define a vector $z = (q_1, p_1, \ldots, q_n, p_n) \in \mathcal{Z}$ simply as (q, p), and for any $a \in \mathbb{R}$ we define

$$(q, p; q_j - a) := (q_1, p_1, \ldots, q_{j-1}, p_{j-1}, q_j - a, p_j, q_{j+1} p_{j+1} \ldots, q_n, p_n), \qquad (11.19)$$

$$(q, p; p_j - a) := (q_1, p_1, \ldots, q_{j-1}, p_{j-1}, q_j, p_j - a, q_{j+1} p_{j+1} \ldots, q_n, p_n). \qquad (11.20)$$

Then for any $a \in \mathbb{R}$ we define the following functions:

$$f_{a,P_j} : (q, p) \mapsto f(q, p; q_j - a); \quad f_{a,Q_j} : (q, p) \mapsto f(q, p; p_j + a) \qquad (11.21)$$

Further, accordingly denoting the Weyl operator W_z as $W_{(q,p)}$, we have for any $\theta \in \mathbb{R}$ and $\alpha, \beta > 0$,

$$\rho_{\theta\alpha, P_j} \ast f_{\theta\beta, P_j} = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(q, p; q_j - \theta\beta) \, W_{(q,p)} \, \rho_{\theta\alpha, P_j} \, W_{(-q,-p)} \, dq \, dp$$

$$= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(q, p) \, W_{(q,p;q_j + \theta\beta)} \, \rho_{\theta\alpha, P_j} \, W^*_{(q,p;q_j + \theta\beta)} \, dq \, dp, \qquad (11.22)$$

where we made a change of variable $q_j - \theta\beta \to q_j$. Then using the Weyl-Segal CCR, we obtain

$$W_{(q,p;q_j + \theta\beta)} \, e^{i\theta\alpha P_j} = e^{ip_j \theta\alpha/2} \, W_{(q,p;q_j + \theta(\alpha+\beta))} \qquad (11.23)$$

Hence,

$$
\begin{aligned}
\rho_{\theta\alpha,P_j} * f_{\theta\beta,P_j} &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(q,p) \, W_{(q,p;q_j+\theta(\alpha+\beta))} \, \rho \, W_{-(q,p;q_j+\theta(\alpha+\beta))} \, dq \, dp \\
&= e^{i\theta(\alpha+\beta)P_j} \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(q,p) \, W_{(q,p)} \, \rho \, W_{-(q,p)} \, dq \, dp \right) e^{-i\theta(\alpha+\beta)P_j} \\
&= \rho_{\theta(\alpha+\beta),P_j} \, .
\end{aligned}
\tag{11.24}
$$

This implies that

$$
D\left(\rho \| \rho_{\theta\alpha,P_j} * f_{\theta\beta,P_j}\right) = D\left(\rho \| \rho_{\theta(\alpha+\beta),P_j}\right).
\tag{11.25}
$$

Similarly, we can show that

$$
D\left(\rho \| \rho_{\theta\alpha,Q_j} * f_{\theta\beta,Q_j}\right) = D\left(\rho \| \rho_{\theta(\alpha+\beta),Q_j}\right).
\tag{11.26}
$$

A direct application of Lemma 11.1.4 leads to:

$$
D\left(\rho \| \rho_{\theta\alpha,R_j} * f_{\theta\beta,R_j}\right) \le D\left(\rho \| \rho_{\theta\alpha,R_j}\right) + D\left(f \| f_{\theta\beta,R_j}\right),
\tag{11.27}
$$

The result then follows by differentiating the above inequality twice, since both sides of the above are functions of θ that are equal up to first order. $\qquad\square$

The quantum integral Fisher information The proofs of the entropic inequalities derived in [Koenig and Smith, 2014, Huber et al., 2017, Datta et al., 2017] all rely on the assumption that the state ρ satisfies the de Bruijn identity (11.8). Classically, the de Bruijn identity holds with mild assumptions on the distribution f (cf. Lemma 1.3.2). However, it is not clear which are the minimal conditions that the state ρ should satisfy for the identity to hold. In [De Palma and Trevisan, 2018], the authors took a different approach by showing that a particular entropic quantity is differentiable and defining the Fisher information as its derivative. In this approach, the de Bruijn identity becomes true by definition. Here, we review the approach of [De Palma and Trevisan, 2018] and relate it to the one we described in the previous paragraph. In the next section, we use this better behaved definition for the quantum Fisher to prove the entropic inequalities already appearing in [Koenig and Smith, 2014, Huber et al., 2017, Datta et al., 2017] without requiring any regularity on the quantum state other than finiteness of its entropy and of its first moments.

Given a quantum state $\rho \in \mathcal{D}(\mathcal{H})$ of finite entropy, its *quantum integral Fisher information* $\tilde{J}(\rho)$ is defined as follows: for $t \ge 0$[2]:

$$
\tilde{J}(\rho)(t) := \mathrm{Ent}\left(\omega_{\mathrm{QC}}(t) \| \omega_{\mathcal{P}^{\mathrm{Heat}}_{*t}(\rho)} \otimes \omega_{g_{t/2}}\right),
$$

where $\omega_{\mathrm{QC}}(t)$ is a state defined on the algebra $\mathcal{B}(\mathcal{H}) \otimes \mathbb{L}_\infty(\mathbb{R}^{2n})$ whose action on tensor product elements $X \otimes f$, $X \in \mathcal{B}(\mathcal{H})$, $f \in \mathbb{L}_\infty(\mathbb{R}^{2n})$, is as follows,:

$$
\omega_{\mathrm{QC}}(t)(X \otimes f) := \int_{\mathbb{R}^{2n}} \mathrm{Tr}(W_z \, \rho \, W_{-z} \, X) \, f(z) \, g_{t/2}(z) \, dz \equiv \omega_\rho \otimes \omega_{g_{t/2}} \circ \Phi(X \otimes f).
\tag{11.28}
$$

In the case when ρ has finite second moments and entropy, one shows that $t \mapsto \tilde{J}(\rho)(t)$ is an increasing, concave continuous function of time (cf. Theorem 2 of [De Palma and Trevisan, 2018]). This provides

---

[2]In fact, [De Palma and Trevisan, 2018] more generally defined a conditional quantum integral Fisher information $\tilde{J}$ that can be expressed as a conditional mutual information.

the existence of the following limit:

$$J(\rho) := 4 \lim_{t \to 0^+} \frac{\tilde{J}(\rho)(t)}{t} \,. \tag{11.29}$$

This new definition of the Fisher information has the advantage that it directly implies the validity of the de Bruijn identity. Indeed, $\tilde{J}(\rho)$ is a mutual information which can be written as [Belavkin and Ohya, 2002]:

$$\tilde{J}(\rho)(t) = S(\mathcal{P}^{\mathrm{Heat}}_{*t}(\rho)) + S(g_{t/2}) - S(\omega_{\mathrm{QC}}(t)) \,,$$

where the entropy $S(\omega_{\mathrm{QC}}(t))$ is defined as [Umegaki, 1962]

$$S(\omega_{\mathrm{QC}}(t)) = -\tau \left(\frac{d\omega_{\mathrm{QC}}(t)}{d\tau} \ln \frac{d\omega_{\mathrm{QC}}(t)}{d\tau} \right),$$

where τ is the trace defined on tensor elements of the semi-finite von Neumann algebra $\mathcal{B}(\mathcal{H}) \bar{\otimes} \mathbb{L}_\infty(\mathbb{R}^{2n})$ as follows: for $X \in \mathcal{T}_1(\mathcal{H})$ and $f \in \mathbb{L}_1(\mathcal{H})$, $\tau(X \otimes f) = \mathrm{Tr}(X) \int_{\mathbb{R}^{2n}} f(z) dz$. The operator $\frac{d\omega_{\mathrm{QC}}(t)}{d\tau}$ is the density operator associated to the state $\omega_{\mathrm{QC}}(t)$. From Equation (11.28), it can be identified as $z \mapsto W_z \rho W_{-z} \, g_{t/2}(z)$. Therefore [De Palma and Trevisan, 2018],

$$\begin{aligned}
\tilde{J}(\rho)(t) &= S(\mathcal{P}^{\mathrm{Heat}}_{*t}(\rho)) + S(g_{t/2}) - S(\omega_{\mathrm{QC}}(t)) \\
&= S(\mathcal{P}^{\mathrm{Heat}}_{*t}(\rho)) + S(g_{t/2}) + \int_{\mathbb{R}^{2n}} \mathrm{Tr}\left(W_z \rho W_{-z} \ln(g_{t/2}(z) W_z \rho W_{-z})\right) g_{t/2}(z) \, dz \\
&= S(\mathcal{P}^{\mathrm{Heat}}_{*t}(\rho)) + S(g_{t/2}) - S(g_{t/2}) + \int_{\mathbb{R}^{2n}} S(W_z \rho W_{-z}) g_{t/2}(z) \, dz \\
&= S(\mathcal{P}^{\mathrm{Heat}}_{*t}(\rho)) - S(\rho) \,.
\end{aligned}$$

Even if formally, the quantities defined in (11.29) and (11.7) coincide (cf. Equation (11.7)), the problem of making this intuition rigorous is still open. Fortunately, the new definition (11.8) can still be shown to satisfy Equation (11.12) for $0 \le \alpha = 1 - \beta \le 1$ as long as ρ and f both have finite second moments. The proof of this shares some similarities with the one of Theorem 11.1.3 and can be found in Theorem 3 of [De Palma and Trevisan, 2018]. In the next subsection, we assume this fact and give a brief summary of the information theoretic inequalities that it implies.

11.1.2. Information theoretic inequalities on the quantum phase space

The classical-quantum entropy power inequality is now a simple consequence of the Blachman-Stam inequality and the de Bruijn identity. It was originally derived in [Huber et al., 2017] using the divergence based Fisher information (11.7). A similar proof which uses the Fisher information defined in Equation (11.29), which also generalizes to the case of conditional entropies, is provided in Theorem 5 of [De Palma and Huber, 2018]:

Theorem 11.1.5 (Classical-quantum EPI). *For any state $\rho \in \mathcal{D}(\mathcal{H})$, any density function $f \in \mathbb{L}_1(\mathbb{R}^{2n})$, $\int f = 1$, of finite entropy and first and second moments, and any $0 \le \lambda \le 1$, the following holds:*

$$\frac{S(\rho * f)}{n} \ge \lambda \frac{S(\rho)}{n} + (1 - \lambda) \frac{S(f)}{n} - \lambda \ln \lambda - (1 - \lambda) \ln(1 - \lambda) \,.$$

After optimization over λ this inequality implies

$$N(\rho * f) \ge N(\rho) + N(f) \,. \tag{CQ-EPI}$$

As expected, the Blachman-Stam inequality also provides quantum generalizations of the concavity of the entropy power, which itself implies a quantum entropic isoperimetric inequality. Our proof of Theorem 11.1.6 can hence be interpreted as a quantum version of the proof of [Dembo, 1989]:

Theorem 11.1.6 (Concavity of the quantum entropy power)**.** *For any state $\rho \in \mathcal{D}(\mathcal{H})$ of finite second moments and finite Fisher information, the entropy power of $\rho_t := \mathcal{P}_{*t}^{\mathrm{qHeat}}(\rho)$ is twice differentiable and satisfies*

$$\frac{d^2}{dt^2}\bigg|_{t=0^+} N(\rho_t) \le 0 \,. \tag{11.30}$$

Proof. Since $J(\rho) < \infty$ by assumption, and since $t \mapsto \tilde{J}(\rho_t)$ is concave, $t \mapsto J(\rho_t)$ is a bounded, decreasing function. Letting $\varepsilon > 0$ and setting $\alpha = \frac{1}{J(\rho_t)}$ and $\beta = \frac{\varepsilon}{4n}$ in (11.12), we get

$$\left(\frac{1}{J(\rho_t)} + \frac{\varepsilon}{4n}\right)^2 J(\rho_{t+\varepsilon}) \le \frac{1}{J(\rho_t)} + \frac{\varepsilon}{4n} \quad \Rightarrow \quad \frac{J(\rho_{t+\varepsilon}) - J(\rho_t)}{\varepsilon} \le \frac{-J(\rho_t)J(\rho_{t+\varepsilon})}{4n} \,. \tag{11.31}$$

Since $t \mapsto J(\rho_t)$ is decreasing, both sides of Equation (11.31) converge as $\varepsilon \to 0^+$, so that $t \mapsto J(\rho_t)$ is differentiable at 0 and:

$$\frac{d}{dt}\bigg|_{t=0^+} J(\rho_t) \le -\frac{J(\rho_t)^2}{4n} \,.$$

It is then an easy calculation to verify that the above inequality together with the definition of J implies (11.30). $\qquad\square$

Theorem 11.1.6 has a simple corollary that can be interpreted as a quantum entropic isoperimetric inequality $e-$Isop. The proof that we provide here only relies on the concavity and asymptotic properties of the entropy power and asymptotics. In contrast, the proof of [De Palma and Trevisan, 2018] makes use of the entropy power inequality.

Theorem 11.1.7 (Isoperimetric inequality for the quantum entropy)**.** *For any state ρ of finite second moments and Fisher information:*

$$J(\rho)\, N(\rho) \ge 2\,\mathrm{e}\,n \,. \tag{q e -- Isop}$$

Proof. From Equation (11.8), for any such ρ evolving under the action of the quantum heat semigroup,

$$\frac{d}{ds} N(\rho_s) = \frac{1}{4n} J(\rho_s)\, N(\rho_s). \tag{11.32}$$

Moreover by the concavity of the entropy power (Theorem 11.1.6),

$$\frac{d}{ds} N(\rho_s)\bigg|_{s=0} \ge \frac{N(\rho_t) - N(\rho)}{t} \,, \qquad \forall\, t > 0 \,. \tag{11.33}$$

However, by Corollary III.4 of [Koenig and Smith, 2014] (see also Theorem 5 of [De Palma and Trevisan, 2018]), $N(\rho_t) = \frac{\mathrm{e}\,t}{2} + \mathcal{O}(1)$. Combining these facts, we get the following inequality:

$$J(\rho) \ge \mathrm{e}^{-\frac{1}{n}S(\rho)}\, 2\,\mathrm{e}\,n, \tag{11.34}$$

which is the statement of Theorem 11.1.7. $\qquad\square$

11.2. Convergence properties of the quantum heat semigroup

In this section, we use the quantum isoperimetric inequality (Theorem 11.1.7) to find a bound on the relative entropy $D(\rho_t \| \rho_t^G)$ for any $t > 0$, where $\rho_t = \mathcal{P}_{*t}^{\mathrm{qHeat}}(\rho)$ and for any state ρ of finite first and second moments, ρ_t^G denotes its 'Gaussification', i.e. the Gaussian state of same mean and covariance as ρ_t. We begin by recalling the following result of [Koenig and Smith, 2014].

Lemma 11.2.1. *For any state $\rho \in \mathcal{D}(\mathcal{H})$ of finite first and second moments, finite entropy and with covariance matrix satisfying the conditions of Theorem 0.2.4, denote $\rho_t := \mathcal{P}_{*t}^{\mathrm{qHeat}}(\rho)$. Then*

$$D(\rho_t \| \rho_t^G) = -S(\rho_t) + S(\rho_t^G). \tag{11.35}$$

Proof. Without loss of generality we can reduce the proof to the case of a centered state ρ. Then:

$$-S(\rho_t) + S(\rho_t^G) = \mathrm{Tr}(\rho_t \ln \rho_t) - \mathrm{Tr}(\rho_t^G \ln \rho_t^G) \tag{11.36}$$

$$= \mathrm{Tr}(\rho_t \ln \rho_t) - \ln C + \mathrm{Tr}(\rho_t^G \mathbf{R}^T \Gamma \mathbf{R}) \tag{11.37}$$

$$= \mathrm{Tr}(\rho_t \ln \rho_t) - \ln C + \mathrm{Tr}(\rho_t \mathbf{R}^T \Gamma \mathbf{R}) \tag{11.38}$$

$$= \mathrm{Tr}(\rho_t \ln \rho_t) - \mathrm{Tr}(\rho_t \ln \rho_t^G) \tag{11.39}$$

$$= D(\rho_t \| \rho_t^G), \tag{11.40}$$

where in the second line we used the form (0.41) for ρ_t^G, and in the third line we used the fact that ρ_t has same covariance matrix as ρ_t^G for each time t. $\qquad \square$

This result enables us to prove that $D(\rho_t \| \rho_t^G)$ converges to 0 as $t \to \infty$, as pointed out By König and Smith in [Koenig and Smith, 2014]. Indeed for any initial state ρ evolving according to the heat semigroup $(\mathcal{P}_t^{\mathrm{qHeat}})_{t \geq 0}$ they proved that the large time behavior of $S(\rho_t)$ is independent of ρ (see Corollary III.4). Therefore $-S(\rho_t)$ and $S(\rho_t^G)$ asymptotically cancel each other, and a direct application of Lemma 11.2.1 provides the convergence claimed. Below, we provide a refined analysis of the behavior of $D(\rho_t \| \rho_t^G)$ as t goes to infinity using the isoperimetric inequality derived in Theorem 11.1.7. We assume that $n = 1$ for sake of simplicity.

Lemma 11.2.2. *Let $\mathcal{H} := \mathbb{L}_2(\mathbb{R})$. Then for any initial state $\rho \in \mathcal{D}(\mathcal{H})$ of finite first and second moments evolving according to the heat semigroup $(\mathcal{P}_t^{\mathrm{qHeat}})_{t \geq 0}$, the following inequality holds for any $t \geq 0$:*

$$D(\rho_t \| \rho_t^G) \leq \mathrm{e}^{-\kappa(t)} D(\rho \| \rho^G) + \mathrm{e}^{-\kappa(t)} \int_0^t \frac{\mathrm{e}^{\kappa(s)}}{N(\rho_s^G)} \Big(\frac{1}{4} J(\rho_s^G) N(\rho_s^G) - \frac{\mathrm{e}}{2} \Big) ds, \tag{11.41}$$

where:

$$\kappa(t) := \frac{\mathrm{e}}{2} \int_0^t N(\rho_s^G)^{-1} ds.$$

Proof. We start by differentiating Equation (11.35):

$$
\begin{aligned}
\frac{dD(\rho_t\|\rho_t^G)}{dt} &= -\frac{dS(\rho_t)}{dt} + \frac{dS(\rho_t^G)}{dt} \\
&= -\frac{1}{4}J(\rho_t) + \frac{1}{4}J(\rho_t^G) \\
&\le -\frac{\mathrm{e}}{2}N(\rho_t)^{-1} + \frac{1}{4}J(\rho_t^G) \\
&= -\frac{\mathrm{e}}{2}\mathrm{e}^{-(S(\rho_t)-S(\rho_t^G))}N(\rho_t^G)^{-1} + \frac{1}{4}J(\rho_t^G) \\
&\le -\frac{\mathrm{e}}{2}N(\rho_t^G)^{-1}D(\rho_t\|\rho_t^G) - \frac{\mathrm{e}}{2}N(\rho_t^G)^{-1} + \frac{1}{4}J(\rho_t^G),
\end{aligned}
\tag{11.42}
$$

where we used Theorem 11.1.7 in the third inequality, as well as the basic inequality $\mathrm{e}^x \ge 1 + x$, $x \in \mathbb{R}$ in the fifth one. Now $t \mapsto \mathrm{e}^{\frac{\mathrm{e}}{2}\int_0^t N(\rho_s^G)^{-1}ds} D(\rho_t\|\rho_t^G)$ is differentiable on $[0,\infty)$, and for all $t \ge 0$:

$$
\frac{d}{dt}\left(\mathrm{e}^{\frac{\mathrm{e}}{2}\int_0^t N(\rho_s^G)^{-1}ds} D(\rho_t\|\rho_t^G)\right) = \left(\frac{\mathrm{e}}{2}N(\rho_t^G)^{-1}D(\rho_t\|\rho_t^G) + \frac{dD(\rho_t\|\rho_t^G)}{dt}\right)\mathrm{e}^{\frac{\mathrm{e}}{2}\int_0^t N(\rho_s^G)^{-1}ds} \tag{11.43}
$$

$$
\le \left(-\frac{\mathrm{e}}{2}N(\rho_t^G)^{-1} + \frac{1}{4}J(\rho_t^G)\right)\mathrm{e}^{\frac{\mathrm{e}}{2}\int_0^t N(\rho_s^G)^{-1}ds}, \tag{11.44}
$$

where we used (11.42) in the last line. Integrating the left hand side and the right hand side of the last inequality, we get

$$
D(\rho_t\|\rho_t^G)\,\mathrm{e}^{\frac{\mathrm{e}}{2}\int_0^t N(\rho_s^G)^{-1}ds} - D(\rho\|\rho^G) \le \int_0^t \frac{\mathrm{e}^{\frac{\mathrm{e}}{2}\int_0^s N(\rho_u^G)^{-1}du}}{N(\rho_s^G)}\left(\frac{1}{4}J(\rho_s^G)N(\rho_s^G) - \frac{\mathrm{e}}{2}\right)ds, \tag{11.45}
$$

which when rearranged gives (11.41). $\qquad\square$

The advantage of the upper bound of (11.41) is that it depends on the initial state ρ only through its covariance matrix. In order to interpret the last result as a large time behavior for the quantum heat semigroup, one however still needs to show that this upper bound converges to 0 as t goes to infinity. Fortunately, useful expressions for the Fisher information and the entropy power of a Gaussian state exist, and can be used to prove this claim (see [Koenig and Smith, 2014, Koenig, 2015]).

Lemma 11.2.3. *Let ρ be a one mode Gaussian state of mean photon number $\nu := \mathrm{Tr}(\rho\, a^* a)$. Then*

$$
J(\rho) = 2\ln\frac{\nu+1}{\nu} \tag{11.46}
$$

$$
S(\rho) = (\nu+1)\ln(\nu+1) - \nu\ln\nu. \tag{11.47}
$$

Lemma 11.2.4. *For any initial Gaussian state of mean photon number ν, and for any time $t \ge 0$, the mean photon number of $\mathcal{P}_{*t}^{\mathrm{qHeat}}(\rho)$ is such that*

$$
\nu(t) = \frac{t}{2} + \mathcal{O}(1). \tag{11.48}
$$

Combining the two above Lemmas, we prove the following proposition.

Proposition 11.2.5. *Let $\mathcal{H} := \mathbb{L}_2(\mathbb{R})$. Then for any state ρ of finite first and second moments, finite entropy and covariance matrix satisfying the conditions of Theorem 0.2.4. Then, given $\rho_t := \mathcal{P}_{*t}^{\mathrm{qHeat}}(\rho)$, for any $0 < \varepsilon < 1$ there exists $t_\varepsilon > 0$ as well as $\alpha_\varepsilon > 0$ such that for any $t \ge t_\varepsilon$,*

$$
D(\rho_t\|\rho_t^G) \le \alpha_\varepsilon\, t^{\varepsilon-1}. \tag{11.49}
$$

Proof. We first study the large time behavior of the function κ:

$$\kappa(t) = \frac{e}{2} \int_0^t N(\rho_s^G)^{-1} ds = \frac{e}{2} \int_0^t \frac{\nu(s)^{\nu(s)}}{(\nu(s)+1)^{\nu(s)+1}} ds \tag{11.50}$$

$$= \frac{e}{2} \int_0^t \frac{1}{\nu(s)+1} \left(1 + \frac{1}{\nu(s)}\right)^{-\nu(s)} ds, \tag{11.51}$$

where we used the expressions given by Equation (11.46) for $N(\rho_s^G)$. Using Lemma 11.2.4, we know that

$$\left(1 + \frac{1}{\nu(s)}\right)^{-\nu(s)} \underset{s\to\infty}{\to} e^{-1},$$

which implies that the integrand in Equation (11.51) is equivalent to $2/(e\,s)$. Therefore $\kappa(t)$ diverges as $t \to \infty$. Therefore, the first term in the right hand side of (11.41) converges to 0 as t goes to infinity. We now proceed to a more refined analysis of each term. Let's first define the C^1 function $f : (0,\infty) \to \mathbb{R}$ by $f(x) := (1+x)^{1/x}$. One can easily checks that f and its derivative f' can be continuously extended to 0, with $f(0) = e$ and $f'(0) = -e/2$, so that the following Taylor expansion up to order 1 around 0 holds

$$f(x) = e(1 - x/2) + o(x). \tag{11.52}$$

Now for any $u \geq 0$, $N(\rho_u^G) = (\nu(u)+1)f(1/\nu(u))$, so that

$$N(\rho_u^G)^{-1} = \frac{1}{e\,\nu(u)} + o\left(\frac{1}{\nu(u)}\right). \tag{11.53}$$

Inserting the expression 11.48 for $\nu(u)$, we find that for any $\varepsilon > 0$ there exists $s_0 > 0$ such that for any $u \geq s_0$,

$$\frac{2-2\varepsilon}{e\,u} \leq N(\rho_u^G)^{-1} \leq \frac{2+2\varepsilon}{e\,u}. \tag{11.54}$$

Integrating from 0 to $s > s_0$, we get

$$e^{\kappa(s_0)}(s/s_0)^{(1-\varepsilon)} \leq e^{\kappa(s)} \leq e^{\kappa(s_0)}(s/s_0)^{(1+\varepsilon)}. \tag{11.55}$$

Finally,

$$\begin{aligned}
J(\rho_s^G)\,N(\rho_s^G) &= 2\ln\left(\frac{\nu(s)+1}{\nu(s)}\right)\frac{(\nu(s)+1)^{\nu(s)+1}}{\nu(s)^{\nu(s)}} \\
&= 2\ln\left((1+1/\nu(s))^{\nu(s)+1}\right)(1+1/\nu(s))^{\nu(s)} \\
&= 2(\ln(f(1/\nu(s))) + \ln(1+1/\nu(s)))f(1/\nu(s)) \\
&= 2\,e\,(1 + 1/(2\nu(s)) + o(1/\nu(s)))(1 - \frac{1}{2\nu(s)} + o(1/\nu(s))) \\
&= 2\,e\,(1 + o\,(1/\nu(s))).
\end{aligned}$$

so that

$$J(\rho_s^G)\,N(\rho_s^G) - 2\,e = o\,(1/\nu(s)). \tag{11.56}$$

This, together with the expression (11.48) for $\nu(s)$, implies that for any $\delta > 0$ there exists a $t_0 > 0$ such

that for any $s \geq t_0$,

$$J(\rho_s^G) N(\rho_s^G) - 2\,\mathrm{e} \leq \frac{2\delta + 2}{s} \tag{11.57}$$

Combining (11.57) with (11.55) and (11.54) we end up with the following: for any $t > \max\{t_0, s_0\}$,

$$
\begin{aligned}
\int_0^t \frac{\mathrm{e}^{\kappa(s)}}{N(\rho_s^G)} \big(\frac{1}{4} J(\rho_s^G) N(\rho_s^G) - \frac{\mathrm{e}}{2}\big) ds \leq{} & \int_0^{\max\{s_0,t_0\}} \frac{\mathrm{e}^{\kappa(s)}}{N(\rho_s^G)} \big(\frac{1}{4} J(\rho_s^G) N(\rho_s^G) - \frac{\mathrm{e}}{2}\big) ds \\
& + \mathrm{e}^{\kappa(s_0)} \frac{(1+\varepsilon)(1+\delta)}{\mathrm{e}\, s_0^{1+\varepsilon}} \int_{\max\{s_0,t_0\}}^t s^{\varepsilon-1}\, ds \\
={} & \int_0^{\max\{s_0,t_0\}} \frac{\mathrm{e}^{\kappa(s)}}{N(\rho_s^G)} \big(\frac{1}{4} J(\rho_s^G) N(\rho_s^G) - \frac{\mathrm{e}}{2}\big) ds \\
& + \mathrm{e}^{\kappa(s_0)} \frac{(1+\delta)(1+1/\varepsilon)}{\mathrm{e}\, s_0^{1+\varepsilon}} \big(t^\varepsilon - \max\{s_0,t_0\}^\varepsilon\big)\,.
\end{aligned}
$$

The claim follows for any $0 < \varepsilon < 1$ by combining this last inequality as well as (11.55) with (11.41). $\qquad\square$

Remark 11.2.6. Equation (11.56) implies that the isoperimetric inequality is saturated for Gaussian states in the limit of an infinite mean photon number. This fact was already noticed in [Huber et al., 2017]. One could similarly carry out the same kind of analysis for systems of a finite number of modes.

11.3. EPI and LSI for the quantum Ornstein Uhlenbeck semigroup

As we saw in Section 4.5, the classical entropic isoperimetric inequality can be used to derive the logarithmic Sobolev inequality for the Ornstein Uhlenbeck semigroup. This is a manifestation of the strong link existing between the latter and the heat semigroup (cf. Section 4.5). The same type of argument seems to fail in the non-commutative regime. However, another entropy power inequality corresponding to a convolution between quantum states does the job, as observed in [De Palma and Huber, 2018]. Here, we briefly review the argument for sake of comparison with Section 4.5.

A quantum convolution via beamsplitters The *quantum-quantum convolution* that we mentioned in the previous paragraph was first introduced in [Koenig and Smith, 2014]. It is motived from physics since defined in terms of the beamsplitter (cf Section 0.2): let ρ_1 and ρ_2 be states in $\mathcal{D}(\mathcal{H})$, $\mathcal{H} \coloneqq \mathbb{L}_2(\mathbb{R}^n)$, and $0 < \lambda < 1$, the λ-*quantum convolution* between ρ_1 and ρ_2 is defined as

$$\rho_1 \boxplus_\lambda \rho_2 \coloneqq \mathrm{Tr}_2\left(U_\lambda\left(\rho_1 \otimes \rho_2\right) U_\lambda^{-1}\right),$$

where U_λ is defined through the symplectic matrix (0.47). The quantum entropy power inequality corresponding to the above convolution was proved in [Koenig and Smith, 2014, De Palma and Trevisan, 2018]:

Theorem 11.3.1 (Quantum EPI). *For any two quantum states $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$ of finite second moments, and any $0 \leq \lambda, \eta \leq 1$,*

$$S(\rho_1 \boxplus_\eta \rho_2) \geq \lambda S(\rho_1) + (1-\lambda) S(\rho_2) + \lambda \ln \frac{\eta}{\lambda} + (1-\lambda) \ln \frac{|1-\eta|}{1-\lambda}\,.$$

After maximization over λ, we end up with the following quantum entropy power inequality:

$$N(\rho_1 \boxplus_\lambda \rho_2) \geq \lambda N(\rho_1) + (1 - \lambda)N(\rho_2). \qquad \text{(qq-EPI)}$$

The quantum entropy power inequality provides a two-lines proof of the exponential convergence in relative entropy of the quantum Ornstein Uhlenbeck semigroup of Section 5.5.2 towards its invariant state σ_ν. This observation was made in Theorem 9 of [De Palma and Huber, 2018]. Indeed, for each $t \geq 0$,

$$\mathcal{P}_{*t}^{\mathrm{qOU}}(\rho) = \mathrm{Tr}_2\left(U_{\lambda(t)}\left(\rho \otimes \sigma_\nu\right)U_{\lambda(t)}^*\right),$$

where $\lambda(t) := \mathrm{e}^{-(\mu^2 - \lambda^2)t}$, which can be easily checked by looking at the action of the semigroup on the creation and anihilation operators. Then

Theorem 11.3.2. *The quantum Ornstein Uhlenbeck semigroup $(\mathcal{P}_t^{\mathrm{qOU}})_{t \geq 0}$ satisfies the following identity: for any initial state ρ of finite first and second moments:*

$$D(\mathcal{P}_{*t}^{\mathrm{qOU}}(\rho)\|\sigma_\nu) \leq \mathrm{e}^{-(\mu^2 - \lambda^2)t}\, D(\rho\|\sigma_\nu).$$

Proof. Let $\rho_t := \mathcal{P}_{*t}^{\mathrm{qOU}}(\rho)$ and $\lambda(t) := \mathrm{e}^{-(\mu^2 - \lambda^2)t}$. First, observe that

$$
\begin{aligned}
D(\rho_t\|\sigma_\nu) &= -S(\rho_t) - \mathrm{Tr}(\rho_t \ln \sigma_\nu) \\
&= -S(\rho_t) - \ln\nu\,\mathrm{Tr}(\rho_t N) - \ln(1 - \nu),
\end{aligned}
$$

where N is the total number operator. Now, a simple calculation made e.g. p. 23 of [Huber et al., 2017] implies

$$\mathrm{Tr}(\rho_t N) = \lambda(t)\,\mathrm{Tr}(\rho N) + (1 - \lambda(t))\,\mathrm{Tr}(\sigma_\nu N).$$

Therefore

$$
\begin{aligned}
D(\rho_t\|\sigma_\nu) &= -S(\rho_t) - \ln\nu\,\left(\lambda(t)\,\mathrm{Tr}(\rho N) + (1 - \lambda(t))\,\mathrm{Tr}(\sigma_\nu N)\right) - \ln(1 - \nu) \\
&\leq -\lambda(t)S(\rho) - (1 - \lambda(t))S(\sigma_\nu) - \lambda(t)\ln\nu\,\mathrm{Tr}(\rho N) - \ln\nu(1 - \lambda(t))\,\mathrm{Tr}(\sigma_\nu N) - \ln(1 - \nu) \\
&= -\lambda(t)D(\rho\|\sigma_\nu) - (1 - \lambda(t))S(\sigma_\nu) - (1 - \lambda(t))\ln\nu\,\mathrm{Tr}(\sigma_\nu N) - (1 - \lambda(t))\ln(1 - \nu) \\
&= -\lambda(t)D(\rho\|\sigma_\nu),
\end{aligned}
$$

where the second line follows from Theorem 11.3.1. $\qquad\square$

Chapter 12.

Quantum optimal transport

In Chapter 4, classical functional inequalities were shown to arise from a single geometric inequality called the Ricci curvature lower bound. In particular, we saw in Section 4.2.2 that the Wasserstein distance of order 2 can be interpreted as a metric on the weak Riemmanian manifold of probability measures $\mathcal{P}(\mathcal{M})$. This lead to a rethinking of the geometry of $\mathcal{M}$ in terms of the convexity properties of a family of functionals on $\mathcal{P}(\mathcal{M})$. From the formal calculus introduced by Otto, the generator of a Markov process naturally arises from the differential structure of $\mathcal{M}$. In the case of Markov chains on non-geodesic, finite sample spaces, the problem is reversed: from the generator of the chain, and its associated transition kernel, [Maas, 2011, Erbar and Maas, 2012] introduced a formal differential calculus and a modified Wasserstein distance that lead to the definition of a Ricci curvature lower bound in terms of the displacement convexity of Boltzman's H functional. Recently, Carlen and Maas introduced an extension of the latter in the context of quantum Markov semigroups acting on finite dimensional C*-algebras [Carlen and Maas, 2014, Carlen and Maas, 2017].

Layout of the chapter: In this chapter, we further analyse the quantum Ricci lower bound $(\mathrm{Ric}(\mathcal{L}) \geq \kappa)$ introduced by Carlen and Maas [Carlen and Maas, 2014, Carlen and Maas, 2017], and derive various equivalent formulations of it. In Section 12.1, we introduce the quantum Wasserstein distance. The quantum version of κ-displacement convexity is studied in Section 12.3. The connection between the Ricci curvature lower bound and the quantum functional and transportation cost inequalities is fleshed out in Section 12.5. This is done by showing that $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ implies a quantum version of the celebrated $\mathrm{HWI}(\kappa)$ inequality which interpolates between the modified logarithmic Sobolev inequality and the a quantum version of the transportation cost inequality of order 2 ($\mathrm{TC}_2(c_2)$). In particular, we show that, in the case of $\kappa > 0$, $\mathrm{HWI}(\kappa) \Rightarrow \mathrm{MLSI}(\kappa)$, recovering the result of [Carlen and Maas, 2017]. On the other hand, we establish that in the case when $\kappa \in \mathbb{R}$, $\mathrm{Ric}(\mathcal{L} \geq \kappa)$ together with $\mathrm{TC}_2(c_2)$ imply $\mathrm{MLSI}(\alpha)$. When $\kappa = 0$, we show that, under the assumption of boundedness of the diameter D of the set of states with respect to the quantum Wasserstein distance $W_{2,\mathcal{L}}$, $\mathrm{Ric}(\mathcal{L}) \geq 0$ implies $\mathrm{PI}(c_1 D^{-2})$ for some universal constant c_1. When the QMS is doubly stochastic (i.e. one which has the completely mixed state as its unique invariant state), we show that $\mathrm{Ric}(\mathcal{L}) \geq 0$ also implies $\mathrm{MLSI}(c_2 D^{-2})$ for some universal positive constant c_2. Quantum concentration inequalities are derived from the Poincaré and the transportation cost inequalities in Section 12.7. Finally, we briefly mention extensions of the theory to infinite dimensional systems in Section 12.8.

12.1. The Wasserstein distance $W_{2,\mathcal{L}}$

In this section, we recall the Benamou-Brenier-like construction of the Wasserstein metric $W_{2,\mathcal{L}}$ first defined in [Carlen and Maas, 2017]. Our goal is to define a Riemmanian manifold structure on the set $\mathcal{D}(\mathcal{H})$ of quantum states. Let $\mathcal{L}$ be the generator of a primitive QMS $(\mathcal{P}_t)_{t\geq 0}$ on $\mathcal{B}(\mathcal{H})$, $d_{\mathcal{H}} < \infty$, with invariant state σ. We further assume that $(\mathcal{P}_t)_{t\geq 0}$ satisfies σ-DBC. We recall from Section 5.2.2 that $\mathcal{L}$ can then be written in the following form (cf. Equation (5.20)):

$$\mathcal{L}(X) = \sum_{j\in\mathcal{J}} e^{-\omega_j/2} \left(\tilde{L}_j^*[X, \tilde{L}_j] + [\tilde{L}_j^*, X]\tilde{L}_j \right), \tag{12.1}$$

where $\Delta_\sigma(\tilde{L}_j) = e^{-\omega_j}\tilde{L}_j$ for some $\omega_j \in \mathbb{R}$.

Remark 12.1.1. Given a primitive quantum Markov semigroup with generator which is self-adjoint with respect to $\langle.,.\rangle_{1,\sigma}$, the operators $\tilde{L}_j$ in Equation (12.1) are uniquely defined up to a unitary transformation, i.e. for any two representations $(\{\tilde{L}_j\}, \{\omega_j\})$ and $(\{\tilde{L}_j'\}, \{\omega_j'\})$ of $\mathcal{L}$, there exists a $|\mathcal{J}| \times |\mathcal{J}|$ unitary matrix U such that, unless $\omega_j = \omega_k'$, $U_{jk} = 0$ and

$$\sum_{k\in\mathcal{J}} \overline{U}_{kj} U_{kl} = \delta_{jl}, \qquad \tilde{L}_l' = \sum_{k\in\mathcal{J}} U_{lk}\tilde{L}_k, \tag{12.2}$$

We also recall from Section 5.3 that, given an operator $X \in \mathcal{B}(\mathcal{H})$, its noncommutative gradient and divergence are respectively defined as:

$$\nabla_{\tilde{\mathbf{L}}}X := (\nabla_{\tilde{L}_j}X)_{j\in\mathcal{J}}, \qquad \operatorname{div}_{\tilde{\mathbf{L}}}(\mathbf{A}) := \sum_{j\in\mathcal{J}}[A_j, \tilde{L}_j^*] \equiv -\sum_{j\in\mathcal{J}} \nabla_{\tilde{L}_j}^* A_j,$$

for $X \in \mathcal{B}(\mathcal{H})$ and $\mathbf{A} \equiv (A_j)_{j\in\mathcal{J}} \in \bigoplus_{j\in\mathcal{J}} \mathcal{B}(\mathcal{H})$, where $\nabla_{\tilde{L}_j}X = [\tilde{L}_j, X]$ and $\nabla_{\tilde{L}_j}^* X := [\tilde{L}_j^*, X]$ for all $j \in \mathcal{J}$. Then, the noncommutative Laplacian takes the following simple form: for $X \in \mathcal{B}(\mathcal{H})$:

$$\Delta(X) := \operatorname{div}_{\tilde{\mathbf{L}}} \circ \nabla_{\tilde{\mathbf{L}}}(X) = -\sum_{j\in\mathcal{J}} \nabla_{\tilde{L}_j}^* \circ \nabla_{\tilde{L}_j}(X).$$

In its Benamou-Brenier formulation, the definition of the commutative Wasserstein distance relies on a continuity equation which admits a unique minimal solution (cf. Section 4.2.1). Similarly, in their finite dimensional quantum setting, [Carlen and Maas, 2017] introduced the following noncommutative version of the continuity equation and showed the existence of a minimal solution. First, since $(\mathcal{P}_t)_{t\geq 0}$ is primitive, $\ker(\Delta) = \ker(\nabla_{\tilde{\mathbf{L}}}) = \mathbb{C}\mathbb{1}_{\mathcal{H}}$, and hence the *noncommutative Poisson equation*

$$\Delta(X) = B \qquad \text{(Poisson equation)}$$

has a solution if and only if B is orthogonal to $\mathbb{1}_{\mathcal{H}}$ with respect to $\langle.,.\rangle_{\mathrm{HS}}$, i.e. $\operatorname{Tr}(B) = \langle \mathbb{1}, B\rangle_{\mathrm{HS}} = 0$. As we will see later, given a full-rank state ρ and family φ of functions $\varphi_k : [0, \infty) \to [0, \infty)$, $k \in \mathcal{J}$, it is convenient to introduce the following operator of multiplication by ρ on $\bigoplus_{j\in\mathcal{J}} \mathcal{B}(\mathcal{H})$:

$$[\rho]_\varphi = \{R_\rho \circ \varphi_k(\Delta_\rho)\}_{k\in\mathcal{J}}. \tag{12.3}$$

where $R_\rho(X) := X\rho$. The choice of the functions φ_k will be fixed by the requirement that the evolution induced by $(\mathcal{P}_t)_{t\geq 0}$ is the gradient-flow associated with the relative entropy functional $D(.\|\sigma)$. For the time being, we keep this degree of freedom untouched, and simply observe that, in the commuting

case,

$$[\rho]_\varphi(\mathbf{A}) = (\varphi_k(1)\, A_k \rho)_{k \in \mathcal{J}}\,.$$

Next, given a continuously differentiable path $(\gamma(s))_{s \in [-\varepsilon,\varepsilon]}$ of full-rank states, with $\partial_s \gamma(s) = \dot{\gamma}(s) \perp_{\mathcal{T}_2(\mathcal{H})} \mathbb{C}\mathbb{1}_\mathcal{H}$, there exists a path $\mathbf{V} : [-\varepsilon, \varepsilon] \to \bigoplus_{j \in \mathcal{J}} \mathcal{B}(\mathcal{H})$ such that

$$-\operatorname{div}_{\tilde{\mathbf{L}}}([\gamma(s)]_\varphi \mathbf{V}(s)) = \dot{\gamma}(s)\,. \tag{12.4}$$

Indeed, the above identity is satisfied by the vector field $\mathbf{V}(s) = [\gamma(s)]_\varphi^{-1} \nabla_{\tilde{\mathbf{L}}}(A)(s)$, where $A : [-\varepsilon, \varepsilon] \to \mathcal{B}(\mathcal{H})$ denotes the solution to the Poisson equation. However, the field $\mathbf{V}$ is not unique in general, since any solution of the form $\mathbf{V} + \mathbf{W}$, where $\mathbf{W} = [\gamma]_\varphi^{-1}(\mathbf{F})$, $\operatorname{div}_{\tilde{\mathbf{L}}} \mathbf{F} = 0$, would satisfy Equation (12.4). However, just like in the classical setting, there exists a minimal such solution, which is constructed as follows: for any $\rho \in \mathcal{D}_+(\mathcal{H})$, define the following *modified weighted $\mathbb{L}_{2,\varphi}(\rho)$ inner product* on $\bigoplus_{j \in \mathcal{J}} \mathcal{B}(\mathcal{H})$: for any $\mathbf{A} = (A_j)_{j \in \mathcal{J}}$ and $\mathbf{B} = (B_j)_{j \in \mathcal{J}}$:

$$\langle \mathbf{A}, \mathbf{B} \rangle_{\mathbb{L}_{2,\varphi}(\rho)} := \sum_{j \in \mathcal{J}} \langle A_j, [\rho]_{\varphi,j}(B_j) \rangle_{\mathrm{HS}}\,. \tag{12.5}$$

Since the solutions to Equation (12.4) span a closed affine subspace of $\bigoplus_{j \in \mathcal{J}} \mathcal{B}(\mathcal{H})$, there exists a unique solution $\mathbf{V}(s)$ of minimal $\mathbb{L}_{2,\varphi}(\gamma(s))$ norm, such that for any $\mathbf{W} = [\gamma]_\varphi^{-1}(\mathbf{F})$, $\operatorname{div}_{\tilde{\mathbf{L}}} \mathbf{F} = 0$,

$$\langle \mathbf{V}, \mathbf{F} \rangle_{\bigoplus_{j \in \mathcal{J}} \mathcal{T}_2(\mathcal{H})} = \langle \mathbf{V}, \mathbf{W} \rangle_{\mathbb{L}_{2,\varphi}(\gamma)} = 0\,.$$

In other words, $\mathbf{V} \perp_{\bigoplus_{j \in \mathcal{J}} \mathcal{T}_2(\mathcal{H})} \ker(\operatorname{div}_{\tilde{\mathbf{L}}})$, which implies that $\mathbf{V} \in \operatorname{im}(\nabla_{\tilde{\mathbf{L}}})$. Since $\ker(\nabla_{\tilde{\mathbf{L}}}) = \mathbb{C}\mathbb{1}_\mathcal{H}$ by primitivity, there exists a unique traceless operator valued function $\Phi : [-\varepsilon, \varepsilon] \to \mathcal{B}(\mathcal{H})$ such that $\mathbf{V} = \nabla_{\tilde{\mathbf{L}}}\Phi$, so that

$$-\operatorname{div}_{\tilde{\mathbf{L}}}([\gamma(s)]_\varphi(\nabla_{\tilde{\mathbf{L}}}(\Phi(s))) = \dot{\gamma}(s)\,. \qquad\text{(n-c continuity equation)}$$

The vector $\nabla_{\tilde{\mathbf{L}}}(\Phi(0))$ should be interpreted as the tangent vector at the origin associated with the path γ in the manifold $\mathcal{D}_+(\mathcal{H})$ of full-rank states.

Remark 12.1.2. If $-\operatorname{div}_{\tilde{\mathbf{L}}}([\rho]_\varphi(\nabla_{\tilde{\mathbf{L}}}(X^*)) = -\operatorname{div}_{\tilde{\mathbf{L}}}([\rho]_\varphi(\nabla_{\tilde{\mathbf{L}}}(X))^*$ for any $\rho \in \mathcal{D}_+(\mathcal{H})$ and $X \in \mathcal{B}(\mathcal{H})$, $\Phi(s)$ is moreover self-adjoint for any $s \in [-\varepsilon, \varepsilon]$, by uniqueness of the solution and self-adjointness of $\dot{\gamma}$. This is the case in particular for the following choice for φ: Given $\boldsymbol{\omega} := (\omega_j)_{j \in \mathcal{J}}$, define the linear operator $[\rho]_\varphi \equiv [\rho]_{\varphi^{\log},\boldsymbol{\omega}}$ on $\bigoplus_{j \in \mathcal{J}} \mathcal{B}(\mathcal{H})$ through

$$[\rho]_{\varphi^{\log},\boldsymbol{\omega}}(\mathbf{A}) := ([\rho]_{\varphi^{\log},\omega_j} A_j)_{j \in \mathcal{J}}\,, \qquad \mathbf{A} \equiv (A_j)_{j \in \mathcal{J}}\,,$$

where for any $\omega \in \mathbb{R}$,

$$[\rho]_{\varphi^{\log},\omega} = R_\rho \circ \varphi_\omega^{\log}(\Delta_\rho)\,, \qquad \varphi_\omega^{\log}(t) := \mathrm{e}^{\omega/2}\, \frac{t - \mathrm{e}^{-\omega}}{\ln t + \omega}\,, \qquad t > 0\,. \tag{12.6}$$

The acronym log stands for relative entropy since the following choice of multiplication by ρ will lead to the gradient-flowness of $(\mathcal{P}_t)_{t \geq 0}$ for the relative entropy functional $D(.\|\sigma)$. The following lemma can be used in order to prove that $\Phi = \Phi^*$ in this case (see Lemma 5.8 of [Carlen and Maas, 2017]):

Lemma 12.1.3. *For any $\omega \in \mathbb{R}$, $\rho \in \mathcal{D}_+(\mathcal{H})$ and $A \in \mathcal{B}(\mathcal{H})$,*

$$[\rho]_{\varphi^{\log},\omega}(A) = \int_0^1 e^{\omega(1/2-s)} \rho^s A \rho^{1-s} ds,$$

$$[\rho]_{\varphi^{\log},\omega}^{-1}(A) = \int_0^\infty (s\,\mathrm{id} + e^{-\omega/2} L_\rho)^{-1}(s\,\mathrm{id} + e^{\omega/2} R_\rho)^{-1}(A)\,ds.$$

Now, given two states $\gamma_0, \gamma_1 \in \mathcal{D}_+(\mathcal{H})$, let $\mathcal{P}(\gamma_0, \gamma_1)$ be the collection of smooth paths $\gamma : [0,1] \to \mathcal{D}_+(\mathcal{H})$ satisfying $\gamma(0) = \gamma_0$ and $\gamma(1) = \gamma_1$. Then, the *quantum Wasserstein distance* associated with the Lindblad generator $\mathcal{L}$ is defined as

$$W_{2,\mathcal{L}}^\varphi(\gamma_0, \gamma_1) := \inf_{\gamma \in \mathcal{P}(\gamma_0,\gamma_1)} \left\{ \int_0^1 \|\nabla_{\tilde{\mathbf{L}}} \Phi(s)\|_{\mathbb{L}_{2,\varphi}(\gamma(s))}^2 \, ds \right\}^{\frac{1}{2}},$$

where, for any path γ, $\Phi : [0,1] \to \mathcal{B}(\mathcal{H})$ is the unique corresponding (traceless) solution to the n-c continuity equation. The Wasserstein distance endows the manifold $\mathcal{D}_+(\mathcal{H})$ with a smooth Riemannian structure. We also denote by $g_{\mathcal{L},\varphi}$ the *Wasserstein metric* associated with $W_{2,\mathcal{L}}^\varphi$. More precisely, for any two smooth paths $\gamma, \gamma' : [-\varepsilon, \varepsilon] \to \mathcal{D}_+(\mathcal{H})$ such that $\gamma(0) = \gamma'(0) = \rho$, with associated tangent vectors $\nabla_{\tilde{\mathbf{L}}} \Phi, \nabla_{\tilde{\mathbf{L}}} \Phi'$, at the origin

$$g_{\mathcal{L},\rho}^\varphi(\dot\gamma, \dot\gamma') = \langle \nabla_{\tilde{\mathbf{L}}} \Phi, \nabla_{\tilde{\mathbf{L}}} \Phi' \rangle_{\mathbb{L}_{2,\varphi}(\rho)}. \tag{12.7}$$

When $\varphi = \varphi^{\log}$, we simply write $W_{2,\mathcal{L}}^{\varphi^{\log}} \equiv W_{2,\mathcal{L}}$, $\mathbb{L}_{2,\varphi^{\log}}(\rho) \equiv \mathbb{L}_{2,\log}(\rho)$ and $g_{\mathcal{L},\rho}^{\varphi^{\log}} \equiv g_{\mathcal{L},\rho}$. The following lemma provides an alternative expression for $W_{2,\mathcal{L}}^\varphi$:

Lemma 12.1.4. *With the above notations, the Wasserstein distance $W_{2,\mathcal{L}}^\varphi$ between two full-rank states γ_0, γ_1 is equal to the minimal length over the smooth paths joining γ_0 and γ_1:*

$$W_{2,\mathcal{L}}^\varphi(\gamma_0, \gamma_1) = \inf_{\gamma \text{ const. speed}} \left\{ \int_0^1 \sqrt{g_{\mathcal{L},\gamma(s)}^\varphi(\dot\gamma(s), \dot\gamma(s))} \, ds : \gamma(0) = \gamma_0, \ \gamma(1) = \gamma_1 \right\}, \tag{12.8}$$

where the infimum is taken over the set of constant speed curves γ, i.e. such that $s \mapsto g_{\mathcal{L},\gamma(s)}^\varphi(\dot\gamma(s), \dot\gamma(s))$ is constant on $[0,1]$.

Proof. By the proof of Lemma 1.1.4 (b) of [Ambrosio et al., 2008], for any smooth path $\gamma : [0,1] \to \mathcal{D}_+(\mathcal{H})$, with $\gamma(0) = \gamma_0$ and $\gamma(1) = \gamma_1$ there exists a constant speed path $\eta : [0,1] \to \mathcal{D}_+(\mathcal{H})$ with the same boundary conditions: for any $s \in [0,1]$,

$$\sqrt{g_{\mathcal{L},\eta(s)}^\varphi(\dot\eta(s), \dot\eta(s))} = \int_0^1 \sqrt{g_{\mathcal{L},\gamma(u)}^\varphi(\dot\gamma(u), \dot\gamma(u))} \, du.$$

Fix such a path γ. By the Cauchy-Schwarz inequality, for any $s \in [0,1]$

$$\sqrt{g_{\mathcal{L},\eta(s)}^\varphi(\dot\eta(s), \dot\eta(s))} = \int_0^1 \sqrt{g_{\mathcal{L},\gamma(u)}^\varphi(\dot\gamma(u), \dot\gamma(u))} \, du \le \left(\int_0^1 g_{\mathcal{L},\gamma(u)}^\varphi(\dot\gamma(u), \dot\gamma(u)) \, du \right)^{1/2}. \tag{12.9}$$

with equality if $\gamma = \eta$. The statement follows after taking the infimum over all smooth paths γ on the right hand side of (12.9). $\qquad\square$

Extension of the metric to $\mathcal{D}(\mathcal{H})$: Here, we provide a procedure to extend the definition of the Wasserstein distance to the set $\mathcal{D}(\mathcal{H})$ of all states on $\mathcal{H}$. The following technical lemma is going to play a crucial role:

Lemma 12.1.5. *For any $\rho \in \mathcal{D}_+(\mathcal{H})$, the map $D_{\varphi^{\log}}(\rho) : \Phi \mapsto -\operatorname{div}_{\tilde{\mathbf{L}}}([\rho]_{\varphi^{\log},\boldsymbol{\omega}} \nabla_{\tilde{\mathbf{L}}}\Phi)$ is invertible and positive on the space of traceless operators. Moreover, if $\rho \geq \varepsilon \mathbb{1}$ for some $\varepsilon > 0$, then:*

$$D_{\varphi^{\log}}(\rho)^{-1} \leq K_{\mathcal{L}} \; \varepsilon^{-1} \; \mathrm{id}, \tag{12.10}$$

where $K_{\mathcal{L}} := \max_{j \in \mathcal{J}} \frac{\omega_j}{\mathrm{e}^{\omega_j/2} - \mathrm{e}^{-\omega_j/2}} \| (-\operatorname{div}_{\tilde{\mathbf{L}}} \circ \nabla_{\tilde{\mathbf{L}}}(.))^{-1} : \mathcal{T}_2(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})\|$, and where the inequality is in the sense of the partial order induced by the convex cone of positive semidefinite operators on $\mathcal{T}_2(\mathcal{H})$.

Proof. Let $\mathcal{W}$ be the space of traceless operators on $\mathcal{H}$. We already saw that, for any $B \in \mathcal{W}$, there exists a unique $\Phi \in \mathcal{W}$ such that $B = -\operatorname{div}_{\tilde{\mathbf{L}}}([\rho]_{\varphi^{\log},\boldsymbol{\omega}}(\nabla_{\tilde{\mathbf{L}}}\Phi))$. We also defined $-\operatorname{div}_{\tilde{\mathbf{L}}}$ as $\nabla_{\tilde{\mathbf{L}}}^*$, where the adjoint is taken with respect to the inner products on $\mathcal{T}_2(\mathcal{H})$ and $\oplus_{j \in \mathcal{J}} \mathcal{T}_2(\mathcal{H})$. The result follows from Lemma 12.1.3, since for any $j \in \mathcal{J}$, $[\rho]_{\varphi^{\log},\omega_j} \geq \varepsilon \frac{\mathrm{e}^{\omega_j/2} - \mathrm{e}^{-\omega_j/2}}{\omega_j} \mathrm{id}$. $\qquad \square$

The above proposition allows us to extend the definition of the Wasserstein distance to non-faithful states:

Proposition 12.1.6 (Extension of the metric to $\mathcal{D}(\mathcal{H})$). *Let $\rho, \omega \in \mathcal{D}(\mathcal{H})$ and let $\{\rho_n\}_{n \in \mathbb{N}}$ and $\{\omega_n\}_{n \in \mathbb{N}}$ be sequences of full-rank states satisfying*

$$\mathrm{Tr}[(\rho - \rho_n)^2] \to 0, \qquad \mathrm{Tr}[(\omega - \omega_n)^2] \to 0, \tag{12.11}$$

as $n \to \infty$. Then the sequence $\{W_{2,\mathcal{L}}(\rho_n, \omega_n)\}_{n \in \mathbb{N}}$ converges.

Proof. The proof is similar to the one given in Proposition 4.5 of [Carlen and Maas, 2014]. It is enough to show that $\{W_{2,\mathcal{L}}(\rho_n, \omega_n)\}_{n \in \mathbb{N}}$ is Cauchy. By the triangle inequality, it is even enough to prove that $W_{2,\mathcal{L}}(\rho_n, \rho_m) \to 0$ as $m, n \to \infty$. Let $\varepsilon \in (0, 1)$ and set $\bar{\rho} := (1 - \varepsilon)\rho + \varepsilon \frac{\mathbb{1}}{d_{\mathcal{H}}}$. Let $N \in \mathbb{N}$ be such that for any $n \geq N$, $\mathrm{Tr}[(\rho - \rho_n)^2] \leq \varepsilon^2$. For $n \geq N$, consider the convex interpolation $\gamma(s) := (1 - s)\rho_n + s\bar{\rho}$. Since $\gamma(s) \geq \varepsilon s \frac{\mathbb{1}}{d_{\mathcal{H}}}$ for $s \in [0, 1]$, we find from Equation (12.8) that

$$
\begin{aligned}
&W_{2,\mathcal{L}}(\rho_n, \bar{\rho}) \\
&\leq \int_0^1 \sqrt{g_{\mathcal{L},\gamma(s)}(\dot{\gamma}(s), \dot{\gamma}(s))} \, ds \\
&= \int_0^1 \left[\sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} (-\operatorname{div}_{\tilde{\mathbf{L}}} \circ [\gamma]_{\varphi^{\log},\boldsymbol{\omega}} \circ \nabla_{\tilde{\mathbf{L}}})^{-1}(\dot{\gamma}), [\gamma]_{\varphi^{\log},\omega_j} \nabla_{\tilde{L}_j} (-\operatorname{div}_{\tilde{\mathbf{L}}} \circ [\gamma]_{\varphi^{\log},\boldsymbol{\omega}} \circ \nabla_{\tilde{\mathbf{L}}})^{-1}(\dot{\gamma}) \rangle_{\mathrm{HS}} \right]^{1/2} ds \\
&= \int_0^1 \sqrt{\langle \dot{\gamma}(s), (-\operatorname{div}_{\tilde{\mathbf{L}}}([\gamma(s)]_{\varphi^{\log},\boldsymbol{\omega}} \nabla_{\tilde{\mathbf{L}}}(.)))^{-1} \dot{\gamma}(s) \rangle_{\mathrm{HS}}} \, ds \\
&\leq \sqrt{\frac{K_{\mathcal{L}} d_{\mathcal{H}}}{\varepsilon}} \int_0^1 s^{-1/2} \sqrt{\mathrm{Tr}[(\dot{\gamma}(s))^2]} \, ds,
\end{aligned}
$$

where we used Lemma 12.1.5 in the second, third and fourth above lines. Now

$$
\begin{aligned}
\mathrm{Tr}[(\dot{\gamma}(s))^2] &= \mathrm{Tr}[(\rho - \rho_n + \varepsilon(\mathbb{1}/d_{\mathcal{H}} - \rho))^2] \\
&\leq 2 \, \mathrm{Tr}[(\rho - \rho_n)^2] + 2\varepsilon^2 \, \mathrm{Tr}[(\mathbb{1}/d_{\mathcal{H}} - \rho)^2] \\
&\leq 2 \left(1 + \mathrm{Tr}[(\mathbb{1}/d_{\mathcal{H}} - \rho)^2] \right) \varepsilon^2.
\end{aligned}
$$

Hence, $W_{2,\mathcal{L}}(\rho_n, \bar{\rho}) \leq C_\rho \sqrt{\varepsilon}$, for some constant C_ρ depending on ρ. Since ε was arbitrary, we conclude by triangle inequality that $W_{2,\mathcal{L}}(\rho_m, \rho_n) \leq W_{2,\mathcal{L}}(\rho_m, \bar{\rho}) + W_{2,\mathcal{L}}(\bar{\rho}, \rho_n) \to 0$. $\qquad \square$

The above proposition justifies the following extension: The quantum Wasserstein distance $W_{2,\mathcal{L}}$

between two states $\rho, \omega \in \mathcal{D}(\mathcal{H})$ is defined as

$$W_{2,\mathcal{L}}(\rho, \omega) := \lim_{n \to \infty} W_{2,\mathcal{L}}(\rho_n, \omega_n),$$

where $\{\rho_n\}_{n \in \mathbb{N}}$ and $\{\omega_n\}_{n \in \mathbb{N}}$ are arbitrary sequences in $\mathcal{D}_+(\mathcal{H})$ satisfying (12.11). Next, we recall from Section 1.3 that the functions $k^{\log} : \mathbb{R}_+ \ni t \mapsto (t-1)^{-1} \ln(t)$ belongs to the class

$$\text{OMM} := \{k : \mathbb{R}_+ \to \mathbb{R} \mid -k \text{ is operator monotone}, \ k(x^{-1}) = xk(x), \ k(1) = 1\}. \tag{12.12}$$

Then, since $(\varphi_{\omega_j}^{\log}(t))^{-1} = e^{-\omega_j/2}(\ln t + \omega_j)(t - e^{-\omega_j})^{-1} = e^{\omega_j/2} k^{\log}(e^{\omega_j} t)$, inequality (1.88) implies that, for any $j \in \mathcal{J}$,

$$[\rho]_{\varphi^{\log}, \omega_j} \leq \frac{L_\rho \, e^{\omega_j/2} + R_\rho \, e^{-\omega_j/2}}{2} \, . \tag{12.13}$$

In the next theorem, we show that $(\mathcal{D}(\mathcal{H}), W_{2,\mathcal{L}})$ forms a complete metric space:

Lemma 12.1.7. *For any* $\rho, \omega \in \mathcal{D}(\mathcal{H})$,

$$\|\rho - \omega\|_1 \leq \sqrt{2} \left(\sum_{j \in \mathcal{J}} (e^{-\omega_j/2} + e^{\omega_j/2}) \|\tilde{L}_j\|_\infty^2 \right)^{1/2} W_{2,\mathcal{L}}(\rho, \omega).$$

Proof. The proof is inspired by the proof of Proposition 2.12 of [Erbar and Maas, 2012]. Fix $\delta > 0$, and $\rho, \omega \in \mathcal{D}_+(\mathcal{H})$ without loss of generality. There exists a smooth path $\gamma : [0,1] \to \mathcal{D}_+(\mathcal{H})$ such that $\gamma(0) = \rho$, $\gamma(1) = \omega$, and by definition of $W_{2,\mathcal{L}}$:

$$\left(\int_0^1 \sum_{j \in \mathcal{J}} \langle V(s)_j, [\gamma(s)]_{\varphi^{\log}, \omega_j} V(s)_j \rangle_{\text{HS}} \right)^{1/2} \equiv \left(\int_0^1 \|\mathbf{V}(s)\|_{\mathbb{L}_{2, \varphi^{\log}}(\gamma(s))}^2 \right)^{1/2} \leq W_{2,\mathcal{L}}(\rho, \omega) + \delta, \tag{12.14}$$

where for each $s \in [0,1]$, $\mathbf{V}(s) \equiv (V(s)_j)_{j \in \mathcal{J}}$ is related to $\gamma(s)$ through the continuity equation:

$$\dot{\gamma}(s) + \text{div}_{\tilde{\mathbf{L}}}([\gamma(s)]_{\varphi^{\log}, \boldsymbol{\omega}} \mathbf{V}(s)) = 0,$$

Hence, for any $X \in \mathcal{B}(\mathcal{H})$,

$$|\text{Tr}(X(\rho - \omega))| \tag{12.15}$$

$$= \left| \text{Tr}\left(X \int_0^1 \frac{d}{ds} \gamma(s) ds \right) \right|$$

$$= \left| \int_0^1 \text{Tr}\left(X \, \text{div}_{\tilde{\mathbf{L}}}([\gamma(s)]_{\varphi^{\log}, \boldsymbol{\omega}} \mathbf{V}(s)) \right) \right|$$

$$= \left| \int_0^1 \sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} X, [\gamma(s)]_{\varphi^{\log}, \omega_j} V(s)_j \rangle_{\text{HS}} \, ds \right|$$

$$\leq \left(\int_0^1 \sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} X, [\gamma(s)]_{\varphi^{\log}, \omega_j} \nabla_{\tilde{L}_j} X \rangle_{\text{HS}} \, ds \right)^{1/2} \left(\int_0^1 \sum_{j \in \mathcal{J}} \langle V(s)_j, [\gamma(s)]_{\varphi^{\log}, \omega_j} V(s)_j \rangle_{\text{HS}} \, ds \right)^{1/2}$$

$$\leq \left(\int_0^1 \sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} X, [\gamma(s)]_{\varphi^{\log}, \omega_j} \nabla_{\tilde{L}_j} X \rangle_{\text{HS}} \, ds \right)^{1/2} (W_{2,\mathcal{L}}(\rho, \omega) + \delta), \tag{12.16}$$

where in the fourth line we used the Cauchy-Schwarz inequality with respect to the inner product $\sum_{j \in \mathcal{J}} \int_0^1 \langle ., [\gamma(s)]_{\varphi^{\log}, \omega_j} . \rangle_{\text{HS}} \, ds$, and the last line comes from (12.14). Now, for each $j \in \mathcal{J}$, and any

$s \in [0,1]$, (12.13) for $\rho = \gamma(s)$ leads to

$$\langle \nabla_{\tilde{L}_j} X, [\gamma(s)]_{\varphi^{\log},\omega_j} \nabla_{\tilde{L}_j} X \rangle_{\mathrm{HS}} \le \frac{1}{2} \left[\mathrm{e}^{-\omega_j/2} \operatorname{Tr}(\gamma(s)(\nabla_{\tilde{L}_j} X)^* \nabla_{\tilde{L}_j} X) + \mathrm{e}^{\omega_j/2} \operatorname{Tr}((\nabla_{\tilde{L}_j} X)^* \gamma(s) \nabla_{\tilde{L}_j} X) \right]$$

$$\le \frac{\mathrm{e}^{-\omega_j/2} + \mathrm{e}^{\omega_j/2}}{2} \|\nabla_{\tilde{L}_j} X\|_\infty^2 \,, \tag{12.17}$$

where the last line follows by Hölder's inequality. Substituting this bound into (12.16), we end up with

$$|\operatorname{Tr}(X(\rho - \omega))| \le 2^{-1/2} \left(\sum_{j \in \mathcal{J}} (\mathrm{e}^{-\omega_j/2} + \mathrm{e}^{\omega_j/2}) \|\nabla_{\tilde{L}_j}(X)\|_\infty^2 \right)^{1/2} (W_{2,\mathcal{L}}(\rho, \omega) + \delta) \tag{12.18}$$

$$\le \sqrt{2} \, \|X\|_\infty \left(\sum_{j \in \mathcal{J}} (\mathrm{e}^{-\omega_j/2} + \mathrm{e}^{\omega_j/2}) \|\tilde{L}_j\|_\infty^2 \right)^{1/2} (W_{2,\mathcal{L}}(\rho, \omega) + \delta) \,.$$

The result follows after taking the limit $\delta \to 0$ and by duality between $\|.\|_1$ and $\|.\|_\infty$.

$\square$

Proposition 12.1.8. *The metric space $(\mathcal{D}(\mathcal{H}), W_{2,\mathcal{L}})$ is complete.*

Proof. This directly follows from Lemma 12.1.7 and Proposition 12.1.6: assume that $\{\rho_n\}_{n \in \mathbb{N}}$ is a Cauchy sequence in $(\mathcal{D}(\mathcal{H}), W_{2,\mathcal{L}})$, that is $W_{2,\mathcal{L}}(\rho_n, \rho_m) \to 0$ as $m, n \to \infty$. Then, by Lemma 12.1.7, $\{\rho_n\}_{n \in \mathbb{N}}$ is also Cauchy with respect to the trace norm $\|.\|_1$. By completeness of the normed vector space $(\mathcal{B}(\mathcal{H}), \|.\|_1)$, this implies existence of $\rho_\infty \in \mathcal{D}(\mathcal{H})$ such that $\|\rho_n - \rho_\infty\|_1 \to 0$ as $n \to \infty$. We conclude that $W_{2,\mathcal{L}}(\rho_n, \rho_\infty) \to W_{2,\mathcal{L}}(\rho_\infty, \rho_\infty) = 0$ by Proposition 12.1.6. $\square$

12.2. Gradient formula

From now on, we restrict ourselves to the study of $W_{2,\mathcal{L}} \equiv W_{2,\mathcal{L}}^{\varphi^{\log}}$, where the functions $\varphi^{\log}_{\omega_j}$ are given in Equation (12.6). In particular, we show that the evolution $(\mathcal{P}_t)_{t \ge 0}$ is the gradient flow for the relative entropy functional $D(.\|\sigma)$ with respect to $W_{2,\mathcal{L}}$. This means that $\mathcal{L}_* \rho = -\operatorname{grad}_{\mathcal{L}} D(\rho\|\sigma)$, where the gradient $\operatorname{grad}_{\mathcal{L}} \mathcal{F}(\rho)$ of a differentiable functional $\mathcal{F} : \mathcal{D}_+(\mathcal{H}) \to \mathbb{R}$ at a point ρ is defined as the unique element in the tangent space at ρ so that

$$\frac{d}{dt} \mathcal{F}(\gamma(t)) \Big|_{t=0} = g_{\mathcal{L},\rho}(\dot{\gamma}, \operatorname{grad}_{\mathcal{L}} \mathcal{F}(\rho)) \tag{12.19}$$

for all smooth paths $\gamma(t)$ defined on $(-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$ with $\gamma(0) = \rho$. In particular, for $\gamma(t) = \rho_t \equiv \mathcal{P}_{*t}(\rho)$,

$$\frac{d}{dt} D(\rho_t\|\sigma) \Big|_{t=0} = -g_{\mathcal{L},\rho}(\mathcal{L}_*(\rho), \mathcal{L}_*(\rho)) \,. \tag{12.20}$$

This result was first proved in [Carlen and Maas, 2017], but we hope our derivation makes the link with the general chain rule of Section 5.3 clearer. First recall that the generator $\mathcal{L}_*$ of $(\mathcal{P}_{*t})_{t \ge 0}$ in the Schrödinger picture takes the form (cf. Theorem 5.10 of [Carlen and Maas, 2017])

$$\mathcal{L}_*(\rho) = -\sum_{j \in \mathcal{J}} \nabla_{\tilde{L}_j^*}(\tilde{L}_j \, \mathrm{e}^{-\omega_j/2} \rho - \mathrm{e}^{\omega_j/2} \rho \tilde{L}_j) \,.$$

Moreover, by simple integral representation of the logarithm, one can show that for any smooth path $\gamma : [-\varepsilon, \varepsilon] \to \mathcal{D}_+(\mathcal{H})$ with $\gamma(0) = \rho$,

$$\frac{d}{ds}D(\gamma(s)\|\sigma)\Big|_{s=0} = \mathrm{Tr}\left[\dot{\gamma}(0)(\ln\rho - \ln\sigma)\right] \tag{12.21}$$

$$= -\mathrm{Tr}\left[\mathrm{div}_{\tilde{\mathbf{L}}}([\rho]_{\varphi^{\log},\boldsymbol{\omega}}\nabla_{\tilde{\mathbf{L}}}(\Phi))(\ln\rho - \ln\sigma)\right] \tag{12.22}$$

$$= \langle \nabla_{\tilde{\mathbf{L}}}\Phi, \nabla_{\tilde{\mathbf{L}}}(\ln\rho - \ln\sigma)\rangle_{\mathbb{L}_{2,\log}(\rho)}, \tag{12.23}$$

where the second line comes from the n-c continuity equation for the couple (γ, Φ) at $s = 0$. Equation (12.19) would follow with $\mathrm{grad}_{\mathcal{L}} D(\rho\|\sigma) = -\mathcal{L}_*(\rho)$ if one can show that

$$\mathcal{L}_*(\rho) = \mathrm{div}_{\tilde{\mathbf{L}}}([\rho]_{\varphi^{\log},\boldsymbol{\omega}}\nabla_{\tilde{\mathbf{L}}}(\ln\rho - \ln\sigma)). \tag{12.24}$$

We prove Equation (12.24) using the general chain rule provided in Theorem 5.3.5. First, notice that $[\rho]_{\varphi^{\log},\boldsymbol{\omega}}$ can be re-expressed as follows:

$$[\rho]_{\varphi^{\log},\omega_j} = \frac{e^{\omega_j/2} L_\rho - e^{-\omega_j/2} R_\rho}{\ln(\Delta_\rho) + \omega_j}$$

$$= \frac{e^{\omega_j/2} L_\rho - e^{-\omega_j/2} R_\rho}{\ln(e^{\omega_j/2} L_\rho) - \ln(e^{-\omega_j/2} R_\rho)}$$

$$= T^{e^{\omega_j/2}\rho, e^{-\omega_j/2}\rho}_{(x,y)\mapsto \frac{x-y}{\ln x - \ln y}}, \tag{12.25}$$

where the second line follows from the commutativity of L_ρ and R_ρ. Moreover, for any $j \in \mathcal{J}$

$$\nabla_{\tilde{L}_j}(\ln\rho - \ln\sigma) = \nabla_{\tilde{L}_j}\ln\rho + \frac{d}{ds}\Big|_{s=0}\Delta_\sigma^s(\tilde{L}_j)$$

$$= \nabla_{\tilde{L}_j}\ln\rho + \frac{d}{ds}\Big|_{s=0} e^{-\omega_j s}\tilde{L}_j$$

$$= \nabla_{\tilde{L}_j}\ln\rho - \omega_j\tilde{L}_j$$

$$= \tilde{L}_j\ln(e^{-\omega_j/2}\rho) - \ln(e^{\omega_j/2}\rho)\tilde{L}_j. \tag{12.26}$$

Therefore, the right hand side of Equation (12.24) reads

$$-\sum_{j\in\mathcal{J}}\nabla_{\tilde{L}_j^*}\left(T^{e^{\omega_j/2}\rho, e^{-\omega_j/2}\rho}_{(x,y)\mapsto \frac{x-y}{\ln x - \ln y}}\left\{\tilde{L}_j\ln(e^{-\omega_j/2}\rho) - \ln(e^{\omega_j/2}\rho)\tilde{L}_j\right\}\right) = -\sum_{j\in\mathcal{J}}\nabla_{\tilde{L}_j^*}(\tilde{L}_j e^{-\omega_j/2}\rho - e^{\omega_j/2}\rho\tilde{L}_j)$$

$$\equiv \mathcal{L}_*(\rho),$$

where the above equality comes from Theorem 5.3.5.

12.3. Quantum displacement convexity

The notion of displacement convexity that we introduced in Section 4.2.3 was extended to the quantum realm by [Carlen and Maas, 2014] for the particular case of the Fermionic Fokker-Planck equation. In [Carlen and Maas, 2017], the same authors took a slightly different path inspired by [Otto and Westdickenberg, 2005, Daneri and Savaré, 2008]. In the latter, the authors showed the geodesic convexity of a function F defined on an abstract metric space under the condition of contractivity of the gradient flow for F for the corresponding distance. Carlen and Maas showed that this approach, when used for a QMS $(\mathcal{P}_{*t})_{t\geq 0}$, that is the gradient flow of the relative entropy functional $D(.\|\sigma)$ in

$(\mathcal{D}_+(\mathcal{H}), W_{2,\mathcal{L}})$, leads to the modified logarithmic Sobolev inequality. In this section, we provide a systematic analysis of the initial notion of displacement convexity in the setting of [Carlen and Maas, 2017], including a study of the geodesic equations on the Riemannian manifold $(\mathcal{D}_+(\mathcal{H}), g_{\mathcal{L}})$.

12.3.1. Geodesic equations

Since $(\mathcal{D}_+(\mathcal{H}), g_{\mathcal{L}})$ defines a valid Riemannian manifold, the local existence and uniqueness of constant speed geodesics is garanteed by standard Riemannian geometry. We first recall that constant speed geodesics $(\gamma(s))_{s \in [0,1]}$, satisfy a Euler-Lagrange equation that we derive in Theorem 12.3.1. This extends Theorem 5.3 of [Carlen and Maas, 2014].

Let $(\mathcal{V}, \langle .,. \rangle)$ be a finite-dimensional real Hilbert space, $\mathcal{W} \subset \mathcal{V}$ be a subspace of $\mathcal{V}$, and $z \in \mathcal{V} \backslash \mathcal{W}$. Consider the affine subspace $\mathcal{W}_z := z + \mathcal{W}$, and let $\mathcal{M} \subset \mathcal{W}_z$ be a relatively open subset. Let $D : \mathcal{M} \to \mathcal{B}(\mathcal{W})$ be a smooth function such that $D(x)$ is self-adjoint and invertible for all $x \in \mathcal{M}$. We shall write $C(x) := D(x)^{-1}$. Consider the Lagrangian $L : \mathcal{W} \times \mathcal{M} \to \mathbb{R}$ defined by $L(p, x) = \langle C(x)p, p \rangle$ and the associated minimization problem:

$$\inf_{u(.) \in C^1([0,1], \mathcal{M})} \left(\int_0^1 L(u'(t), u(t)) dt : \quad u(0) = u_0, \ u(1) = u_1 \right),$$

where $u_0, u_1 \in \mathcal{M}$ are given boundary values. Then the Euler-Lagrange equations are equivalent to the following system of equations:

$$\begin{cases} u'(t) - D(u(t))v(t) = 0, \\ v'(t) + \dfrac{1}{2} \langle \partial_x D(u(t))v(t), v(t) \rangle = 0. \end{cases} \tag{12.27}$$

Here, we apply this abstract result to the case where $\mathcal{V} = \mathcal{B}_{sa}(\mathcal{H})$, with inner product $\langle .,. \rangle$ the usual Hilbert-Schmidt inner product, $\mathcal{W} = \{ A \in \mathcal{V} : \operatorname{Tr}(A) = 0 \}$, $z := \mathbb{1}/\dim(\mathcal{H})$, and $\mathcal{M} = \mathcal{D}_+(\mathcal{H})$. Indeed any density operator ρ can be written as $\rho = \mathbb{1}/\dim \mathcal{H} + K$, for some self-adjoint and traceless operator K. For any $\rho \in \mathcal{D}_+(\mathcal{H})$, we already proved in Lemma 12.1.5 that $D_{\varphi^{\log}}(\rho) : \Phi \mapsto -\operatorname{div}_{\tilde{\mathbf{L}}}([\rho]_{\varphi^{\log}, \boldsymbol{\omega}} \nabla_{\tilde{\mathbf{L}}} \Phi)$ is invertible and self-adjoint. Now we use the following identity (see [Carlen and Maas, 2014] p. 21):

$$\frac{d}{dt}(\rho + tA)^{\alpha} \Big|_{t=0} = \int_0^1 \int_0^{\alpha} \frac{\rho^{\alpha - \beta}}{(1-s)\mathbb{1} + s\rho} A \frac{\rho^{\beta}}{(1-s)\mathbb{1} + s\rho} \, d\beta ds \tag{12.28}$$

for any $0 < \alpha < 1$, $\rho \in \mathcal{D}_+(\mathcal{H})$ and $A \in \mathcal{W}$. Hence for all $A, \Phi \in \mathcal{W}$,

$$\begin{aligned} \frac{d}{dt}\Big|_{t=0} \langle D_{\varphi^{\log}}(\rho + tA)[\Phi], \Phi \rangle_{\mathrm{HS}} &= \frac{d}{dt}\Big|_{t=0} \sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} \Phi, [\rho + tA]_{\varphi^{\log}, \omega_j} \nabla_{\tilde{L}_j} \Phi \rangle_{\mathrm{HS}} \\ &= \frac{d}{dt}\Big|_{t=0} \sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} \Phi, \int_0^1 e^{\omega_j(1/2 - \alpha)} (\rho + tA)^{\alpha} \nabla_{\tilde{L}_j} \Phi (\rho + tA)^{1-\alpha} \rangle_{\mathrm{HS}} \, d\alpha \\ &= \langle A, \nabla_{\tilde{\mathbf{L}}} \Phi \cdot_{\rho} \nabla_{\tilde{\mathbf{L}}} \Phi \rangle_{\mathrm{HS}}, \end{aligned} \tag{12.29}$$

where for two vectors $\vec{V}_1, \vec{V}_2$ in $\bigoplus_{j \in \mathcal{J}} \mathcal{B}(\mathcal{H})$,

$$\vec{V}_1 \cdot_{\rho} \vec{V}_2 := \sum_{j \in \mathcal{J}} \int_0^1 \int_0^1 e^{\omega_j(1/2 - \alpha)} (\chi_j(\vec{V}_1, \vec{V}_2^*, \rho, \alpha, s) + \chi_j(\vec{V}_1^*, \vec{V}_2, \rho, 1 - \alpha, s)) \, d\alpha ds,$$

where

$$\chi_j(\vec{V}_1, \vec{V}_2, \rho, \alpha, s) := \int_0^\alpha \frac{\rho^\beta}{(1-s)\mathbb{1}+s\rho}(V_1)_j\, \rho^{1-\alpha}\,(V_2)_j \frac{\rho^{\alpha-\beta}}{(1-s)\mathbb{1}+s\rho}d\beta.$$

Therefore, in our context the Euler-Lagrange equations (12.27) reduce to the following:

Theorem 12.3.1. *The geodesic equations in the Riemannian manifold* $(\mathcal{D}_+(\mathcal{H}), W_{2,\mathcal{L}})$ *are given by*

$$\begin{cases} \dfrac{d}{ds}\gamma(s) + \operatorname{div}_{\tilde{\mathbf{L}}}\big([\gamma(s)]_{\varphi^{\log},\boldsymbol{\omega}}\nabla_{\tilde{\mathbf{L}}}\Phi(s)\big) = 0\,, \\[2mm] \dfrac{d\Phi(s)}{ds} + \dfrac{1}{2}\nabla_{\tilde{\mathbf{L}}}\Phi(s)._{\gamma(s)}\nabla_{\tilde{\mathbf{L}}}\Phi(s) = 0\,. \end{cases} \tag{12.30}$$

12.3.2. Equivalent formulations for displacement convexity

In analogy with [Erbar and Maas, 2012], we say that a primitive quantum Markov semigroup $(\mathcal{P}_t)_{t\geq 0}$ with associated invariant state σ and generator $\mathcal{L}$ of the form of Equation (12.1), with $\Delta_\sigma(\tilde{L}_j) = \mathrm{e}^{-\omega_j}\tilde{L}_j$, has *Ricci curvature bounded from below* by a constant $\kappa \in \mathbb{R}$ if the following inequality holds:

$$\frac{d^2}{ds^2}\bigg|_{s=0} D(\gamma(s)\|\sigma) \geq \kappa\, g_{\mathcal{L},\rho}(\dot\gamma(0), \dot\gamma(0))\,, \tag{Ric($\mathcal{L}$) $\geq \kappa$}$$

for any constant speed geodesic $(\gamma(s))_{s\in(-\varepsilon,\varepsilon)}$ on $\mathcal{D}_+(\mathcal{H})$ such that $\gamma(0) = \rho$. This inequality will also be referred to as the *quantum Ricci lower bound*. For a classical intuition behind this definition for the quantum Ricci curvature lower bound, we refer to Equation (4.7). Theorem 12.3.1 is useful to derive an expression for the second derivative of the relative entropy $D(\gamma(s)\|\sigma)$ with respect to s. We already know from the gradient flow equation 12.20 that

$$\begin{aligned} \frac{d}{ds}D(\gamma(s)\|\sigma) &= -g_{\mathcal{L},\gamma(s)}(\dot\gamma(s), \mathcal{L}_*(\gamma(s))) \\ &= \sum_{j\in\mathcal{J}}\langle\nabla_{\tilde{L}_j}\Phi(s), [\gamma(s)]_{\varphi^{\log},\omega_j}\nabla_{\tilde{L}_j}(\ln\gamma(s) - \ln\sigma)\rangle_{\mathrm{HS}} \\ &= \sum_{j\in\mathcal{J}}\langle\nabla_{\tilde{L}_j}\Phi(s), [\gamma(s)]_{\varphi^{\log},\omega_j}(\tilde{L}_j\ln(\mathrm{e}^{-\omega_j/2}\gamma(s)) - \ln(\mathrm{e}^{\omega_j/2}\gamma(s))\tilde{L}_j)\rangle_{\mathrm{HS}}, \end{aligned}$$

where the second line comes from Equation (12.24), and the last identity comes from Equation (12.26). Using the chain rule formula of Theorem 5.3.5 together with Equation (12.25)

$$[\gamma(s)]_{\varphi^{\log},\omega_j}(\tilde{L}_j\ln(\mathrm{e}^{-\omega_j/2}\gamma(s)) - \ln(\mathrm{e}^{\omega_j/2}\gamma(s))\tilde{L}_j) = \mathrm{e}^{-\omega_j/2}\tilde{L}_j\gamma(s) - \mathrm{e}^{\omega_j/2}\gamma(s)\tilde{L}_j,$$

so that we finally get

$$\frac{d}{ds}D(\gamma(s)\|\sigma) = \sum_{j\in\mathcal{J}}\langle\nabla_{\tilde{L}_j}\Phi(s), \mathrm{e}^{-\omega_j/2}\tilde{L}_j\gamma(s) - \mathrm{e}^{\omega_j/2}\gamma(s)\tilde{L}_j\rangle_{\mathrm{HS}}.$$

Differentiating once more, we get:

$$\begin{aligned} \frac{d^2}{ds^2}D(\gamma(s)\|\sigma)\bigg|_{s=0} = \sum_{j\in\mathcal{J}}\bigg\{ &\langle\nabla_{\tilde{L}_j}\frac{d}{ds}\Phi(s)\bigg|_{s=0}, \mathrm{e}^{-\omega_j/2}\tilde{L}_j\rho - \mathrm{e}^{\omega_j/2}\rho\tilde{L}_j\rangle_{\mathrm{HS}} \\ &+ \langle\nabla_{\tilde{L}_j}\Phi(s), \mathrm{e}^{-\omega_j/2}\tilde{L}_j\dot\gamma(0) - \mathrm{e}^{\omega_j/2}\dot\gamma(0)\tilde{L}_j\rangle_{\mathrm{HS}}\bigg\}. \end{aligned} \tag{12.31}$$

We first take care of the second line of Equation (12.31). Using Theorem 12.3.1, we find

$$\langle \nabla_{\tilde{L}_j} \Phi, \mathrm{e}^{-\omega_j/2} \tilde{L}_j \dot{\gamma}(0) - \mathrm{e}^{\omega_j/2} \dot{\gamma}(0) \tilde{L}_j \rangle_{\mathrm{HS}}$$

$$= -\langle \nabla_{\tilde{L}_j} \Phi, \mathrm{e}^{-\omega_j/2} \tilde{L}_j \operatorname{div}_{\tilde{\mathbf{L}}}([\rho]_{\varphi^{\log},\boldsymbol{\omega}} \nabla \Phi) - \mathrm{e}^{\omega_j/2} \operatorname{div}_{\tilde{\mathbf{L}}}([\rho]_{\varphi^{\log},\boldsymbol{\omega}} \nabla \Phi) \tilde{L}_j \rangle_{\mathrm{HS}}$$

$$= -\langle \nabla_{\tilde{L}_j} \Phi, \mathrm{e}^{-\omega_j/2} \tilde{L}_j \sum_{k \in \mathcal{J}} [[\rho]_{\varphi^{\log},\omega_k} \nabla_{\tilde{L}_k} \Phi, \tilde{L}_k^*] - \mathrm{e}^{\omega_j/2} \sum_{k \in \mathcal{J}} [[\rho]_{\varphi^{\log},\omega_k} \nabla_{\tilde{L}_k} \Phi, \tilde{L}_k^*] \tilde{L}_j \rangle_{\mathrm{HS}}$$

$$= \sum_{k \in \mathcal{J}} \left(\mathrm{e}^{-\omega_j/2} \langle \nabla_{\tilde{L}_k} (\tilde{L}_j^* \nabla_{\tilde{L}_j} \Phi), [\rho]_{\varphi^{\log},\omega_k} \nabla_{\tilde{L}_k} \Phi \rangle_{\mathrm{HS}} - \mathrm{e}^{\omega_j/2} \langle \nabla_{\tilde{L}_k} (\nabla_{\tilde{L}_j} \Phi \tilde{L}_j^*), [\rho]_{\varphi^{\log},\omega_k} \nabla_{\tilde{L}_k} \Phi \rangle_{\mathrm{HS}} \right)$$

$$= \sum_{k \in \mathcal{J}} \langle \nabla_{\tilde{L}_k} \left(\mathrm{e}^{-\omega_j/2} \tilde{L}_j^* \nabla_{\tilde{L}_j} \Phi - \mathrm{e}^{\omega_j/2} \nabla_{\tilde{L}_j} \Phi \tilde{L}_j^* \right), [\rho]_{\varphi^{\log},\omega_k} \nabla_{\tilde{L}_k} \Phi \rangle_{\mathrm{HS}}.$$

Hence by Equation (12.1),

$$\sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} \Phi, \mathrm{e}^{-\omega_j/2} \tilde{L}_j \dot{\gamma}(0) - \mathrm{e}^{\omega_j/2} \dot{\gamma}(0) \tilde{L}_j \rangle = - \sum_{k \in \mathcal{J}} \langle \nabla_{\tilde{L}_k} \mathcal{L}(\Phi), [\rho]_{\varphi^{\log},\omega_k} \nabla_{\tilde{L}_k} \Phi \rangle_{\mathrm{HS}}$$

$$= -\langle \nabla_{\tilde{\mathbf{L}}} \mathcal{L}(\Phi), \nabla_{\tilde{\mathbf{L}}} \Phi \rangle_{\mathbb{L}_{2,\log}(\rho)}. \tag{12.32}$$

By (12.30), the first line of (12.31) is equal to

$$\frac{1}{2} \sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} (\nabla_{\tilde{\mathbf{L}}} \Phi \cdot_\rho \nabla_{\tilde{\mathbf{L}}} \Phi), \mathrm{e}^{\omega_j/2} \rho \tilde{L}_j - \mathrm{e}^{-\omega_j/2} \tilde{L}_j \rho \rangle_{\mathrm{HS}} \tag{12.33}$$

$$= \frac{1}{2} \sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{\mathbf{L}}} \Phi \cdot_\rho \nabla_{\tilde{\mathbf{L}}} \Phi, [\tilde{L}_j^*, \rho \tilde{L}_j] \mathrm{e}^{\omega_j/2} - \mathrm{e}^{-\omega_j/2} [\tilde{L}_j^*, \tilde{L}_j \rho] \rangle_{\mathrm{HS}}$$

$$= \frac{1}{2} \langle \nabla_{\tilde{\mathbf{L}}} \Phi \cdot_\rho \nabla_{\tilde{\mathbf{L}}} \Phi, \mathcal{L}_*(\rho) \rangle_{\mathrm{HS}}, \tag{12.34}$$

where we used that, replacing $\tilde{L}_j$ by $\tilde{L}_j^*$ so that $\omega_j \to -\omega_j$,

$$\mathcal{L}_*(\rho) = \sum_{j \in \mathcal{J}} \left(\mathrm{e}^{\omega_j/2} [\tilde{L}_j^* \rho, \tilde{L}_j] + \mathrm{e}^{-\omega_j/2} [\tilde{L}_j, \rho \tilde{L}_j^*] \right) = \sum_{j \in \mathcal{J}} \left(\mathrm{e}^{-\omega_j/2} [\tilde{L}_j \rho, \tilde{L}_j^*] + \mathrm{e}^{\omega_j/2} [\tilde{L}_j^*, \rho \tilde{L}_j] \right),$$

Hence, using (12.32) and (12.34), (12.31) reduces to

$$\frac{d^2}{ds^2} D(\gamma(s) \| \sigma) \Big|_{s=0} = \frac{1}{2} \langle \nabla_{\tilde{\mathbf{L}}} \Phi \cdot_\rho \nabla_{\tilde{\mathbf{L}}} \Phi, \mathcal{L}_*(\rho) \rangle_{\mathrm{HS}} - \langle \nabla_{\tilde{\mathbf{L}}} \mathcal{L}(\Phi), \nabla_{\tilde{\mathbf{L}}} \Phi \rangle_{\mathbb{L}_{2,\log}(\rho)}. \tag{12.35}$$

One can compare this expression with the one derived in Proposition 4.3 of [Erbar and Maas, 2012]. To make this analogy more clear, we denote the quantity on the right hand side of Equation (12.35) by

$$B(\rho, \Phi) \equiv \frac{1}{2} \langle \nabla_{\tilde{\mathbf{L}}} \Phi \cdot_\rho \nabla_{\tilde{\mathbf{L}}} \Phi, \mathcal{L}_*(\rho) \rangle_{\mathrm{HS}} - \langle \nabla_{\tilde{\mathbf{L}}} \mathcal{L}(\Phi), \nabla_{\tilde{\mathbf{L}}} \Phi \rangle_{\mathbb{L}_{2,\log}(\rho)}. \tag{12.36}$$

The above *quantum Bochner formula* is analogous to the one derived by the formal Otto calculus in Equation (4.4), and together with Equation (12.35) justifies the interpretation of $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ as a quantum Ricci curvature lower bound. Next, the following lemma extends Lemma 4.6 of [Erbar and Maas, 2012] to the quantum regime, as well as part of the proof of Proposition 5.11 of [Carlen and Maas, 2014], and will prove to be useful in what follows:

Lemma 12.3.2. *Let $(\gamma(s))_{s \in [0,1]}$ be a smooth curve in $\mathcal{D}_+(\mathcal{H})$. For each $t \geq 0$, set $\gamma(s,t) := \mathcal{P}_{*st}(\gamma(s))$, and let $(\Phi(s,t))_{s \in [0,1]}$ be a smooth curve satisfying the continuity equation*

$$\partial_s(\gamma(s,t)) + \operatorname{div}_{\tilde{\mathbf{L}}}([\gamma(s,t)]_{\varphi^{\log},\boldsymbol{\omega}} \nabla_{\tilde{\mathbf{L}}} \Phi(s,t)) = 0, \qquad s \in [0,1]. \tag{12.37}$$

Therefore,

$$\frac{1}{2}\partial_t\, g_{\mathcal{L},\gamma(s,t)}(\partial_s\gamma(s,t),\partial_s\gamma(s,t)) + \partial_s D(\gamma(s,t)\|\sigma) = -sB(\gamma(s,t),\Phi(s,t)).$$

Proof. Start by noticing that

$$
\begin{aligned}
\partial_s D(\gamma(s,t)\|\sigma) &= \partial_s \operatorname{Tr}(\gamma(s,t)(\ln\gamma(s,t)-\ln\sigma))\\
&= \operatorname{Tr}(\partial_s\gamma(s,t)(\ln\gamma(s,t)-\ln\sigma))\\
&= -\operatorname{Tr}((\ln\gamma(s,t)-\ln\sigma)\operatorname{div}_{\tilde{\mathbf{L}}}([\gamma(s,t)]_{\varphi^{\log},\boldsymbol\omega}\nabla_{\tilde{\mathbf{L}}}\Phi(s,t)))\\
&= -\langle\ln\gamma(s,t)-\ln\sigma,\operatorname{div}_{\tilde{\mathbf{L}}}([\gamma(s,t)]_{\varphi^{\log},\boldsymbol\omega}\nabla_{\tilde{\mathbf{L}}}\Phi(s,t))\rangle_{\mathrm{HS}}\\
&= -\sum_{j\in\mathcal{J}}\langle\ln\gamma(s,t)-\ln\sigma,[[\gamma(s,t)]_{\varphi^{\log},\omega_j}(\nabla_{\tilde{L}_j}\Phi(s,t)),\tilde{L}_j^*]\rangle_{\mathrm{HS}}\\
&= \sum_{j\in\mathcal{J}}\langle\nabla_{\tilde{L}_j}(\log\gamma(s,t)-\log\sigma),[\gamma(s,t)]_{\varphi^{\log},\omega_j}(\nabla_{\tilde{L}_j}\Phi(s,t))\rangle_{\mathrm{HS}}\\
&= \sum_{j\in\mathcal{J}}\langle[\gamma(s,t)]_{\varphi^{\log},\omega_j}(\nabla_{\tilde{L}_j}(\log\gamma(s,t)-\log\sigma)),\nabla_{\tilde{L}_j}\Phi(s,t)\rangle_{\mathrm{HS}}\\
&= \sum_{j\in\mathcal{J}}\langle\nabla_{\tilde{L}_j}^*[\gamma(s,t)]_{\varphi^{\log},\omega_j}(\nabla_{\tilde{L}_j}(\log\gamma(s,t)-\log\sigma)),\Phi(s,t)\rangle_{\mathrm{HS}}\\
&= -\langle\mathcal{L}_*(\gamma(s,t)),\Phi(s,t)\rangle_{\mathrm{HS}},
\end{aligned}
$$

where in the third line we used (12.37), in last line Equation (12.24), and in the second line we used the integral representation of the logarithm to prove that $\operatorname{Tr}(\gamma(s,t)\partial_s\log\gamma(s,t))=0$. Moreover,

$$
\begin{aligned}
\frac{1}{2}\partial_t\, g_{\mathcal{L},\gamma(s,t)}(\partial_s\gamma(s,t),\partial_s\gamma(s,t)) &= \frac{1}{2}\partial_t\sum_{j\in\mathcal{J}}\langle\nabla_{\tilde{L}_j}\Phi(s,t),[\gamma(s,t)]_{\varphi^{\log},\omega_j}\nabla_{\tilde{L}_j}\Phi(s,t)\rangle_{\mathrm{HS}}\\
&= \sum_{j\in\mathcal{J}}\Big(\langle\partial_t(\nabla_{\tilde{L}_j}\Phi(s,t)),[\gamma(s,t)]_{\varphi^{\log},\omega_j}\nabla_{\tilde{L}_j}\Phi(s,t)\rangle_{\mathrm{HS}}\\
&\quad + \frac{1}{2}\langle\nabla_{\tilde{L}_j}\Phi(s,t),\partial_t([\gamma(s,t)]_{\varphi^{\log},\omega_j})\nabla_{\tilde{L}_j}\Phi(s,t)\rangle_{\mathrm{HS}}\Big). \quad(12.38)
\end{aligned}
$$

From (12.29),

$$
\sum_{j\in\mathcal{J}}\langle\nabla_{\tilde{L}_j}\Phi(s,t),\partial_t([\gamma(s,t)]_{\varphi^{\log},\omega_j})\nabla_{\tilde{L}_j}\Phi(s,t)\rangle_{\mathrm{HS}} = \langle\partial_t\gamma(s,t),\nabla\Phi(s,t)\cdot_{\gamma(s,t)}\nabla\Phi(s,t)\rangle_{\mathrm{HS}}
$$

$$
= s\langle\mathcal{L}_*(\gamma(s,t)),\nabla\Phi(s,t)\cdot_{\gamma(s,t)}\nabla\Phi(s,t)\rangle_{\mathrm{HS}}. \quad(12.39)
$$

Moreover,

$$
\sum_{j\in\mathcal{J}}\langle\partial_t(\nabla_{\tilde{L}_j}\Phi(s,t)),[\gamma(s,t)]_{\varphi^{\log},\omega_j}\nabla_{\tilde{L}_j}\Phi(s,t)\rangle_{\mathrm{HS}} \quad\quad\quad\quad\quad\quad\quad(12.40)
$$

$$
\begin{aligned}
&= -\sum_{j\in\mathcal{J}}\langle\partial_t\Phi(s,t),[[\gamma(s,t)]_{\varphi^{\log},\omega_j}(\nabla_{\tilde{L}_j}\Phi(s,t)),\tilde{L}_j^*]\rangle_{\mathrm{HS}}\\
&= -\langle\partial_t\Phi(s,t),\operatorname{div}_{\tilde{\mathbf{L}}}([\gamma(s,t)]_{\varphi^{\log},\boldsymbol\omega}(\nabla\Phi(s,t)))\rangle_{\mathrm{HS}}\\
&= \langle\partial_t\Phi(s,t),\partial_s\gamma(s,t)\rangle_{\mathrm{HS}}\\
&= \partial_t(\langle\Phi(s,t),\partial_s\gamma(s,t)\rangle_{\mathrm{HS}}) - \langle\Phi(s,t),\partial_s\partial_t\gamma(s,t)\rangle_{\mathrm{HS}}\\
&= \partial_t\, g_{\mathcal{L},\gamma(s,t)}(\partial_s\gamma(s,t),\partial_s\gamma(s,t)) - \langle\Phi(s,t),\partial_s\partial_t\gamma(s,t)\rangle_{\mathrm{HS}}\\
&= \partial_t\, g_{\mathcal{L},\gamma(s,t)}(\partial_s\gamma(s,t),\partial_s\gamma(s,t)) - \langle\Phi(s,t),\partial_s(s\mathcal{L}_*(\gamma(s,t)))\rangle_{\mathrm{HS}}, \quad(12.41)
\end{aligned}
$$

where we used once again 12.37 in the third and fifth lines above. Therefore, using 12.39 and 12.41, the right hand side of 12.38 reduces to

$$\frac{1}{2}\partial_t\, g_{\mathcal{L},\gamma(s,t)}(\partial_s\gamma(s,t),\partial_s\gamma(s,t))$$
$$= \langle \Phi(s,t), \partial_s(s\mathcal{L}_*(\gamma(s,t)))\rangle_{\mathrm{HS}} - \frac{1}{2}s\langle \mathcal{L}_*(\gamma(s,t)), \nabla\Phi(s,t)\cdot_{\gamma(s,t)}\nabla\Phi(s,t)\rangle_{\mathrm{HS}}.$$

Hence,

$$\frac{1}{2}\partial_t\, g_{\mathcal{L},\gamma(s,t)}(\partial_s\gamma(s,t),\partial_s\gamma(s,t)) + \partial_s D(\gamma(s,t)\|\sigma)$$
$$= s\,\langle \Phi(s,t), \mathcal{L}_*\partial_s(\gamma(s,t))\rangle_{\mathrm{HS}} - \frac{1}{2}s\,\langle \mathcal{L}_*\gamma(s,t).\nabla\Phi(s,t)\cdot_{\gamma(s,t)}\nabla\Phi(s,t)\rangle_{\mathrm{HS}}$$
$$= -s\,\langle \mathcal{L}(\Phi(s,t)), \mathrm{div}_{\tilde{\mathbf{L}}}([\gamma(s,t)]_{\varphi^{\log},\boldsymbol{\omega}}\nabla\Phi(s,t)))\rangle_{\mathrm{HS}} - \frac{1}{2}s\,\langle \mathcal{L}_*\gamma(s,t).\nabla\Phi(s,t)\cdot_{\gamma(s,t)}\nabla\Phi(s,t)\rangle_{\mathrm{HS}}$$
$$= s\,\langle \nabla\mathcal{L}(\Phi(s,t)), \nabla\Phi(s,t)\rangle_{\mathbb{L}_{2,\log}(\gamma(s,t))} - \frac{1}{2}s\,\langle \mathcal{L}_*(\gamma(s,t)), \nabla\Phi(s,t)\cdot_{\gamma(s,t)}\nabla\Phi(s,t)\rangle_{\mathrm{HS}}$$
$$= -s\,B(\gamma(s,t), \Phi(s,t)),$$

which is what needed to be proved. $\qquad\square$

Theorem 12.3.3. *Let $\mathcal{L}$ be the generator of a primitive QMS $(\mathcal{P}_t)_{t\geq 0}$, with unique invariant state σ, of the form of Equation (12.1). Then, for $\kappa \in \mathbb{R}$, the following are equivalent:*

(i) *Ricci curvature lower bound: for any constant speed geodesic $(\gamma(s))_{s\in(-\varepsilon,\varepsilon)}$ on $\mathcal{D}_+(\mathcal{H})$ such that $\gamma(0) = \rho$:*

$$\left.\frac{d^2}{ds^2}\right|_{s=0} D(\gamma(s)\|\sigma) \geq \kappa\, g_{\mathcal{L},\rho}(\dot\gamma(0),\dot\gamma(0)), \qquad\qquad (\mathrm{Ric}(\mathcal{L})\geq\kappa)$$

(ii) *For all $\rho \in \mathcal{D}_+(\mathcal{H})$, and $\Phi \in \mathcal{W}$,*

$$B(\rho,\Phi) \geq \kappa\|\nabla\Phi\|^2_{\mathbb{L}_{2,\log}(\rho)}$$

(iii) *Evolution variational inequality: for all $\rho,\omega \in \mathcal{D}_+(\mathcal{H})$ and all $t \geq 0$, writing $\rho_t := \mathcal{P}_{*t}(\rho)$:*

$$\frac{1}{2}\left.\frac{d}{dt}\right|_{t^+} (W_{2,\mathcal{L}}(\rho_t,\omega))^2 + \frac{\kappa}{2} W_{2,\mathcal{L}}(\rho_t,\omega)^2 \leq D(\omega\|\sigma) - D(\rho_t\|\sigma). \qquad (12.42)$$

(iv) *Equation (12.42) holds for any $\rho,\omega \in \mathcal{D}(\mathcal{H})$.*

(v) *κ-displacement convexity of the relative entropy: for any constant speed geodesic $(\gamma(s))_{s\in[0,1]}$ in $\mathcal{D}(\mathcal{H})$,*

$$D(\gamma(s)\|\sigma) \leq (1-s)D(\gamma(0)\|\sigma) + sD(\gamma(1)\|\sigma) - \frac{\kappa}{2}s(1-s)\,W_{2,\mathcal{L}}(\gamma(0),\gamma(1))^2. \qquad (12.43)$$

Proof. The proof is inspired by the one of Theorem 4.5 of [Erbar and Maas, 2012]. That $(i) \Leftrightarrow (ii)$ follows from Equation (12.36). We use Lemma 12.3.2 to show that $(ii) \Rightarrow (iii)$: Take a smooth path $(\gamma(s))_{s\in[0,1]}$ such that $\gamma(0) = \omega$, $\gamma(1) = \rho$ and

$$\int_0^1 g_{\mathcal{L},\gamma(s)}(\dot\gamma(s),\dot\gamma(s))\,ds \leq W_{2,\mathcal{L}}(\rho,\omega)^2 + \varepsilon. \qquad (12.44)$$

With the notations of Lemma 12.3.2,

$$\frac{1}{2}\partial_t\left(e^{2\kappa st}\,g_{\mathcal{L},\gamma(s,t)}(\partial_s\gamma(s,t),\partial_s\gamma(s,t))\right)+\partial_s\left(e^{2\kappa st}\,D(\gamma(s,t)\|\sigma)\right)\le 2\,\kappa t\,e^{2\kappa st}\,D(\gamma(s,t)\|\sigma)\,.$$

Integrating with respect to $t\in[0,h]$, for some $h>0$, and $s\in[0,1]$,

$$2\kappa\int_0^1 ds\int_0^h dt\,t\,e^{2\kappa st}\,D(\gamma(s,t)\|\sigma)\ge\int_0^h\left(e^{2\kappa t}\,D(\gamma(1,t)\|\sigma)-D(\gamma(0,t)\|\sigma)\right)dt$$
$$+\frac{1}{2}\int_0^1 e^{2\kappa sh}\,g_{\mathcal{L},\gamma(s,h)}(\partial_s\gamma(s,h),\partial_s\gamma(s,h))-g_{\mathcal{L},\gamma(s,0)}(\partial_s\gamma(s,0),\partial_s\gamma(s,0))\,ds\,.$$
$$(12.45)$$

The following inequality, for which a classical equivalent is given in the proof of Theorem 4.5 of [Erbar and Maas, 2012], can be derived similarly to Lemma 5.1 of [Daneri and Savaré, 2008]:

$$m(\kappa h)W_{2,\mathcal{L}}(\rho_h,\omega)^2\le\int_0^1 e^{2\kappa sh}\,g_{\mathcal{L},\gamma(s,h)}(\partial_s\gamma(s,h),\partial_s\gamma(s,h))\,ds,\qquad(12.46)$$

where $m(x):=x\,e^x/\sinh(x)$. Indeed, define $f:s\mapsto e^{2\kappa sh}$, and denote $L_f:=\int_0^1\frac{1}{f(s)}ds$. Then, let $g:[0,1]\mapsto[0,1]$ be the smooth increasing map defined as $g(s)=L_f^{-1}\int_0^s\frac{1}{f(u)}du$, and denote its inverse k such that $k'(g(s))=L_f f(s)$. Then define the reparametrized curve $(\gamma(k(r),h),k'(r)\,\Phi(k(r),h))_{r\in[0,1]}$ which satisfies the continuity equation:

$$\partial_r\gamma(k(r),h)=k'(r)\partial_1\gamma(k(r),h)$$
$$=-k'(r)\operatorname{div}_{\tilde{\mathbf{L}}}\left([\gamma(k(r),h)]_{\varphi^{\log},\boldsymbol{\omega}}\nabla\Phi(k(r),h)\right),$$

where we used Equation (12.37) in order to established the second line. This curve satisfies $\gamma(k(0),h)=\omega$ and $\gamma(k(1),h)=\rho_h$, so that

$$W_{2,\mathcal{L}}(\rho_h,\omega)^2\le\int_0^1 g_{\mathcal{L},\gamma(k(r),h)}(\partial_r\gamma(k(r),h),\partial_r\gamma(k(r),h))\,dr$$
$$=\int_0^1 k'(r)^2\|\nabla_{\tilde{\mathbf{L}}}\Phi(k(r),h)\|_{\mathbb{L}_{2,\log}(\gamma(k(r),h))}^2 dr$$
$$=\int_0^1 k'(g(s))\|\nabla_{\tilde{\mathbf{L}}}\Phi(s,h)\|_{\mathbb{L}_{2,\log}(\gamma(s,h))}^2 ds$$
$$=L_f\int_0^1 f(s)\,g_{\mathcal{L},\gamma(s,h)}(\partial_s\gamma(s,h),\partial_s\gamma(s,h))\,ds\,,$$

which directly leads to 12.46. This inequality, together with 12.44, implies

$$\frac{m(h\kappa)}{2}W_{2,\mathcal{L}}(\rho_h,\omega)^2-\frac{1}{2}W_{2,\mathcal{L}}(\rho,\omega)^2-\varepsilon+\int_0^h e^{2\kappa t}\,dt\,D(\rho_h\|\sigma)-hD(\omega\|\sigma)$$
$$\le\frac{1}{2}\int_0^1 e^{2\kappa sh}\,g_{\mathcal{L},\gamma(s,t)}(\partial_s\gamma(s,t),\partial_s\gamma(s,t))\,ds-\frac{1}{2}\int_0^1 g_{\mathcal{L},\gamma(s)}(\dot\gamma(s),\dot\gamma(s))\,ds$$
$$+\int_0^h e^{2\kappa t}\,D(\rho_t\|\sigma)\,dt-hD(\omega\|\sigma)$$
$$\le 2\kappa\int_0^1 ds\int_0^h dt\,t\,e^{2\kappa st}\,D(\gamma(s,t)\|\sigma)\,.$$

where, in the first inequality, we also used the monotonicity of the relative entropy so that $D(\rho_h\|\sigma)=D(\rho_h\|\mathcal{P}_{*h}\sigma)\le D(\rho_t\|\mathcal{P}_{*t}\sigma)=D(\rho_t\|\sigma)$, and in the second one that for all $t>0$, $\gamma(1,t)=\rho_t$, $\gamma(0,t)=\omega$,

as well as 12.45. Since for all $s \in [0,1]$, $t \mapsto D(\gamma(s,t)\|\sigma)$ is bounded,

$$\lim_{h\to 0} \frac{1}{h} \int_0^1 \int_0^h t\, \mathrm{e}^{2\kappa st}\, D(\gamma(s,t)\|\sigma)dt\, ds = 0.$$

Moreover,

$$\lim_{h\to 0} \frac{1}{h} \left(\int_0^h \mathrm{e}^{2\kappa t}\, dt\, D(\rho_h\|\sigma) - hD(\omega\|\sigma) \right) = D(\rho\|\sigma) - D(\omega\|\sigma)$$

Since $\varepsilon > 0$ was arbitrary, we arrive at

$$\frac{d}{dh}\Big|_{h=0^+} \left(\frac{m(\kappa h)}{2} W_{2,\mathcal{L}}(\rho_h, \omega)^2 \right) + D(\rho\|\sigma) - D(\omega\|\sigma) \le 0.$$

The result for $t = 0$ follows from the fact that the first term in the left hand side above is equal to $\frac{\kappa}{2} W_{2,\mathcal{L}}(\rho_h, \omega)^2 + \frac{1}{2} \frac{d}{dh}\big|_{h=0^+} W_{2,\mathcal{L}}(\rho_h, \omega)^2$. The case $t \ge 0$ directly follows from the case $t = 0$.

$(iii) \Rightarrow (iv)$ follows from Theorem 3.3 of [Daneri and Savaré, 2008] together with the fact that $(\mathcal{D}(\mathcal{H}), W_{2,\mathcal{L}})$ is complete (cf. Proposition 12.1.8).

$(iv) \Rightarrow (v)$ follows directly from Theorem 3.2 of [Daneri and Savaré, 2008].

$(v) \Rightarrow (i)$ can easily be proved as follows: let $0 < \varepsilon < \varepsilon'$, and without loss of generality, let $\gamma : (-\varepsilon', \varepsilon') \to \mathcal{D}_+(\mathcal{H})$ be speed 1 geodesic, and that $\gamma(0) = \rho$. Then, construct the following constant speed geodesic $\tilde{\gamma} : [0,1] \to \mathcal{D}_+(\mathcal{H})$ as follows: for any $s \in [0,1]$, $\tilde{\gamma}(s) := \gamma(2\varepsilon s - \varepsilon)$. It then follows that $W_{2,\mathcal{L}}(\tilde{\gamma}(0), \tilde{\gamma}(1)) = 2\varepsilon$. Moreover, by applying (12.43) to $\tilde{\gamma}$, we find, after a suitable rearrangement of the terms:

$$\frac{D(\gamma(\varepsilon)\|\sigma) - 2D(\rho\|\sigma) + D(\gamma(-\varepsilon)\|\sigma)}{\varepsilon^2} \ge \kappa.$$

The result follows after taking the limit $\varepsilon \to 0$. $\qquad\square$

Other equivalent formulations of displacement convexity Here, we provide other characterizations of the Ricci curvature lower bound in terms of some contraction properties of the Wasserstein metric along the semigroup $(\mathcal{P}_t)_{t\ge 0}$. In the next theorem, the characterization of displacement convexity in terms of gradient estimates can be interpreted as a non-commutative version of Bakry-Émery's original gradient bound (see Theorem 4.7.2 of [Bakry et al., 2014]): in particular they showed that the Ricci curvature lower bound is equivalent to the following pointwise inequality for smooth enough functions (e.g. functions satisfying Condition 2.2.1):

$$\Gamma(P_t(f), P_t(f)) \le \mathrm{e}^{-2\kappa t} P_t(\Gamma(f, f)), \tag{12.47}$$

where Γ stands for the carré du champ operator introduced in Section 2.2.

Proposition 12.3.4 (Gradient estimate)**.** *For any $\kappa \in \mathbb{R}$, $\mathrm{Ric}(\mathcal{L}) \ge \kappa$ is equivalent to the following gradient estimate: for any $\rho \in \mathcal{D}_+(\mathcal{H})$, any traceless $\Phi \in \mathcal{B}_{sa}(\mathcal{H})$ and all $t > 0$:*

$$\|\nabla_{\tilde{\mathbf{L}}}(\mathcal{P}_t(\Phi))\|^2_{\mathbb{L}_{2,\log}(\rho)} \le \mathrm{e}^{-2\kappa t} \|\nabla_{\tilde{\mathbf{L}}}\Phi\|^2_{\mathbb{L}_{2,\log}(\mathcal{P}_{*t}(\rho))}. \tag{12.48}$$

Proof. We first prove the direct implication: define for $u \in [0,t]$ $\rho_u \equiv \mathcal{P}_{*u}(\rho)$ and $\Phi(u) \equiv \mathcal{P}_u(\Phi)$. Next, define

$$\psi(s) := e^{-2\kappa s} \|\nabla_{\tilde{\mathbf{L}}} \Phi(t-s)\|^2_{\mathbb{L}_{2,\log}(\rho_s)} \equiv e^{-2\kappa s} \sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j}(\Phi(t-s)), [\rho_s]_{\varphi^{\log},\omega_j} \nabla_{\tilde{L}_j}(\Phi(t-s)) \rangle_{\mathrm{HS}}.$$

Then, $\psi(0) = \|\nabla_{\tilde{\mathbf{L}}} \Phi(t)\|^2_{\mathbb{L}_{2,\log}(\rho)}$ and $\psi(t) = e^{-2\kappa t} \|\nabla_{\tilde{\mathbf{L}}} \Phi\|^2_{\mathbb{L}_{2,\log}(\rho_t)}$. It is then enough to prove that ψ has non-negative derivative to prove the claim. But:

$$\psi'(s) = 2 e^{-2\kappa s} \left[-\kappa \|\nabla_{\tilde{\mathbf{L}}} \Phi(t-s)\|^2_{\mathbb{L}_{2,\log}(\rho_s)} + \frac{1}{2} \frac{\partial}{\partial s} \|\nabla_{\tilde{\mathbf{L}}} \Phi(t-s)\|^2_{\mathbb{L}_{2,\log}(\rho_s)} \right]$$

$$= 2 e^{-2\kappa s} \left[-\kappa \|\nabla_{\tilde{\mathbf{L}}} \Phi(t-s)\|^2_{\mathbb{L}_{2,\log}(\rho_s)} + \sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} \partial_s \Phi(t-s), [\rho_s]_{\varphi^{\log},\omega_j} \nabla_{\tilde{L}_j} \Phi(t-s) \rangle_{\mathrm{HS}} \right.$$

$$\left. + \frac{1}{2} \sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} \Phi(t-s), \partial_s([\rho_s]_{\varphi^{\log},\omega_j}) \nabla_{\tilde{L}_j} \Phi(t-s) \rangle_{\mathrm{HS}} \right]$$

$$= 2 e^{-2\kappa s} \left[-\kappa \|\nabla_{\tilde{\mathbf{L}}} \Phi(t-s)\|^2_{\mathbb{L}_{2,\log}(\rho_s)} - \langle \nabla_{\tilde{\mathbf{L}}} \mathcal{L}(\Phi(t-s)), \nabla_{\tilde{\mathbf{L}}} \Phi(t-s) \rangle_{\mathcal{L},\rho_s} \right.$$

$$\left. + \frac{1}{2} \langle \mathcal{L}_*(\rho_s), \nabla_{\tilde{\mathbf{L}}} \Phi(t-s) \cdot_{\rho_s} \nabla_{\tilde{\mathbf{L}}} \Phi(t-s) \rangle_{\mathrm{HS}} \right]$$

$$= 2 e^{-2\kappa s} \left[-\kappa \|\nabla_{\tilde{\mathbf{L}}} \Phi(t-s)\|^2_{\mathbb{L}_{2,\log}(\rho_s)} + B(\rho_s, \Phi(t-s)) \right]$$

where we used Equation (12.29) in the second line. We conclude by a use of (ii) of Theorem 12.3.3. For the reverse implication, assume that Equation (12.48) holds. Then,

$$0 \leq e^{-2\kappa t} \|\nabla_{\tilde{\mathbf{L}}} \Phi\|^2_{\mathbb{L}_{2,\log}(\mathcal{P}_{*t}(\rho))} - \|\nabla_{\tilde{\mathbf{L}}}(\mathcal{P}_t(\Phi))\|^2_{\mathbb{L}_{2,\log}(\rho)}$$

$$= (e^{-2\kappa t} - 1) \|\nabla_{\tilde{\mathbf{L}}} \Phi\|^2_{\mathbb{L}_{2,\log}(\rho)} + e^{-2\kappa t} (\|\nabla_{\tilde{\mathbf{L}}} \Phi\|^2_{\mathbb{L}_{2,\log}(\mathcal{P}_{*t}(\rho))} - \|\nabla_{\tilde{\mathbf{L}}} \Phi\|^2_{\mathbb{L}_{2,\log}(\rho)})$$

$$- \|\nabla_{\tilde{\mathbf{L}}}(\mathcal{P}_t(\Phi))\|^2_{\mathbb{L}_{2,\log}(\rho)} + \|\nabla_{\tilde{\mathbf{L}}} \Phi\|^2_{\mathbb{L}_{2,\log}(\rho)}.$$

By dividing by t and letting $t \to 0$, we once again obtain that $-\kappa \|\nabla_{\tilde{\mathbf{L}}} \Phi\|^2_{\mathbb{L}_{2,\log}(\rho)} + B(\rho, \Phi) \geq 0$, which allows us to conclude from Equation (12.29). $\square$

In the commutative smooth setting, the Ricci curvature lower bound is also known to be equivalent to the contraction of the Wasserstein distance along the semigroup $(P_t)_{t \geq 0}$ (see Theorem 9.7.2 of [Bakry et al., 2014]):

$$W_2(P_{t*}(\nu), P_{t*}(\nu')) \leq e^{-\kappa t} W_2(\nu, \nu').$$

This still holds true in the non-commutative, finite dimensional setting:

Proposition 12.3.5. *For any $\kappa \in \mathbb{R}$, $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ is equivalent to the contraction of the Wasserstein distance along the flow generated by $(\mathcal{P}_t)_{t \geq 0}$: for any $\rho, \omega \in \mathcal{D}_+(\mathcal{H})$*

$$W_{2,\mathcal{L}}(\mathcal{P}_{t*}(\rho), \mathcal{P}_{t*}(\omega)) \leq e^{-\kappa t} W_{2,\mathcal{L}}(\rho, \omega). \tag{12.49}$$

Proof. The direct implication follows from Proposition 3.1 of [Daneri and Savaré, 2008], which holds in great generality for gradient flows on metric spaces, and Theorem 12.3.3(iii). The reverse implication is proved as in inequality (2.12) of [Daneri and Savaré, 2008], using the smooth Riemannian structure provided by $(\mathcal{D}_+(\mathcal{H}), W_{2,\mathcal{L}})$ in the finite dimensional case. $\square$

In the commutative diffusive setting, the contraction of (12.47) is actually known to be equivalent

to its "square root" version, usually referred to as the *strong gradient bound*:

$$\sqrt{\Gamma(P_t(f), P_t(f))} \le e^{-\kappa t} P_t(\sqrt{\Gamma(f, f)}) . \tag{12.50}$$

The proof of (12.50)⇒(12.47) follows by a simple use of Jensen's inequality, the converse being the content of Theorem 3.3.18 of [Bakry et al., 2014]. The advantage of this formulation arises from the fact that some canonical semigroups (e.g. the quantum Ornstein Uhlenbeck semigroup on $\mathbb{R}^n$) saturate the inequality, or equivalently:

$$[L, \nabla] = \kappa \nabla .$$

Therefore, the Ricci lower bound is equivalent to comparing the commutation of a semigroup with the gradient to the one of a canonical semigroup. This is similar in spirit to the geometric inequalities briefly mentioned in Chapter 11 that involve the comparison of the curvature of a space to the one of a model space of constant curvature (e.g. the sphere). A similar reasoning lead [Johnson, 2017] to formulate a Bakry-Émery condition for birth and death processes on $\mathbb{N}$ in terms of a comparison to the Poisson process. Going back to our non-commutative setting, [Carlen and Maas, 2017] showed that the quantum Ornstein-Uhlenbeck semigroup, as well as its fermionic version on the Clifford algebra, do satisfy such a commutation relation. They used this fact to derive the modified logarithmic Sobolev constant for these QMS via the contraction (12.49). In the next proposition, we recall their argument:

Proposition 12.3.6. *Assume that the following equalities hold: there exists $\kappa \in \mathbb{R}$ such that, for any $j \in \mathcal{J}$ and any $t \ge 0$,*

$$\nabla_{\tilde{L}_j} \circ \mathcal{P}_t = e^{-\kappa t} \mathcal{P}_t \circ \nabla_{\tilde{L}_j} . \tag{12.51}$$

Then, $\mathrm{Ric}(\mathcal{L}) \ge \kappa$ holds.

Proof. From Proposition 12.3.5, it is enough to prove that (12.49) holds. Assume that $(\gamma(s))_{s \in [0,1]}$ is a minimal geodesic relating ρ to σ and denote by $(\mathbf{A}(s))_{s \in [0,1]}$ the unique solution of the continuity equation

$$\dot{\gamma}(s) = \mathrm{div}\, \mathbf{A}(s) .$$

By duality, Equation (12.51) implies that $\mathcal{P}_{t*}\dot{\gamma}(s) = e^{-\kappa t} \mathrm{div}\, \vec{\mathcal{P}}_{t*}\mathbf{A}(s)$, where $\vec{\mathcal{P}}_{t*}\mathbf{A} := (\mathcal{P}_{t*}A_j)_{j \in \mathcal{J}}$. Then, denoting $\gamma(s, t) := \mathcal{P}_{t*}(\gamma(s))$,

$$\begin{aligned}
g_{\mathcal{L}, \gamma(s,t)}(\dot{\gamma}(s,t), \dot{\gamma}(s,t)) &= e^{-2\kappa t} \sum_{j \in \mathcal{J}} \langle \mathcal{P}_{*t}(A_j(s)), [\gamma(s,t)]_{\varphi^{\log}, \omega_j}^{-1}(\mathcal{P}_{t*}(A_j(s))) \rangle_{\mathrm{HS}} \\
&\le e^{-2\kappa t} \sum_{j \in \mathcal{J}} \langle A_j(s), [\gamma(s)]_{\varphi^{\log}, \omega_j}^{-1} A_j(s) \rangle_{\mathrm{HS}} \\
&= e^{-2\kappa t} g_{\mathcal{L}, \gamma(s)}(\dot{\gamma}(s), \dot{\gamma}(s)) ,
\end{aligned}$$

where the inequality arises from the property of monotonicity of Fisher information metrics (cf. Theorem 1.3.3). The result follows after taking the integral over the geodesic path. □

12.4. Example: the quantum depolarizing semigroup

In this section, we derive a Ricci curvature lower bound for the depolarizing semigroup (cf. Section 5.5.1):

Theorem 12.4.1. *The quantum depolarizing semigroup* $(\mathcal{P}_t^{\mathrm{depol}})_{t\geq 0}$ *satisfies* $\mathrm{Ric}(\mathcal{L}^{\mathrm{depol}}) \geq \frac{1}{2}$.

Proof. In the Schrödinger picture, the generator $\mathcal{L}_*^{\mathrm{depol}}$ can be written as

$$\mathcal{L}_*^{\mathrm{depol}}(\rho) = \frac{\mathbb{1}}{d_{\mathcal{H}}} - \rho = \frac{1}{d_{\mathcal{H}}^2}\sum_{j=1}^{d_{\mathcal{H}}^2}(U_j\,\rho\,U_j^* - \rho), \qquad \rho \in \mathcal{D}(\mathcal{H}),$$

where the operators U_j can be chosen to be unitary (e.g. generalized Pauli matrices [Wolf, 2012]). In this case, $\tilde{L}_j = (\sqrt{2}d_{\mathcal{H}})^{-1}\,U_j$ and $\omega_j = 0$, $j = 1, ..., d_{\mathcal{H}}^2$. Now, given a vector $\vec{V} = (V_1, ..., V_{d_{\mathcal{H}}^2}) \in \bigoplus_j \mathcal{B}(\mathcal{H})$ and $\rho \in \mathcal{D}_+(\mathcal{H})$,

$$\langle \vec{V}, \vec{V}\rangle_{\mathbb{L}_{2,\varphi^{\log}}(\rho)} = \sum_{j=1}^{d_{\mathcal{H}}^2}\langle V_j, [\rho]_{\varphi^{\log}}(V_j)\rangle_{\mathrm{HS}} \equiv \sum_{j=1}^{d_{\mathcal{H}}^2}\int_0^1 \mathrm{Tr}(V_j^*\,\rho^s\,V_j\rho^{1-s})\,ds$$

Now, given $\Phi \in \mathcal{B}_{sa}(\mathcal{H})$ and $\rho \in \mathcal{D}_+(\mathcal{H})$, the following holds:

$$\nabla_{\tilde{\mathbf{L}}}\Phi\cdot_\rho\nabla_{\tilde{\mathbf{L}}}\Phi = \sum_{j=1}^{d_{\mathcal{H}}^2}\int_0^1\int_0^1\int_0^\alpha \frac{\rho^\beta}{(1-s)\mathbb{1}+s\rho}\nabla_{\tilde{L}_j}\Phi\rho^{1-\alpha}(\nabla_{\tilde{L}_j}\Phi)^*\frac{\rho^{\alpha-\beta}}{(1-s)\mathbb{1}+s\rho}\,d\beta\,ds\,d\alpha$$
$$+ \int_0^1\int_0^1\int_0^{1-\alpha}\frac{\rho^\beta}{(1-s)\mathbb{1}+s\rho}(\nabla_{\tilde{L}_j}\Phi)^*\rho^\alpha\nabla_{\tilde{L}_j}\Phi\frac{\rho^{1-\alpha-\beta}}{(1-s)\mathbb{1}+s\rho}\,d\beta\,ds\,d\alpha,$$

Then:

$$\frac{1}{2}\langle \nabla_{\tilde{\mathbf{L}}}\Phi\cdot_\rho\nabla_{\tilde{\mathbf{L}}}\Phi, \mathcal{L}_*^{\mathrm{depol}}(\rho)\rangle_{\mathrm{HS}}$$

$$= \frac{1}{2}\sum_{j=1}^{d_{\mathcal{H}}^2}\int_0^1 ds\int_0^1 d\alpha\left\{\int_0^\alpha \mathrm{Tr}\left[\left(\frac{\mathbb{1}}{d_{\mathcal{H}}}-\rho\right)\frac{\rho^\beta}{(1-s)\mathbb{1}+s\rho}\nabla_{\tilde{L}_j}\Phi\rho^{1-\alpha}(\nabla_{\tilde{L}_j}\Phi)^*\frac{\rho^{\alpha-\beta}}{(1-s)\mathbb{1}+s\rho}\right]d\beta\right.$$
$$\left.+\int_0^{1-\alpha}\mathrm{Tr}\left[\left(\frac{\mathbb{1}}{d_{\mathcal{H}}}-\rho\right)\frac{\rho^\beta}{(1-s)\mathbb{1}+s\rho}(\nabla_{\tilde{L}_j}\Phi)^*\rho^\alpha\nabla_{\tilde{L}_j}\Phi\frac{\rho^{1-\alpha-\beta}}{(1-s)\mathbb{1}+s\rho}\right]d\beta\right\}.$$

By cyclicity of the trace, and since $\int_0^1\frac{1}{((1-s)\mathbb{1}+s\rho)^2}ds = \rho^{-1}$, forgetting about the positive contributions coming from the terms in $d_{\mathcal{H}}^{-1}\mathbb{1}$ the above expression can be lower bounded as follows:

$$\frac{1}{2}\langle \nabla_{\tilde{\mathbf{L}}}\Phi\cdot_\rho\nabla_{\tilde{\mathbf{L}}}\Phi, \mathcal{L}_*^{\mathrm{depol}}(\rho)\rangle_{\mathrm{HS}}$$

$$\geq -\frac{1}{2}\sum_{j=1}^{d_{\mathcal{H}}^2}\int_0^1\left\{\int_0^\alpha \mathrm{Tr}(\rho^\beta\nabla_{\tilde{L}_j}\Phi\rho^{1-\alpha}(\nabla_{\tilde{L}_j}\Phi)^*\rho^{\alpha-\beta}) + \int_0^{1-\alpha}\mathrm{Tr}(\rho^\beta(\nabla_{\tilde{L}_j}\Phi)^*\rho^\alpha\nabla_{\tilde{L}_j}\Phi\rho^{1-\alpha-\beta})d\beta\right\}d\alpha$$

$$= -\frac{1}{2}\sum_{j=1}^{d_{\mathcal{H}}^2}\int_0^1\left\{\alpha\,\mathrm{Tr}(\nabla_{\tilde{L}_j}\Phi\rho^{1-\alpha}(\nabla_{\tilde{L}_j}\Phi)^*\rho^\alpha) + (1-\alpha)\,\mathrm{Tr}((\nabla_{\tilde{L}_j}\Phi)^*\rho^\alpha\nabla_{\tilde{L}_j}\Phi\rho^{1-\alpha})\right\}d\alpha$$

$$= -\frac{1}{2}\sum_{j=1}^{d_{\mathcal{H}}^2}\int_0^1\mathrm{Tr}((\nabla_{\tilde{L}_j}\Phi)^*\rho^\alpha\nabla_{\tilde{L}_j}\Phi\rho^{1-\alpha})\,d\alpha.$$

On the other hand,

$$\langle \nabla_{\tilde{\mathbf{L}}}\mathcal{L}^{\mathrm{depol}}(\Phi), \nabla_{\tilde{\mathbf{L}}}\Phi\rangle_{\mathbb{L}_{2,\varphi^{\log}}(\rho)} = \sum_{j=1}^{d_{\mathcal{H}}^2}\int_0^1\mathrm{Tr}\left[\left\{\nabla_{\tilde{L}_j}\left(\frac{\mathbb{1}}{d_{\mathcal{H}}}\mathrm{Tr}(\Phi)-\Phi\right)\right\}^*\rho^s\,\nabla_{\tilde{L}_j}\Phi\rho^{1-s}\right]ds$$

$$= -\sum_{j=1}^{d_{\mathcal{H}}^2}\int_0^1\mathrm{Tr}\left[(\nabla_{\tilde{L}_j}\Phi)^*\rho^s\nabla_{\tilde{L}_j}\Phi\rho^{1-s}\right]ds$$

Therefore,

$$B(\rho, \Phi) \geq \frac{1}{2} \sum_{j=1}^{d_{\mathcal{H}}^2} \int_0^1 \operatorname{Tr}\left[(\nabla_{\tilde{L}_j} \Phi)^* \rho^s \nabla_{\tilde{L}_j} \Phi \, \rho^{1-s}\right] ds = \frac{1}{2} \|\nabla_{\tilde{\mathbf{L}}} \Phi\|_{\mathbb{L}_{2,\log}(\rho)}^2,$$

and the result follows. □

12.5. Link to quantum functional inequalities

In this section, we relate the Ricci curvature lower bound to other quantum functional inequalities. In Section 12.5, we show that the above displacement convexity implies a quantum generalization of the HWI inequality of [Erbar and Maas, 2012], which itself implies MLSI in the case of a positive curvature. The quantum transportation cost inequality of order 2 is introduced, and its link to MLSI and PI is strengthened in Section 12.5.2.

12.5.1. The quantum HWI inequality

In [Erbar and Maas, 2012], the authors proved that, in the classical discrete setting, $\operatorname{Ric}(\mathcal{L}) \geq \kappa$ for $\kappa \in \mathbb{R}$ implies an HWI-like inequality (cf. Theorem 4.3.1). Here, we provide a quantum generalization of their result and relate it to the modified logarithmic Sobolev inequality. First of all, we need the following lemma that will also prove its usefulness in Section 12.7:

Lemma 12.5.1. *Let $\rho, \tau \in \mathcal{D}_+(\mathcal{H})$. Then for all $t > 0$, $\rho_t \equiv \mathcal{P}_{*t}(\rho)$ satisfies*

$$\frac{d}{dt} W_{2,\mathcal{L}}(\rho_t, \tau) \leq \sqrt{\operatorname{EP}_\sigma(\rho_t)}.$$

Proof. We proceed similarly to the proof of Proposition 7.1 of [Erbar and Maas, 2012]: Firstly, by the triangle inequality,

$$\frac{d}{dt} W_{2,\mathcal{L}}(\rho_t, \tau) = \lim_{s \to 0} \frac{1}{s}(W_{2,\mathcal{L}}(\rho_{t+s}, \tau) - W_{2,\mathcal{L}}(\rho_t, \tau)) \leq \lim_{s \to 0} \frac{1}{s} W_{2,\mathcal{L}}(\rho_t, \rho_{t+s}), \tag{12.52}$$

Now, by Lemma 12.1.4,

$$W_{2,\mathcal{L}}(\rho_t, \rho_{t+s}) = \inf_{\gamma(s)} \left\{ \int_0^1 g_{\mathcal{L},\gamma(u)}(\dot{\gamma}(u), \dot{\gamma}(u)) \, du : \ \gamma(0) = \rho_t, \ \gamma(1) = \rho_{t+s} \right\},$$

This implies by a change of variable $v = t + us$ that for any smooth curve γ such that $\gamma(t) = \rho_t$ and $\gamma(t+s) = \rho_{t+s}$,

$$W_{2,\mathcal{L}}(\rho_t, \rho_{t+s}) \leq \int_t^{t+s} g_{\mathcal{L},\gamma(v)}(\dot{\gamma}(v), \dot{\gamma}(v)) \, dv. \tag{12.53}$$

Moreover, from Equation (12.20):

$$g_{\mathcal{L},\rho_t}(\dot{\rho}_t, \dot{\rho}_t) = -\frac{d}{dt} D(\rho_t \| \sigma) = \operatorname{EP}_\sigma(\rho_t), \tag{12.54}$$

where the second identity holds by Theorem 7.3.1. Hence, choosing $\gamma(v) = \rho_v$, we bound the right hand side of (12.52) as follows:

$$\frac{d}{dt} W_{2,\mathcal{L}}(\rho_t, \tau) \leq \lim_{s \to 0} \frac{1}{s} \int_t^{t+s} \sqrt{\operatorname{EP}_\sigma(\rho_v)} \, dv = \sqrt{\operatorname{EP}_\sigma(\rho_t)},$$

where the last equality holds since $t \to \sqrt{\mathrm{EP}_\sigma(\rho_t)}$ is continuous. $\square$

Theorem 12.5.2. *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq \kappa$, *for some* $\kappa \in \mathbb{R}$. *Then* $\mathcal{L}$ *satisfies the following inequality*

$$\forall \rho \in \mathcal{D}_+(\mathcal{H}), \qquad D(\rho\|\sigma) \leq W_{2,\mathcal{L}}(\rho,\sigma)\sqrt{\mathrm{EP}_\sigma(\rho)} - \frac{\kappa}{2}W_{2,\mathcal{L}}(\rho,\sigma)^2. \qquad (\mathrm{HWI}(\kappa))$$

Proof. By Theorem 12.3.3, for any $\rho, \omega \in \mathcal{D}_+(\mathcal{H})$

$$\frac{1}{2}\frac{d}{dt}\bigg|_{t=0^+} W_{2,\mathcal{L}}(\rho_t,\omega)^2 + \frac{\kappa}{2}W_{2,\mathcal{L}}(\rho,\omega)^2 \leq D(\omega\|\sigma) - D(\rho\|\sigma).$$

Taking $\omega := \sigma$, this implies that

$$D(\rho\|\sigma) \leq -\frac{1}{2}\frac{d}{dt}\bigg|_{t=0^+} W_{2,\mathcal{L}}(\rho_t\|\sigma)^2 - \frac{\kappa}{2}W_{2,\mathcal{L}}(\rho,\sigma)^2. \qquad (12.55)$$

Then,

$$
\begin{aligned}
-\frac{1}{2}\frac{d}{dt}\bigg|_{t=0^+} W_{2,\mathcal{L}}(\rho_t,\sigma)^2 &= \liminf_{s\to 0^+} \frac{1}{2s}\left(W_{2,\mathcal{L}}(\rho,\sigma)^2 - W_{2,\mathcal{L}}(\rho_s,\sigma)^2\right) \\
&\leq \limsup_{s\to 0^+} \frac{1}{2s}\left(W_{2,\mathcal{L}}(\rho,\rho_s)^2 + 2W_{2,\mathcal{L}}(\rho,\rho_s)W_{2,\mathcal{L}}(\rho_s,\sigma)\right) \\
&\leq \limsup_{s\to 0^+} \frac{1}{2s}W_{2,\mathcal{L}}(\rho,\rho_s)^2 + W_{2,\mathcal{L}}(\rho,\sigma)\sqrt{\mathrm{EP}_\sigma(\rho)} \\
&= W_{2,\mathcal{L}}(\rho,\sigma)\sqrt{\mathrm{EP}_\sigma(\rho)}.
\end{aligned}
$$

where the second inequality follows from Lemma 12.5.1. The result follows from inserting this back into (12.55). $\square$

In the case when $\kappa > 0$, we recover the modified logarithmic Sobolev inequality:

Corollary 12.5.3 (Quantum Bakry-Émery theorem). *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq \kappa$, *for some* $\kappa > 0$. *Then* $(\mathcal{P}_t)_{t\geq 0}$ *satisfies* $\mathrm{MLSI}(\alpha_1)$ *with* $\alpha_1 = \frac{\kappa}{2}$.

Proof. By Theorem 12.5.2, $\mathcal{L}$ satisfies $\mathrm{HWI}(\kappa)$. $\mathrm{MLSI}(\kappa)$ follows from an application of Young's inequality:

$$xy \leq cx^2 + \frac{1}{4c}y^2, \qquad \forall x, y \in \mathbb{R}, c > 0, \qquad (12.56)$$

in which we set $x = W_{2,\mathcal{L}}(\rho,\sigma)$, $y = \sqrt{\mathrm{EP}_\sigma(\rho)}$, and $c = \frac{\kappa}{2}$. $\square$

12.5.2. The quantum transportation cost inequality

In analogy with Section 4.6, we introduce the concept of a quantum transportation cost inequality associated to the Wasserstein distance $W_{2,\mathcal{L}}$. More precisely, given a primitive QMS $(\mathcal{P}_t)_{t\geq 0}$ on $\mathcal{B}(\mathcal{H})$ satisfying σ-DBC $(\mathcal{P}_t)_{t\geq 0}$ is said to satisfy a *transportation cost inequality of order* 2 with constant $c_2 > 0$ if for all $\rho \in \mathcal{D}_+(\mathcal{H})$,

$$W_{2,\mathcal{L}}(\rho,\sigma) \leq \sqrt{2\,c_2\,D(\rho\|\sigma)}. \qquad (\mathrm{TC}_2(c_2))$$

The QMS is also said to satisfy the *quantum* $\mathrm{MLSI} + \mathrm{TC}_2(c)$ *inequality* of constant $c > 0$ if all $\rho \in \mathcal{D}_+(\mathcal{H})$,

$$W_{2,\mathcal{L}}(\rho, \sigma) \leq c \sqrt{\mathrm{EP}_\sigma(\rho)}. \qquad\qquad (\mathrm{MLSI} + \mathrm{TC}_2(c))$$

In the case $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ for $\kappa \in \mathbb{R}$, $\mathrm{HWI}(\kappa)$ still implies the modified log-Sobolev inequality under the further condition that TC_2 holds. This is a direct quantum generalization of Theorem 7.8 of [Erbar and Maas, 2012] (see also Corollary 3.1 of [Otto and Villani, 2000a]):

Corollary 12.5.4. *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq \kappa$, $\kappa \in \mathbb{R}$, *and that* $\mathrm{TC}_2(c_2)$ *holds with* $c_2^{-1} \geq \max(0, -\kappa)$. *then* $\mathrm{MLSI}(\alpha_1)$ *holds for*

$$\alpha_1 = \max\left[\frac{1}{8\,c_2}\left(1 + c_2\kappa\right)^2, \frac{\kappa}{2}\right]$$

Proof. The proof is rudimentary and identical to the one of Corollary 3.1 of [Otto and Villani, 2000a]. $\qquad\square$

Similarly, we can show that $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ for $\kappa \in \mathbb{R}$ implies $\mathrm{MLSI}(\alpha_1)$ as long as $\mathrm{MLSI} + \mathrm{TC}_2(c)$ holds.

Corollary 12.5.5. *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq \kappa$, $\kappa \in \mathbb{R}$, *and that* $\mathrm{MLSI} + \mathrm{TC}_2(c)$ *holds with* $c^{-1} \geq \max(\kappa, 0)$, *then* $\mathrm{MLSI}(\alpha_1)$ *holds, with*

$$\alpha_1 = \frac{1}{2\,c\,(2 - \kappa c)}.$$

Proof. See Corollary 3.2 of [Otto and Villani, 2000a]. $\qquad\square$

The following result, namely that $\mathrm{MLSI}(\alpha_1)$ implies $\mathrm{TC}_2(c_2)$ was first proven in the classical, continuous case by Otto and Villani in [Otto and Villani, 2000a] (see also [Bobkov et al., 2001, Gozlan, 2009] for alternative proofs, Theorem 7.5 of [Erbar and Maas, 2012] for the classical, discrete case, and [Carlen and Maas, 2014] for the case of the fermionic Fokker-Planck semigroup).

Theorem 12.5.6. *If* $(\mathcal{P}_t)_{t\geq 0}$ *satisfies* $\mathrm{MLSI}(\alpha_1)$, *then* $\mathrm{TC}_2(c_2)$ *holds with* $2\,c_2 = \alpha_1^{-1}$.

Proof. Fix $\rho \in \mathcal{D}_+(\mathcal{H})$, and set $\rho_t = \mathcal{P}_{*t}(\rho)$. First note that as $t \to \infty$,

$$D(\rho_t\|\sigma) \to 0 \text{ and } W_{2,\mathcal{L}}(\rho, \rho_t) \to W_{2,\mathcal{L}}(\rho, \sigma) \qquad\qquad (12.57)$$

Define now the function

$$F(t) := W_{2,\mathcal{L}}(\rho_t, \rho) + \sqrt{\frac{1}{\alpha_1} D(\rho_t\|\sigma)}.$$

Obviously $F(0) = \sqrt{D(\rho\|\sigma)/\alpha_1}$, and by (12.57), $F(t) \to W_{2,\mathcal{L}}(\sigma, \rho)$ as $t \to \infty$. Hence it is sufficient to prove that F is non-increasing. In order to do so, we only need to show that its derivative is non-positive. If $\rho_t \neq \sigma$, we know from Lemma 12.5.1 that

$$\frac{d}{dt}F(t) \leq \sqrt{\mathrm{EP}_\sigma(\rho_t)} + \sqrt{\frac{1}{\alpha_1}} \frac{\frac{d}{dt}D(\rho_t\|\sigma)}{2\sqrt{D(\rho_t\|\sigma)}} = \sqrt{\mathrm{EP}_\sigma(\rho_t)} - \frac{\mathrm{EP}_\sigma(\rho_t)}{\sqrt{4\alpha_1 D(\rho_t\|\sigma)}} \leq 0.$$

where we used $\mathrm{MLSI}(\alpha_1)$ in the last inequality. If $\rho_t = \sigma$, then the relation also holds true, since this implies that $\rho_r = \sigma$ for all $r \geq t$. $\qquad\square$

Remark 12.5.7. In the last version of [Carlen and Maas, 2017], the authors independently added a slightly different proof of Theorem 12.5.6.

The following corollary of Theorem 12.5.6 easily follows

Corollary 12.5.8. *Assume that* $(\mathcal{P}_t)_{t\geq 0}$ *satisfies* $\mathrm{MLSI}(\alpha_1)$ *for some* $\alpha_1 > 0$. *Then it also satisfies* $\mathrm{MLSI} + \mathrm{TC}_2(c)$ *with* $2c = \alpha_1^{-1}$.

A new proof of Pinsker's inequality We show that a refinement of Lemma 12.1.7 as well as Theorem 12.5.6 can be used to provide a new proof of the quantum Pinsker inequality (cf. Equation (1.57)).

Theorem 12.5.9 (Quantum Pinsker's inequality). *For any* $\rho, \sigma \in \mathcal{D}_+(\mathcal{H})$,

$$\|\rho - \sigma\|_1 \leq \sqrt{2D(\rho\|\sigma)}.$$

Proof. Let $\mathcal{L}_{1/d_{\mathcal{H}}}$ be the generator of the quantum depolarizing semigroup with unique invariant state $\mathbb{1}/d_{\mathcal{H}}$: for any $X \in \mathcal{B}(\mathcal{H})$:

$$\mathcal{L}_{1/d_{\mathcal{H}}}(X) = \frac{1}{d_{\mathcal{H}}} \mathrm{Tr}(X)\mathbb{1} - X.$$

It is shown in Equation (5.42) of that $\mathcal{L}_{1/d_{\mathcal{H}}}$ can take the following form:

$$\mathcal{L}_{1/d_{\mathcal{H}}}(X) = \frac{1}{2d_{\mathcal{H}}} \sum_{k,l=1}^{d_{\mathcal{H}}} |e_k\rangle\langle e_l| \left[X, |e_l\rangle\langle e_k|\right] + \left[|e_k\rangle\langle e_l|, X\right] |e_l\rangle\langle e_k|,$$

for any orthonormal basis $\{e_i\}$. Recall the last line (12.17) of the proof of Lemma 12.1.7, where we showed that for any $\delta > 0$, any smooth path $(\gamma(s))_{s\in[0,1]}$ such that $\gamma(0) = \rho$, $\gamma(1) = \sigma$, and

$$\left(\int_0^1 \|\dot{\gamma}(s)\|_{g_{\mathcal{L}_{1/d_{\mathcal{H}}},\gamma(s)}}\right)^{1/2} \leq W_{2,\mathcal{L}_{1/d_{\mathcal{H}}}}(\rho,\sigma) + \delta,$$

and for any self-adjoint operator X,

$$|\mathrm{Tr}(X(\rho - \sigma))|$$
$$\leq \left(\int_0^1 \left\{\sum_{j\in\mathcal{J}} \mathrm{Tr}(\gamma(s)(\nabla_{\tilde{L}_j}X)^*\nabla_{\tilde{L}_j}X) + \mathrm{Tr}(\gamma(s)\nabla_{\tilde{L}_j}X(\nabla_{\tilde{L}_j}X)^*)\right\} ds\right)^{\frac{1}{2}} \left(W_{2,\mathcal{L}_{\frac{1}{d_{\mathcal{H}}}}}(\rho,\sigma) + \delta\right),$$

$$(12.58)$$

where for the depolarizing semigroup, the index $j \in \mathcal{J}$ represents a couple (k,l), so that $\tilde{L}_{kl} = \frac{1}{\sqrt{2d_{\mathcal{H}}}} |e_k\rangle\langle e_l|$. Choosing the basis $\{e_k\}_{k=1}^{d_{\mathcal{H}}}$ to be the one diagonalizing the operator $X \coloneqq \sum_{k=1}^{d_{\mathcal{H}}} \varphi(k)|e_k\rangle\langle e_k|$, the term in brackets on the right hand side of (12.58) reduces to

$$\frac{1}{2d_{\mathcal{H}}} \sum_{k,l} (\varphi(l) - \varphi(k))^2 \left(\langle e_k, \gamma(s)e_k\rangle + \langle e_l, \gamma(s)e_l\rangle\right) \leq \|\varphi\|_{\mathrm{lip,H}}^2,$$

where $\|\varphi\|_{\mathrm{lip,H}} \coloneqq \sup_{k\neq l} |\varphi(k) - \varphi(l)|$. Therefore, letting δ tend to 0,

$$|\mathrm{Tr}(X(\rho - \sigma))| \leq \|\varphi\|_{\mathrm{lip,H}} W_{2,\mathcal{L}_{1/d_{\mathcal{H}}}}(\rho,\sigma).$$

Assuming, moreover, that $0 \leq X \leq \mathbb{1}$, this implies that, for any $k \neq l$, $|\varphi(k) - \varphi(l)| \leq 1$, and thus $\|\varphi\|_{\mathrm{lip,H}} \leq 1$. By duality,

$$\|\rho - \sigma\|_1 \equiv \sup_{0 \leq X \leq \mathbb{1}} |\mathrm{Tr}(X(\rho - \sigma))| \leq \sup_{X = \sum_j \varphi(j)|e_j\rangle\langle e_j|: \|\varphi\|_{\mathrm{lip,H}} \leq 1} |\mathrm{Tr}\, X(\rho - \sigma)| \leq W_{2, \mathcal{L}_{1/d_{\mathcal{H}}}}(\rho, \sigma).$$

We conclude using Theorem 12.5.6 as well as the well-known fact that in the case of the depolarizing semigroup, $\alpha_1 = 1/4$ (see e.g. Lemma 25 of [Kastoryano and Temme, 2013]). □

From transportation cost to Poincaré inequality   In the classical, Riemannian case, [Otto and Villani, 2000a] proved that $\mathrm{TC}_2(c_2)$ implies the Poincaré inequality. In the discrete setting, this was proved in Proposition 7.6 of [Erbar and Maas, 2012]. Theorem 12.5.10 below extends these results to the quantum regime.

Theorem 12.5.10. *If $(\mathcal{P}_t)_{t \geq 0}$ satisfies $\mathrm{TC}_2(c_2)$, then $\mathrm{PI}(\lambda)$ holds with respect to $\langle .,. \rangle_\sigma$, with $\lambda = (c_2\, \kappa_{\mathcal{L}})^{-1}$, where $\kappa_{\mathcal{L}} = \sup_{j \in \mathcal{J}} \|[\sigma]_{\varphi^{\log}, -\omega_j}^{-\frac{1}{2}} \circ [\sigma]_{\varphi^{\log}, \omega_j} \circ [\sigma]_{\varphi^{\log}, -\omega_j}^{-\frac{1}{2}} : \mathcal{T}_2(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})\|$.*

Remark 12.5.11. In the commutative setting, Theorem 12.5.10 reduces to Proposition 7.6 of [Erbar and Maas, 2012]. Indeed, in this case, one can easily verify that for any $j \in \mathcal{J}$, $[\sigma]_{\varphi^{\log}, \omega_j}(X) = \frac{1}{2}\sinh(\omega_j/2)\sigma X$ and $[\sigma]_{\varphi^{\log}, -\omega_j}^{-1}(X) = 2X/(\sigma \sinh(\omega_j/2))$. Therefore $[\sigma]_{\varphi^{\log}, \omega_j} \circ [\sigma]_{\varphi^{\log}, -\omega_j}^{-1}(X) = X$, so that $\kappa_{\mathcal{L}} = 1$ and the result follows.

Proof. Let $X \in \mathcal{B}(\mathcal{H})$ such that $\mathrm{Tr}(\sigma X) = 0$, and for some ε small enough, define $X^\varepsilon := \mathbb{1} + \varepsilon X > 0$. Then, define the completely positive, trace-preserving map Ξ_σ through the following equation: for any $Y \in \mathcal{B}(\mathcal{H})$,

$$\Xi_\sigma(Y) := \int_0^\infty \frac{\sigma^{1/2}}{t\mathbb{1} + \sigma} Y \frac{\sigma^{1/2}}{t\mathbb{1} + \sigma} dt$$

In order to get the result we will need the following two technical lemmas:

Lemma 12.5.12. *With the notations of Equation (12.1),*

$$\tilde{L}_j(t\mathbb{1} + \sigma)^{-1} = \frac{\mathrm{e}^{-\omega_j}\,\sigma^{-1}}{\mathbb{1} + t\,\mathrm{e}^{-\omega_j}\,\sigma^{-1}} \tilde{L}_j \tag{12.59}$$

$$(t\mathbb{1} + \sigma)^{-1}\tilde{L}_j = \tilde{L}_j \frac{\sigma^{-1}\,\mathrm{e}^{\omega_j}}{\mathbb{1} + t\sigma^{-1}\,\mathrm{e}^{\omega_j}}. \tag{12.60}$$

Proof. These identities follow directly from the fact that $\tilde{L}_j$ is an eigenvector of Δ_σ with associated eigenvalue $\mathrm{e}^{-\omega_j}$. □

Lemma 12.5.13. *For $\Gamma_\sigma(X) \equiv \sigma^{\frac{1}{2}} X \sigma^{\frac{1}{2}}$,*

$$\nabla_{\tilde{L}_j}(\Xi_\sigma(X)) = [\sigma]_{\varphi^{\log}, -\omega_j}^{-1} \circ \Gamma_\sigma \circ \nabla_{\tilde{L}_j} X. \tag{12.61}$$

Proof. Using that for each $j \in \mathcal{J}$, $\Delta_\sigma^{\pm\frac{1}{2}}(\tilde{L}_j) = \mathrm{e}^{\mp\omega_j/2}\,\tilde{L}_j$,

$$\nabla_{\tilde{L}_j}(\Xi_\sigma(X)) = \mathrm{e}^{\omega_j/2}\,\sigma^{1/2}\tilde{L}_j \int_0^\infty (t\mathbb{1} + \sigma)^{-1}X(t\mathbb{1} + \sigma)^{-1}\sigma^{1/2}dt \tag{12.62}$$

$$- \mathrm{e}^{-\omega_j/2}\,\sigma^{1/2} \int_0^\infty (t\mathbb{1} + \sigma)^{-1}X(t\mathbb{1} + \sigma)^{-1}\tilde{L}_j\sigma^{1/2}dt$$

Let us first consider the first term on the right hand side of Equation (12.62). By Equation (12.59) it is equal to

$$
\int_0^\infty \frac{e^{\omega_j/2}}{e^{\omega_j}\sigma + t\mathbb{1}} \sigma^{1/2}\tilde{L}_j X \sigma^{1/2}(t\mathbb{1}+\sigma)^{-1} dt = \int_0^\infty \frac{1}{e^{\omega_j}\sigma + e^{\omega_j/2} u\mathbb{1}}\Gamma_\sigma(\tilde{L}_j X)\frac{e^{\omega_j}}{e^{\omega_j/2} u\mathbb{1}+\sigma} du
$$

$$
= \int_0^\infty \frac{e^{-\omega_j/2}}{e^{\omega_j/2}\sigma + u\mathbb{1}}\Gamma_\sigma(\tilde{L}_j X)\frac{e^{-\omega_j/2}}{u\mathbb{1}+e^{-\omega_j/2}\sigma} e^{\omega_j} du
$$

$$
= [\sigma]_{\varphi^{\log},-\omega_j}^{-1}\circ\Gamma_\sigma(\tilde{L}_j X),
$$

where we made the change of variable $e^{\omega_j/2}u = t$ in the first line, and used Lemma 12.1.3 in the last line. Similarly, using (12.60), the second term on the right hand side of (12.62) is equal to

$$
\int_0^\infty \frac{e^{-\omega_j/2}}{t\mathbb{1}+\sigma}\Gamma_\sigma(X\tilde{L}_j)\frac{1}{e^{-\omega_j}\sigma + t\mathbb{1}} dt = \int_0^\infty \frac{1}{u+e^{\omega_j/2}\sigma}\Gamma_\sigma(X\tilde{L}_j)\frac{1}{e^{-\omega_j/2}\sigma + u} du
$$

$$
= [\sigma]_{\varphi^{\log},-\omega_j}^{-1}(\Gamma_\sigma(X\tilde{L}_j)),
$$

where we made the change of variable $t = e^{-\omega_j/2}u$. Hence Equation (12.61) follows. $\square$

We are now ready to prove Theorem 12.5.10. Start by the following:

$$
\langle X, \Xi_\sigma(X)\rangle_\sigma = \mathrm{Tr}(\Gamma_\sigma(X)\Xi_\sigma(X)) = \frac{1}{\varepsilon}\mathrm{Tr}(\sigma^{1/2}\Xi_\sigma(X)\sigma^{1/2}(X^\varepsilon - \mathbb{1}))
$$

$$
= \frac{1}{\varepsilon}\mathrm{Tr}(\Xi_\sigma(X)(\sigma^{1/2}X^\varepsilon\sigma^{1/2}-\sigma)).
$$

For any $\delta > 0$, there exists a smooth path $(\gamma^\varepsilon(s))_{s\in[0,1]}$, with associated vector field $(\mathbf{V}^\varepsilon(s) = \nabla_{\tilde{\mathbf{L}}}(\Phi(s)))_{s\in[0,1]}$, interpolating between $\rho^\varepsilon := \Gamma_\sigma(X^\varepsilon)$ and σ, and such that

$$
\left(\int_0^1 g_{\mathcal{L},\gamma^\varepsilon(s)}(\dot{\gamma}^\varepsilon(s),\dot{\gamma}^\varepsilon(s))\,ds\right)^{1/2} \le W_{2,\mathcal{L}}(\rho^\varepsilon,\sigma)+\delta \tag{12.63}
$$

This implies that

$$
\langle X, \Xi_\sigma(X)\rangle_\sigma = -\frac{1}{\varepsilon}\mathrm{Tr}\left(\Xi_\sigma(X)\int_0^1 \frac{d}{ds}\gamma^\varepsilon(s)ds\right)
$$

$$
= \frac{1}{\varepsilon}\mathrm{Tr}\left(\Xi_\sigma(X)\int_0^1 \mathrm{div}([\gamma^\varepsilon(s)]_{\varphi^{\log},\boldsymbol{\omega}}\mathbf{V}^\varepsilon(s))\,ds\right)
$$

$$
= -\frac{1}{\varepsilon}\int_0^1 \sum_{j\in\mathcal{J}}\langle\nabla_{\tilde{L}_j}\Xi_\sigma(X),[\gamma^\varepsilon(s)]_{\varphi^{\log},\omega_j}(V^\varepsilon(s))_j\rangle_{\mathrm{HS}}\,ds
$$

$$
\le \left(\sum_{j\in\mathcal{J}}\int_0^1\langle\nabla_{\tilde{L}_j}\Xi_\sigma(X),[\gamma^\varepsilon(s)]_{\varphi^{\log},\omega_j}\circ\nabla_{\tilde{L}_j}\Xi_\sigma(X)\rangle_{\mathrm{HS}}\,ds\right)^{1/2}\frac{\left(\int_0^1\|\mathbf{V}^\varepsilon(s)\|_{\mathcal{L},\gamma^\varepsilon(s)}^2 ds\right)^{1/2}}{\varepsilon}
$$

$$
\le \left(\sum_{j\in\mathcal{J}}\int_0^1\langle\nabla_{\tilde{L}_j}\Xi_\sigma(X),[\gamma^\varepsilon(s)]_{\varphi^{\log},\omega_j}\circ\nabla_{\tilde{L}_j}\Xi_\sigma(X)\rangle_{\mathrm{HS}}\,ds\right)^{1/2}\frac{W_{2,\mathcal{L}}(\rho^\varepsilon,\sigma)+\delta}{\varepsilon}
$$

$$
\le \left(\sum_{j\in\mathcal{J}}\int_0^1\langle\nabla_{\tilde{L}_j}\Xi_\sigma(f),[\gamma^\varepsilon(s)]_{\varphi^{\log},\omega_j}\circ\nabla_{\tilde{L}_j}\Xi_\sigma(X)\rangle_{\mathrm{HS}}\,ds\right)^{1/2}\frac{\sqrt{2c_2 D(\rho^\varepsilon\|\sigma)}+\delta}{\varepsilon},
$$

$$\tag{12.64}$$

where the first inequality comes from a use of the Cauchy-Schwarz inequality with respect to the inner product $\sum_{j\in\mathcal{J}}\langle.\,,\int_0^1[\gamma^\varepsilon(s)]_{\varphi^{\log},\omega_j}ds\,.\rangle_{\mathrm{HS}}$, the second from Equation (12.63) and the last one

from $\mathrm{TC}_2(c_2)$. As $\varepsilon \to 0$, the term in brackets in (12.64) converges to

$$\sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} \Xi_\sigma(X), [\sigma]_{\varphi^{\log}, \omega_j} \circ \nabla_{\tilde{L}_j} \Xi_\sigma(X) \rangle_{\mathrm{HS}}$$

$$= \sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} X, \Gamma_\sigma \circ [\sigma]_{\varphi^{\log}, -\omega_j}^{-1} \circ [\sigma]_{\varphi^{\log}, \omega_j} \circ [\sigma]_{\varphi^{\log}, -\omega_j}^{-1} \circ \Gamma_\sigma(\nabla_{\tilde{L}_j} X) \rangle_{\mathrm{HS}}$$

$$\leq \sup_{j \in \mathcal{J}} \| [\sigma]_{\varphi^{\log}, -\omega_j}^{-\frac{1}{2}} \circ [\sigma]_{\varphi^{\log}, \omega_j} \circ [\sigma]_{\varphi^{\log}, -\omega_j}^{-\frac{1}{2}} : \mathcal{T}_2(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H}) \| \, \mathcal{E}_{2,\mathcal{L}}(X, \Xi_\sigma(X))$$

$$\equiv \kappa_\mathcal{L} \, \mathcal{E}_{2,\mathcal{L}}(X, \Xi_\sigma(X))$$

where we used Lemma 12.5.13 on the first line. As $\delta > 0$ was chosen arbitrarily, we can now take the limit $\delta \to 0$. Moreover, following the approach of the proof of Theorem 16 of [Kastoryano and Temme, 2013], one can prove that

$$\frac{D(\rho^\varepsilon \| \sigma)}{\varepsilon^2} \to \frac{1}{2} \, \mathrm{Tr}(\Gamma_\sigma(X) \, \Xi_\sigma(X))$$

Substituting into (12.64), we get

$$\frac{1}{c_2 \, \kappa_\mathcal{L}} \langle X, \Xi_\sigma(X) \rangle_\sigma \leq \mathcal{E}_{2,\mathcal{L}}(X, \Xi_\sigma(X)).$$

This is exactly the form that was derived at the end of the proof of Theorem 16 of [Kastoryano and Temme, 2013] which led to the Poincaré inequality. □

12.6. Diameter estimates

The *diameter* of $\mathcal{D}(\mathcal{H})$ in the Wasserstein distance $W_{2,\mathcal{L}}$ is defined as follows:

$$\mathrm{Diam}_\mathcal{L}(\mathcal{D}(\mathcal{H})) := \sup_{\rho, \sigma \in \mathcal{D}(\mathcal{H})} W_{2,\mathcal{L}}(\rho, \sigma).$$

Another straightforward consequence of the κ-displacement convexity of the quantum relative entropy for $\kappa > 0$ is the following estimate on the diameter of $\mathrm{Diam}_\mathcal{L}(\mathcal{D}(\mathcal{H}))$, which is a quantum analogue of the Bonnet-Myers theorem (cf. (4.10), see also Proposition 7.3 of [Erbar and Fathi, 2018]).

Proposition 12.6.1. *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ *holds for* $\kappa > 0$. *Then for any two states* $\rho, \omega \in \mathcal{D}(\mathcal{H})$,

$$W_{2,\mathcal{L}}(\rho, \omega)^2 \leq \frac{4}{\kappa} (D(\rho \| \sigma) + D(\omega \| \sigma)).$$

Therefore,

$$\mathrm{Diam}_\mathcal{L}(\mathcal{D}(\mathcal{H})) \leq \sqrt{\frac{8 \ln \| \sigma^{-1} \|_\infty}{\kappa}} \,.$$

Proof. The result follows directly from the convexity of the quantum relative entropy (cf. (v) of Theorem 12.3.3):

$$0 \leq D(\gamma(1/2) \| \sigma) \leq \frac{1}{2} D(\rho \| \sigma) + \frac{1}{2} D(\omega \| \sigma) - \frac{\kappa}{8} W_{2,\mathcal{L}}(\rho, \omega)^2.$$

for a given constant speed geodesic $(\gamma(s))_{s \in [0,1]}$ relating ρ and ω. □

In the next two subsections, we show how one can recover the Poincaré and modified logarithmic

Sobolev inequalities in the case when $\mathrm{Ric}(\mathcal{L}) \geq 0$ by further assuming a bound on the diameter $\mathrm{Diam}_{\mathcal{L}}$.

12.6.1. From diameter bound to the Poincaré inequality

In this section we show that $\mathrm{Ric}(\mathcal{L}) \geq 0$ together with a condition of finiteness of the diameter of $\mathcal{D}(\mathcal{H})$ with respect to the distance $W_{2,\mathcal{L}}$ implies the Poincaré inequality, hence extending Proposition 5.9 of [Erbar and Fathi, 2018] to our non-commutative setting. Throughout this section, we fix $(\mathcal{P}_t)_{t\geq 0}$ to be a primitive QMS on $\mathcal{B}(\mathcal{H})$, $\mathcal{H}$ finite dimensional, with unique invariant state σ and associated generator $\mathcal{L}$, satisfying σ-DBC. The next result is a non-commutative extension of the third equivalent statement in Theorem 4.7.2 of [Bakry et al., 2014]:

Proposition 12.6.2 (Reverse quantum Poincaré inequality)**.** *Assume that $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ holds for some $\kappa \in \mathbb{R}$. Then for any $\rho \in \mathcal{D}_+(\mathcal{H})$, any $\Phi \in \mathcal{B}_{sa}(\mathcal{H})$ and all $t > 0$:*

$$\mathrm{Tr}(\mathcal{P}_{*t}(\rho)\Phi^2) - \mathrm{Tr}(\rho\,(\mathcal{P}_t(\Phi))^2) \geq \frac{\mathrm{e}^{2\kappa t}-1}{\kappa}\|\nabla_{\tilde{\mathbf{L}}}\mathcal{P}_t(\Phi)\|^2_{\mathbb{L}_{2,\log}(\rho)} \tag{12.65}$$

Proof. The proof is similar to the one of Theorem 3.5 of [Erbar and Fathi, 2018]. For $u \geq 0$, let $\rho_u \equiv \mathcal{P}_{*u}(\rho)$ and $\Phi(u) \equiv \mathcal{P}_u(\Phi)$. Then, from Proposition 12.3.4,

$$2\,\mathrm{e}^{2\kappa s}\|\nabla_{\tilde{\mathbf{L}}}\Phi(t)\|^2_{\mathbb{L}_{2,\log}(\rho)}$$

$$= 2\,\mathrm{e}^{2\kappa s}\,\|\nabla_{\tilde{\mathbf{L}}}(\mathcal{P}_s\Phi(t-s))\|^2_{\mathbb{L}_{2,\log}(\rho)}$$

$$\leq 2\,\|\nabla_{\tilde{\mathbf{L}}}\Phi(t-s)\|^2_{\mathbb{L}_{2,\log}(\rho_s)}$$

$$= 2\sum_{j\in\mathcal{J}}\langle\nabla_{\tilde{L}_j}\Phi(t-s),[\rho_s]_{\varphi^{\log},\omega_j}\nabla_{\tilde{L}_j}\Phi(t-s)\rangle_{\mathrm{HS}}$$

$$\leq \sum_{j\in\mathcal{J}}\langle\nabla_{\tilde{L}_j}\Phi(t-s),(\mathrm{e}^{-\omega_j/2}R_{\rho_s}+\mathrm{e}^{\omega_j/2}L_{\rho_s})\nabla_{\tilde{L}_j}\Phi(t-s)\rangle_{\mathrm{HS}}$$

$$= \sum_{j\in\mathcal{J}}\left(\mathrm{e}^{\omega_j/2}\,\mathrm{Tr}[\rho_s\nabla_{\tilde{L}_j}\Phi(t-s)(\nabla_{\tilde{L}_j}\Phi(t-s))^*]+\mathrm{e}^{-\omega_j/2}\,\mathrm{Tr}[\rho_s(\nabla_{\tilde{L}_j}\Phi(t-s))^*\nabla_{\tilde{L}_j}\Phi(t-s)]\right)$$

$$= \sum_{j\in\mathcal{J}}\left(\mathrm{e}^{\omega_j/2}\,\mathrm{Tr}(\rho_s[\tilde{L}_j,\Phi(t-s)][\Phi(t-s),\tilde{L}_j^*])+\mathrm{e}^{-\omega_j/2}\,\mathrm{Tr}(\rho_s[\Phi(t-s),\tilde{L}_j^*][\tilde{L}_j,\Phi(t-s)])\right)$$

$$= \sum_{j\in\mathcal{J}}\left(\mathrm{Tr}(\rho_s(\tilde{L}_j\Phi(t-s)^2\tilde{L}_j^*\,\mathrm{e}^{\omega_j/2}+\mathrm{e}^{-\omega_j/2}\,\tilde{L}_j^*\Phi(t-s)^2\tilde{L}_j))\right.$$

$$+\mathrm{e}^{\omega_j/2}\,\mathrm{Tr}(\rho_s(-\tilde{L}_j\Phi(t-s)\tilde{L}_j^*\Phi(t-s)-\Phi(t-s)\tilde{L}_j\Phi(t-s)\tilde{L}_j^*+\Phi(t-s)\tilde{L}_j\tilde{L}_j^*\Phi(t-s)))$$

$$\left.+\mathrm{e}^{-\omega_j/2}\,\mathrm{Tr}(\rho_s(\Phi(t-s)\tilde{L}_j^*\tilde{L}_j\Phi(t-s)-\tilde{L}_j^*\Phi(t-s)\tilde{L}_j\Phi(t-s)-\Phi(t-s)\tilde{L}_j^*\Phi(t-s)\tilde{L}_j))\right)$$

$$= \mathrm{Tr}(\rho_s\mathcal{L}(\Phi(t-s)^2))$$

$$+\sum_{j\in\mathcal{J}}\mathrm{e}^{\omega_j/2}\,\mathrm{Tr}\left[\rho_s(-\tilde{L}_j\Phi(t-s)\tilde{L}_j^*\Phi(t-s)-\Phi(t-s)\tilde{L}_j\Phi(t-s)\tilde{L}_j^*\right.$$

$$\left.+\Phi(t-s)\tilde{L}_j\tilde{L}_j^*\Phi(t-s)+\Phi(t-s)^2\tilde{L}_j\tilde{L}_j^*)\right]$$

$$+\mathrm{e}^{-\omega_j/2}\,\mathrm{Tr}\left[\rho_s(\Phi(t-s)\tilde{L}_j^*\tilde{L}_j\Phi(t-s)-\tilde{L}_j^*\Phi(t-s)\tilde{L}_j\Phi(t-s)\right.$$

$$\left.-\Phi(t-s)\tilde{L}_j^*\Phi(t-s)\tilde{L}_j+\tilde{L}_j^*\tilde{L}_j\Phi(t-s)^2)\right]$$

$$= \mathrm{Tr}(\rho_s\mathcal{L}(\Phi(t-s)^2))-\mathrm{Tr}(\rho_s\Phi(t-s)\mathcal{L}(\Phi(t-s)))-\mathrm{Tr}(\rho_s\mathcal{L}(\Phi(t-s))\Phi(t-s))$$

$$= \frac{\partial}{\partial s}\mathrm{Tr}(\rho_s\Phi^2(t-s))$$

where we used (12.13) in the fifth line. The claim follows after integrating from 0 to t. $\qquad\square$

Theorem 12.6.3. $\mathrm{Ric}(\mathcal{L}) \geq 0 \; + \; \mathrm{Diam}_{\mathcal{L}}(\mathcal{D}(\mathcal{H})) \leq D \Rightarrow \mathrm{PI}(\frac{1}{\mathrm{e}\,D^2})$.

Proof. Let $X \in \mathcal{B}_{sa}(\mathcal{H})$ an eigenvector of $\mathcal{L}$ with associated eigenvalue opposite to the spectral gap λ of $\mathcal{L}$. Without loss of generality, $\|X\|_\infty = 1$, and by primitivity of $(\mathcal{P}_t)_{t \geq 0}$, $\mathrm{Tr}(\sigma X) = 0$. Now, note that $\mathcal{P}_t(X) = \mathrm{e}^{-\lambda t} X$. Therefore, the reverse Poincaré inequality (12.65) in the case when $\kappa = 0$ implies that for any $\rho \in \mathcal{D}_+(\mathcal{H})$,

$$\|\nabla_{\tilde{\mathbf{L}}} X\|^2_{\mathbb{L}_{2,\log}(\rho)} \leq \frac{\mathrm{e}^{2\lambda t}}{2t} \|X\|^2_\infty .$$

Optimizing in t and using $\|X\|_\infty = 1$, we find

$$\|\nabla_{\tilde{\mathbf{L}}} X\|^2_{\mathbb{L}_{2,\log}(\rho)} \leq \mathrm{e}\,\lambda .$$

Given the following spectral decomposition of $X = \sum_\mu \mu P_\mu$, since $\mathrm{Tr}(\sigma X) = 0$, the minimum and maximum eigenvalues of X, respectively denoted by $\mu_{\min}$ and $\mu_{\max}$, obey $\mu_{\min} < 0 < \mu_{\max}$. Since we assumed $\|X\|_\infty = 1$, this implies that given a path $(\gamma(s))_{s \in [0,1]}$ in $\mathcal{D}(\mathcal{H})$ joining the states $\gamma(0) = \frac{P_{\mu_{\max}}}{\mathrm{Tr}(P_{\mu_{\max}})}$ and $\gamma(1) = \frac{P_{\mu_{\min}}}{\mathrm{Tr}(P_{\mu_{\min}})}$ such that $\int_0^1 g_{\mathcal{L},\gamma(s)}(\dot{\gamma}(s),\dot{\gamma}(s))\, ds \leq W_{2,\mathcal{L}}(\gamma(0),\gamma(1))^2 + \varepsilon$,

$$
\begin{aligned}
1 \leq |\mu_{\max} - \mu_{\min}| &= \left| \mathrm{Tr}\, X \left(\frac{P_{\mu_{\min}}}{\mathrm{Tr}(P_{\mu_{\min}})} - \frac{P_{\mu_{\max}}}{\mathrm{Tr}(P_{\mu_{\max}})} \right) \right| \\
&= \left| \mathrm{Tr}\left(X \int_0^1 \dot{\gamma}(s)\, ds \right) \right| \\
&= \left| \int_0^1 \sum_{j \in \mathcal{J}} \langle \nabla_{\tilde{L}_j} X, [\gamma(s)]_{\varphi^{\log},\omega_j} \nabla_{\tilde{L}_j} \Phi(s) \rangle_{\mathrm{HS}}\, ds \right| \\
&\leq \sqrt{(D^2 + \varepsilon)} \left(\int_0^1 \|\nabla_{\tilde{\mathbf{L}}} X\|^2_{\mathbb{L}_{2,\log}(\gamma(s))} ds \right)^{1/2} \leq \sqrt{(D^2+\varepsilon)\lambda\, \mathrm{e}},
\end{aligned}
$$

where in the last line we used the Cauchy-Schwarz inequality with respect to the inner product $\sum_{j \in \mathcal{J}} \langle\,.\,, \int_0^1 [\gamma(s)]_{\varphi^{\log},\omega_j} ds\,.\rangle_{\mathrm{HS}}$, and the result directly follows. $\qquad\square$

12.6.2. From diameter bound to modified log-Sobolev inequality

In [Erbar and Fathi, 2018], a modified logarithmic Sobolev inequality was proved to holds under the conditions that $\mathrm{Ric}(\mathcal{L}) \geq 0$ and of boundedness of the diameter of the underlying space under the modified Wasserstein distance. Here, we extend their results to the quantum regime under the further assumption that the semigroup $(\mathcal{P}_t)_{t \geq 0}$ is unital, leaving the study of the general case to later. The idea of the proof is to get a non-tight logarithmic Sobolev inequality from HWI(0), and then to tighten it using ideas borrowed from [Barthe and Kolesnikov, 2008].

Given two states $\rho, \omega \in \mathcal{D}(\mathcal{H})$, with associated spectral decompositions $\rho = \sum_{i \in E} \lambda_i P_i$, $\omega = \sum_{j \in F} \mu_j Q_j$, where E and F are two finite index sets, a *coupling* of ρ and ω is a probability distribution q on $E \times F$ such that

$$
\begin{aligned}
\sum_{i \in E} q(i,j) &= \mu_j\, \mathrm{Tr}(Q_j) \\
\sum_{j \in F} q(i,j) &= \lambda_i\, \mathrm{Tr}(P_i).
\end{aligned}
$$

The set of couplings between ρ and ω is denoted by $\Pi(\rho,\omega)$. In analogy with the classical literature (see e.g. [Erbar and Maas, 2012]), given an primitive semigroup $(\mathcal{P}_t)_{t \geq 0}$ with associated generator $\mathcal{L}$,

the *coupling Wasserstein distance of order two* between ρ and ω is defined as follows:

$$W_{2,\mathcal{L},c}(\rho,\omega)^2 := \inf_{q \in \Pi(\rho,\sigma)} \sum_{i \in E, \, j \in F} q(i,j) \, W_{2,\mathcal{L}}(\rho_j,\omega_j)^2,$$

where

$$\rho_i := \frac{P_i}{\mathrm{Tr}\, P_i}, \qquad \omega_j := \frac{Q_j}{\mathrm{Tr}(Q_j)}, \qquad i \in E, \; j \in F.$$

The following result is a quantum generalization of Proposition 2.14 of [Erbar and Maas, 2012]:

Proposition 12.6.4. *Let $(\mathcal{P}_t)_{t\geq 0}$ be a primitive QMS, with unique invariant state σ and associated generator $\mathcal{L}$, satisfying the detailed balance condition. Then, for any $\rho, \omega \in \mathcal{D}_+(\mathcal{H})$,*

$$W_{2,\mathcal{L}}(\rho,\omega) \leq W_{2,\mathcal{L},c}(\rho,\omega).$$

Proof. Given $\rho = \sum_{i \in E} \lambda_i P_i$, $\omega = \sum_{j \in F} \mu_j Q_j$ the spectral decompositions of the states ρ and ω. For $(i,j) \in E \times F$, define $\rho_i := \frac{P_i}{\mathrm{Tr}(P_i)}$, $\omega_j := \frac{Q_j}{\mathrm{Tr}(Q_j)}$, and let $\varepsilon > 0$. By definition of the Wasserstein distance $W_{2,\mathcal{L}}$, there exists a curve $\gamma_{ij} : [0,1] \mapsto \mathcal{D}(\mathcal{H})$ from ρ_i to ω_j such that

$$\int_0^1 g_{\mathcal{L},\gamma_{ij}(s)}(\dot\gamma_{ij}(s),\dot\gamma_{ij}(s))\, ds \leq W_{2,\mathcal{L}}(\rho_i,\omega_j)^2 + \varepsilon.$$

For any coupling $q : E \times F \to \mathbb{R}_+$ of the states ρ and ω, define the path $(\gamma(s))_{s\in[0,1]}$ on $\mathcal{D}(\mathcal{H})$ as

$$\gamma(s) = \sum_{i \in E, \, j \in F} q(i,j)\, \gamma_{ij}(s),$$

so that $\gamma(0) = \rho$ and $\gamma(1) = \omega$. Now,

$$\begin{aligned}
W_{2,\mathcal{L}}(\rho,\omega)^2 &\leq \int_0^1 g_{\mathcal{L},\gamma(s)}(\dot\gamma(s),\dot\gamma(s))\, ds \\
&\leq \sum_{i \in E, \, j \in F} q(i,j) \int_0^1 g_{\mathcal{L},\gamma(s)}(\dot\gamma_{ij}(s),\dot\gamma_{ij}(s))\, ds \\
&\leq \sum_{i \in E, \, j \in F} q(i,j)\, W_{2,\mathcal{L}}(\rho_i,\omega_j)^2 + \varepsilon.
\end{aligned}$$

where we used the convexity of $g_{\mathcal{L}}$ in the second line (see e.g. equation (8.15) of [Carlen and Maas, 2017]). As ε was arbitrary, the result follows after optimizing over the couplings q. $\qquad\square$

In what follows, we assume that the semigroup satisfies $\sigma = \mathbb{1}_{\mathcal{H}}/d_{\mathcal{H}}$-DBC. In order to prove the main result of this section, we need the following two lemmas that are extensions of Lemmas 6.2 and 6.3 of [Erbar and Fathi, 2018]:

Lemma 12.6.5. *Assume that $\mathrm{Ric}(\mathcal{L}) \geq 0$ and $\mathrm{Diam}_{\mathcal{L}}(\mathcal{D}(\mathcal{H})) \leq D$. Then for any $\delta > 0$ and $X \in \mathcal{B}_{sa}(\mathcal{H})$ such that $\mathrm{Tr}(X^2) = d_{\mathcal{H}}$:*

$$D\left(X^2/d_{\mathcal{H}} \,\|\, \mathbb{1}/d_{\mathcal{H}}\right) \leq \delta D^2 \, \mathrm{EP}_{\mathbb{1}/d_{\mathcal{H}}}(X^2/d_{\mathcal{H}}) + \frac{1}{4d_{\mathcal{H}}\,\delta}\, \mathrm{Tr}(X^2 \mathbb{1}_{(1,\infty)}(X^2)).$$

Proof. The case when X is not of full support is trivial, as then $\mathrm{EP}_{\sigma}(X^2/d_{\mathcal{H}}) = \infty$. Without loss of generality, we assume that X has full support, so that $X^2/d_{\mathcal{H}} \in \mathcal{D}_+(\mathcal{H})$. Write $X = \sum_{i \in E} \varphi(i) P_i$ the spectral decomposition of X, for some index set E. From HWI(0), and Young's inequality (12.56)

with $c = \delta D^2$, $x = \sqrt{\mathrm{EP}_\sigma(X^2/d_\mathcal{H})}$, and $y = W_{2,\mathcal{L}}(X^2/d_\mathcal{H}, \mathbb{1}/d_\mathcal{H})$:

$$D(X^2/d_\mathcal{H} \| \mathbb{1}/d_\mathcal{H}) \le \delta D^2 \, \mathrm{EP}_{\mathbb{1}/d_\mathcal{H}}(X^2/d_\mathcal{H}) + \frac{1}{4\delta D^2} W_{2,\mathcal{L}}(X^2/d_\mathcal{H}, \mathbb{1}/d_\mathcal{H})^2.$$

From Proposition 12.6.4, for any coupling $q : E \times F \to \mathbb{R}_+$ between $X^2/d_\mathcal{H}$ and $\mathbb{1}/d_\mathcal{H}$ such that $q(i,j) = 0$ whenever $\varphi(i)^2 \le 1$,

$$D(X^2/d_\mathcal{H} \| \mathbb{1}/d_\mathcal{H}) \le \delta D^2 \, \mathrm{EP}_{\mathbb{1}/d_\mathcal{H}}(X^2/d_\mathcal{H}) + \frac{1}{4\delta D^2} \sum_{i,j:\ \varphi(i)^2 > 1} q(i,j) W_{2,\mathcal{L}}\left(\frac{P_i}{\mathrm{Tr}(P_i)}, \frac{P_j}{\mathrm{Tr}(P_j)} \right)^2$$

$$\le \delta D^2 \, \mathrm{EP}_{\mathbb{1}/d_\mathcal{H}}(X^2/d_\mathcal{H}) + \frac{1}{4d_\mathcal{H}\delta} \mathrm{Tr}(X^2 \mathbb{1}_{(1,\infty)}(X^2)),$$

which is what was needed to be proved. $\qquad\square$

Lemma 12.6.6. *For any $A > 1$ there exists $\gamma > 0$ such that for any $X \in \mathcal{B}_{sa}(\mathcal{H})$ with $\mathrm{Tr}(X^2) = d_\mathcal{H}$,*

$$\frac{1}{d_\mathcal{H}} \mathrm{Tr}(X^2 \mathbb{1}_{[A^2,\infty)}(X^2)) \le \left(\frac{A}{A-1} \right)^2 \mathrm{Var}_{\mathbb{1}/d_\mathcal{H}}(X), \tag{12.66}$$

$$D(X^2/d_\mathcal{H} \| \mathbb{1}/d_\mathcal{H}) \le \gamma \, \mathrm{Var}_{\mathbb{1}/d_\mathcal{H}}(X) + \frac{1}{d_\mathcal{H}} \mathrm{Tr}(X^2 \ln X^2 \mathbb{1}_{[A^2,\infty)}(X^2)). \tag{12.67}$$

Proof. This is a direct rewriting of Lemma 2.5 of [Barthe and Kolesnikov, 2008]. $\qquad\square$

Theorem 12.6.7. *Let $(\mathcal{P}_t)_{t\ge0}$ be a primitive semigroup that satisfies $\mathbb{1}/d_\mathcal{H}$-DBC with associated generator $\mathcal{L}$. Assume that $\mathrm{Ric}(\mathcal{L}) \ge 0$ and that $\mathrm{Diam}_\mathcal{L}(\mathcal{D}(\mathcal{H})) \le D$. Then $\mathrm{MLSI}(cD^{-2})$ holds, for some universal constant c.*

Proof. Let $A > 1$ and $X \in \mathcal{B}_{sa}(\mathcal{H})$ of spectral decomposition $X = \sum_{i \in E} \varphi(i) P_i$, with $\mathrm{Tr}(X^2) = d_\mathcal{H}$. Without loss of generality, we can assume X positive definite. Then, set $X_A := X \vee A \equiv \sum_{i:\ \varphi(i)\ge A} \varphi(i) P_i + A \mathbb{1}_{(-\infty,A)}(X)$. Define the state $\rho_A = X_A^2/\mathrm{Tr}(X_A^2)$. By (12.67),

$$D(X^2/d_\mathcal{H} \| \mathbb{1}/d_\mathcal{H}) \le \gamma \, \mathrm{Var}_{\mathbb{1}/d_\mathcal{H}}(X) + \frac{1}{d_\mathcal{H}} \mathrm{Tr}(X^2 \ln X^2 \mathbb{1}_{[A^2,\infty)}(X^2)). \tag{12.68}$$

By Theorem 12.6.3,

$$\gamma \, \mathrm{Var}_{\mathbb{1}/d_\mathcal{H}}(X) \le -\mathrm{e}\, D^2 \gamma \frac{1}{d_\mathcal{H}} \langle X, \mathcal{L}(X) \rangle_{\mathrm{HS}} \le \frac{\gamma \mathrm{e}\, D^2}{4} \mathrm{EP}_{\mathbb{1}/d_\mathcal{H}}(X^2/d_\mathcal{H}), \tag{12.69}$$

where in the last inequality, we used the strong regularity of Dirichlet forms of unital semigroups (see [Kastoryano and Temme, 2013]). Moreover,

$$\frac{1}{d_\mathcal{H}} \mathrm{Tr}(X^2 \ln X^2 \mathbb{1}_{[A^2,\infty)}(X^2)) = \frac{1}{d_\mathcal{H}} \mathrm{Tr}(X_A^2 \ln X_A^2) - \frac{1}{d_\mathcal{H}} A^2 \ln A^2 \, \mathrm{Tr}(\mathbb{1}_{(-\infty,A)}(X))$$

$$\le (1 + A^2) D(\rho_A \| \mathbb{1}/d_\mathcal{H}) + \frac{\mathrm{Tr}(X_A^2)}{d_\mathcal{H}} \ln\left(\frac{\mathrm{Tr} X_A^2}{d_\mathcal{H}} \right) - \frac{A^2 \ln A^2}{d_\mathcal{H}} \mathrm{Tr}(\mathbb{1}_{(-\infty,A)}(X)), \tag{12.70}$$

where in the last line we used that $\frac{1}{d_\mathcal{H}} \mathrm{Tr}(X_A^2) \le 1 + A^2$. However, from Lemma 12.6.5 applied to X_A, since $\mathrm{EP}_{\mathbb{1}/d_\mathcal{H}}(\rho_A) \le \frac{d}{\mathrm{Tr}(X_A^2)} \mathrm{EP}_{\mathbb{1}/d}(X^2/d)$ by convexity of monotone Riemannian metrics (see e.g.

equation (8.16) of [Carlen and Maas, 2017]),

$$D(\rho_A \| \mathbb{1}/d_{\mathcal{H}}) \leq \frac{d_{\mathcal{H}} \, \delta \, D^2}{\mathrm{Tr}(X_A^2)} \mathrm{EP}_{\mathbb{1}/d_{\mathcal{H}}}(X^2/d_{\mathcal{H}}) + \frac{1}{4 \, \mathrm{Tr}(X_A^2)\delta} \mathrm{Tr}(X_A^2 \mathbb{1}_{[\frac{1}{d_{\mathcal{H}}} \mathrm{Tr}(X_A^2),\infty)}(X_A^2))$$

$$\leq \frac{\delta D^2}{A^2} \mathrm{EP}_{\mathbb{1}/d_{\mathcal{H}}}(X^2/d_{\mathcal{H}}) + \frac{1}{4d_{\mathcal{H}} \, \delta \, A^2} \mathrm{Tr}(X^2 \mathbb{1}_{[A^2,\infty)}(X^2)), \qquad (12.71)$$

where in the last line we used that $A^2 \leq \frac{1}{d_{\mathcal{H}}} \mathrm{Tr}(X_A^2)$. Using (12.66) and (12.69) together with Theorem 12.6.3,

$$D(\rho_A \| \mathbb{1}/d_{\mathcal{H}}) \leq \frac{\delta D^2}{A^2} \mathrm{EP}_{\mathbb{1}/d_{\mathcal{H}}}(X^2/d_{\mathcal{H}}) + \frac{1}{4\delta(A-1)^2} \mathrm{Var}_{\mathbb{1}/d_{\mathcal{H}}}(X)$$

$$\leq \frac{\delta D^2}{A^2} \mathrm{EP}_{\mathbb{1}/d_{\mathcal{H}}}(X^2/d_{\mathcal{H}}) - \frac{\mathrm{e} \, D^2}{2\delta(A-1)^2} \frac{1}{d}\langle X, \mathcal{L}(X)\rangle_{\mathrm{HS}}$$

$$\leq \left(\frac{\delta D^2}{A^2} + \frac{\mathrm{e} \, D^2}{8\delta(A-1)^2}\right) \mathrm{EP}_{\mathbb{1}/d_{\mathcal{H}}}(X^2/d_{\mathcal{H}}).$$

Now,

$$\frac{\mathrm{Tr}(X_A^2)}{d_{\mathcal{H}}} \ln\left(\frac{\mathrm{Tr}(X_A^2)}{d_{\mathcal{H}}}\right) - \frac{A^2 \ln(A^2)}{d_{\mathcal{H}}} \mathrm{Tr}(\mathbb{1}_{(-\infty,A)}(X))$$

$$\leq \frac{1}{d_{\mathcal{H}}}\left(A^2 \mathrm{Tr}(\mathbb{1}_{(-\infty,A)}(X)) + \mathrm{Tr}(X^2 \mathbb{1}_{[A^2,\infty)}(X^2))\right) \ln\left[\frac{A^2 \mathrm{Tr}(\mathbb{1}_{(-\infty,A)}(X)) + \mathrm{Tr}(X^2 \mathbb{1}_{[A^2,\infty)}(X^2))}{d_{\mathcal{H}}}\right]$$

$$\qquad - \frac{A^2 \ln(A^2)}{d_{\mathcal{H}}} \mathrm{Tr}(\mathbb{1}_{(-\infty,A)}(X))$$

$$= A^2 \frac{1}{d_{\mathcal{H}}} \mathrm{Tr}(\mathbb{1}_{(-\infty,A)}(f)) \ln\left(\frac{1}{d_{\mathcal{H}}} \mathrm{Tr}(\mathbb{1}_{(-\infty,A)}(X)) + \frac{\mathrm{Tr}(X^2 \mathbb{1}_{[A^2,\infty)}(X^2))}{d_{\mathcal{H}} A^2}\right)$$

$$\qquad + \frac{1}{d_{\mathcal{H}}} \mathrm{Tr}(X^2 \mathbb{1}_{[A^2,\infty)}(X^2)) \ln\left(A^2 \frac{1}{d_{\mathcal{H}}} \mathrm{Tr}(\mathbb{1}_{(-\infty,A]}(X)) + \frac{1}{d_{\mathcal{H}}} \mathrm{Tr}(X^2 \mathbb{1}_{[A^2,\infty)}(X^2))\right)$$

$$\leq A^2 \ln\left(1 + \frac{\mathrm{Tr}(X^2 \mathbb{1}_{[A^2,\infty)}(X^2))}{d_{\mathcal{H}} A^2}\right) + \frac{1}{d_{\mathcal{H}}} \mathrm{Tr}(X^2 \mathbb{1}_{[A^2,\infty)}(X^2)) \ln(1 + A^2)$$

$$\leq (1 + \ln(1 + A^2)) \frac{1}{d_{\mathcal{H}}} \mathrm{Tr}(X^2 \mathbb{1}_{[A^2,\infty)}(X^2)),$$

where in the fourth line we used that $\frac{1}{d_{\mathcal{H}}} \mathrm{Tr}(X^2) = 1$. Using once more (12.66) and (12.69) together with Theorem 12.6.3, we find

$$\frac{\mathrm{Tr}(X_A^2)}{d_{\mathcal{H}}} \ln\left(\frac{\mathrm{Tr}(X_A^2)}{d_{\mathcal{H}}}\right) - \frac{A^2 \ln(A^2)}{d_{\mathcal{H}}} \mathrm{Tr}(\mathbb{1}_{(-\infty,A]}(X)) \leq \frac{\mathrm{e} \, D^2 A^2 (1 + \ln(1 + A^2))}{2(1-A)^2} \mathrm{EP}_{\mathbb{1}/d_{\mathcal{H}}}(X^2/d_{\mathcal{H}}).$$
$$(12.72)$$

The result follows after combining (12.72), (12.68), (12.69), (12.70) and (12.71). □

12.7. Quantum concentration inequalities

12.7.1. The quantum Wasserstein distance of order 1

The Wasserstein distance of order 1 can also be extended to the quantum setting. We recall from Section 4.6 that the classical Wasserstein distance $W_1(\mu, \nu)$ between two probability measures μ, ν on

a metric space has the following Kantorovich-Rubinstein dual representation:

$$W_1(\mu, \nu) = \sup_{\varphi:\, \|\varphi\|_{\mathrm{lip}} \leq 1} \left\{ \int \varphi \, d\mu - \int \varphi \, d\nu \right\},$$

where $\|\varphi\|_{\mathrm{lip}}$ denotes the Lipschitz constant of the function φ. In the same spirit, given a primitive quantum Markov semigroup $(\mathcal{P}_t = e^{t\mathcal{L}})_{t \geq 0}$ on $\mathcal{B}(\mathcal{H})$, $\mathcal{H}$ finite dimensional, satisfying σ-DBC, we define the following Lipschitz constant of a operator $X \in \mathcal{B}(\mathcal{H})$:

$$\|X\|_{\mathrm{Lip}} := \left(\sum_{j \in \mathcal{J}} (e^{-\omega_j/2} + e^{\omega_j/2}) \|\nabla_{\tilde{L}_j} X\|_\infty^2 \right)^{1/2}. \tag{12.73}$$

Next we define the *non-commutative* 1-*Wasserstein distance* between two states ρ and σ associated to the generator $\mathcal{L}$ to be

$$W_{1,\mathcal{L}}(\rho, \omega) := \sup_{\|X\|_{\mathrm{Lip}} \leq 1} |\mathrm{Tr}(X(\rho - \omega))|. \tag{12.74}$$

Note that similar distances were recently defined in [Chen et al., 2017a, Chen et al., 2017b]. The difference with our definition (12.74) lies in the definition of the Lipschitz constant.

Lemma 12.7.1. *The non-commutative* 1-*Wasserstein distance* $W_{1,\mathcal{L}}$ *defines a distance on* $\mathcal{D}(\mathcal{H})$.

Proof. Symmetry and non-negativity are obvious by definition. If $W_{1,\mathcal{L}}(\rho, \omega) = 0$, then for all $X \in \mathcal{B}(\mathcal{H})$, $\mathrm{Tr}(X(\rho - \omega)) = 0$, which implies $\rho = \omega$. Finally, let $\rho, \omega, \tau \in \mathcal{D}(\mathcal{H})$. Then, for any $X \in \mathcal{B}(\mathcal{H})$ such that $\|X\|_{\mathrm{Lip}} \leq 1$, we have by the triangle inequality

$$|\mathrm{Tr}(X(\rho - \tau))| \leq |\mathrm{Tr}(X(\rho - \sigma))| + |\mathrm{Tr}(X(\sigma - \tau))| \leq W_{1,\mathcal{L}}(\rho, \tau) + W_{1,\mathcal{L}}(\sigma, \tau).$$

The result follows by taking the supremum over such operators X on the right hand side of the above inequality. $\qquad \square$

Next proposition justifies the definition of the quantum Liptchitz norm:

Lemma 12.7.2. *For any* $\rho, \omega \in \mathcal{D}_+(\mathcal{H})$,

$$\sqrt{2} \, W_{1,\mathcal{L}}(\rho, \omega) \leq W_{2,\mathcal{L}}(\rho, \omega).$$

Proof. The proof is a direct consequence of inequality (12.18): for any $X \in \mathcal{B}(\mathcal{H})$:

$$\sqrt{2} \, |\mathrm{Tr}(X(\rho - \omega))| \leq \left(\sum_{j \in \mathcal{J}} (e^{-\omega_j/2} + e^{\omega_j/2}) \|\nabla_{\tilde{L}_j} X\|_\infty^2 \right)^{1/2} W_{2,\mathcal{L}}(\rho, \omega). \tag{12.75}$$

$$\square$$

Other possible definitions for $W_{1,\mathcal{L}}$ Notice that the expression (12.73) depends on the choice of Lindblad operators. An alternative definition of the quantum Lipschitz constant which is independent of the representation is as follows:

$$\|X\|_{\mathrm{Lip},2} := \left(\sum_{j \in \mathcal{J}} (e^{-\omega_j/2} + e^{\omega_j/2}) \|\nabla_{\tilde{L}_j} X\|_2^2 \right)^{1/2}.$$

$\|X\|_{\mathrm{Lip},2}$ does not depend on the representation $(\{\tilde{L}_j\}, \{\omega_j\})$ in (12.1). Indeed, given two such representations $(\{\tilde{L}_j\}, \{\omega_j\})$ and $(\{\tilde{L}'_j\}, \{\omega'_j\})$,

$$
\begin{aligned}
\sum_{j \in \mathcal{J}} (\mathrm{e}^{-\omega'_j/2} + \mathrm{e}^{\omega'_j/2}) \|\nabla_{\tilde{L}'_j} X\|_2^2 &= \sum_{j \in \mathcal{J}} (\mathrm{e}^{-\omega'_j/2} + \mathrm{e}^{\omega'_j/2}) \operatorname{Tr}([X^*, (\tilde{L}'_j)^*][\tilde{L}'_j, X]) \\
&= \sum_{j,l,k \in \mathcal{J}} (\mathrm{e}^{-\omega'_j/2} + \mathrm{e}^{\omega'_j/2}) U_{jk} \overline{U}_{jl} \operatorname{Tr}([X^*, \tilde{L}_l^*][\tilde{L}_k, X]) \\
&= \sum_{j,l,k \in \mathcal{J}} U_{jk} \overline{U}_{jl} (\mathrm{e}^{-\omega_k/2} + \mathrm{e}^{\omega_k/2}) \operatorname{Tr}([X^*, \tilde{L}_l^*][\tilde{L}_k, X]) \\
&= \sum_{l,k \in \mathcal{J}} \delta_{kl}(\mathrm{e}^{-\omega_k/2} + \mathrm{e}^{\omega_k/2}) \operatorname{Tr}([X^*, \tilde{L}_l^*][\tilde{L}_k, X]) \\
&= \sum_{l \in \mathcal{J}} (\mathrm{e}^{-\omega_l/2} + \mathrm{e}^{\omega_l/2}) \operatorname{Tr}([X^*, \tilde{L}_l^*][\tilde{L}_l, X]) \\
&= \sum_{j \in \mathcal{J}} (\mathrm{e}^{-\omega_j/2} + \mathrm{e}^{\omega_j/2}) \|\nabla_{\tilde{L}_j} X\|_2^2 \,,
\end{aligned}
$$

where we used Remark 12.1.1 in the above lines. Moreover, since $\|.\|_\infty \le \|.\|_2$, it follows that $\|X\|_{\mathrm{Lip}} \le \|X\|_{\mathrm{Lip},2}$. One can also defined the Wasserstein distance of order 1 corresponding to this Lipschitz norm:

$$
W_{1,\mathcal{L},2}(\rho, \omega) := \sup_{\|X\|_{\mathrm{Lip},2} \le 1} |\operatorname{Tr} X(\rho - \omega)|, \tag{12.76}
$$

for which Lemma 12.7.1 extends, i.e. $W_{1,\mathcal{L},2}$ defines a distance on $\mathcal{D}_+(\mathcal{H})$. The following proposition follows directly from the direct observation that $\|X\|_{\mathrm{Lip}} \le \|X\|_{\mathrm{Lip},2}$.

Proposition 12.7.3. *For any $\rho, \omega \in \mathcal{D}_+(\mathcal{H})$,*

$$
W_{1,\mathcal{L},2}(\rho, \omega) \le W_{1,\mathcal{L}}(\rho, \omega) \,.
$$

12.7.2. Transportation cost inequality and Gaussian concentration

A primitive QMS $(\mathcal{P}_t)_{t \ge 0}$ in σ-DBC is said to satisfy a *transportation cost inequality of order* 1 with constant $c_1 > 0$ if for all $\rho \in \mathcal{D}_+(\mathcal{H})$

$$
W_{1,\mathcal{L}}(\rho, \sigma) \le \sqrt{2 c_1 D(\rho \| \sigma)} \,. \tag{$\mathrm{TC}_1(c_1)$}
$$

The following result is a direct consequence of Lemma 12.7.2:

Theorem 12.7.4. *If $(\mathcal{P}_t)_{t \ge 0}$ satisfies $\mathrm{TC}_2(c_2)$, then it satisfies $\mathrm{TC}_1(c_1)$ with $2c_1 = c_2$.*

The first proof that the classical transportation cost inequality of order 1 implies Gaussian concentration is due to Marton [Marton, 1996a]. The following theorem is a quantum generalization of Bobkov-Götze's proof [Bobkov and Goetze, 1999] which relies on the variational representation of the 1 Wasserstein distance (see also Theorem 36 of [Raginsky and Sason, 2014] or Proposition 7.7 of [Erbar and Maas, 2012]):

Theorem 12.7.5. *Let $(\mathcal{P}_t)_{t \ge 0}$ be a primitive QMS on $\mathcal{B}(\mathcal{H})$ in σ-DBC. If $(\mathcal{P}_t)_{t \ge 0}$ satisfies $\mathrm{TC}_1(c_1)$, the following quantum Gaussian concentration inequality holds: for any $X \in \mathcal{B}_{sa}(\mathcal{H})$,*

$$
\operatorname{Tr}(\sigma \mathbb{1}_{[r, \infty)}(X - \operatorname{Tr}(\sigma X))) \le \exp\left(\frac{-r^2}{2 c_1 \|(\Delta_\sigma^{-1/2} X)\|_{\mathrm{Lip}}^2} \right). \tag{qGauss}
$$

Proof. Here we follow the lines of the proof of Theorem 36 of [Raginsky and Sason, 2014]. Let $Y \in \mathcal{B}(\mathcal{H})$ be such that $\text{Tr}(\sigma Y) = 0$ and $\|Y\|_{\text{Lip}} \le 1$. From $\text{TC}_1(c_1)$, we know that for any $\rho \in \mathcal{D}_+(\mathcal{H})$,

$$|\text{Tr}(\rho Y)| \le W_{1,\mathcal{L}}(\rho, \sigma) \le \sqrt{2\, c_1 D(\rho\|\sigma)}\,.$$

Next, from the fact that

$$\inf_{\theta>0}\left(\frac{a}{\theta} + \frac{b\theta}{2}\right) = \sqrt{2ab}$$

for any $a, b \ge 0$, we see that any such Y must satisfy the following bound for any $\theta > 0$:

$$|\text{Tr}(\rho Y)| \le \frac{1}{\theta}D(\rho\|\sigma) + \frac{c_1\theta}{2}\,.$$

Rearranging the terms, we obtain for any $\theta > 0$:

$$2\theta|\text{Tr}(\rho Y)| - c_1\theta^2 \le 2D(\rho\|\sigma) \le 2\hat{D}(\rho\|\sigma)\,, \tag{12.77}$$

where we have used (1.66), and $\hat{D}(\rho\|\sigma)$ is the maximal divergence defined through Equation (1.67). Define $\rho := \sigma^{1/2}\, e^{\theta X}\, \sigma^{1/2}/(\text{Tr}(\sigma\, e^{\theta X}))$, where $X \in \mathcal{B}_{sa}(\mathcal{H})$ is to be specified later. Hence Equation (12.77) becomes

$$2\theta|\text{Tr}(\rho Y)| - c_1\theta^2 \le 2\theta\frac{\text{Tr}(\sigma\, e^{\theta X}\, X)}{\text{Tr}(\sigma\, e^{\theta X})} - 2\ln(\text{Tr}(\sigma\, e^{\theta X}))\,.$$

Now for $X = \Delta_\sigma^{1/2}(Y)$, the last expression further simplifies into

$$-\frac{c_1\theta^2}{2} \le -\ln(\text{Tr}(\sigma\, e^{\theta X})) \quad \Rightarrow \quad M_X(\theta) \equiv \text{Tr}(\sigma\, e^{\theta X}) \le e^{c_1\theta^2/2}\,.$$

We conclude by Markov's inequality:

$$\text{Tr}(\sigma\mathbb{1}_{[r,\infty)}(X)) \le e^{-r\theta}\, M_X(\theta) \le e^{-r\theta}\, e^{c_1\theta^2/2}\,. \tag{12.78}$$

Optimizing over all $\theta > 0$,

$$\text{Tr}(\sigma\mathbb{1}_{[r,\infty)}(X)) \le e^{-\frac{r^2}{2c_1}}\,.$$

In order to achieve this bound we assumed that $Y = \Delta_\sigma^{-1/2}(X) \in \mathcal{B}(\mathcal{H})$ is such that $\|Y\|_{\text{Lip}} \le 1$ and $\text{Tr}(\sigma Y) = 0$. This implies that $\text{Tr}(\sigma X) = \text{Tr}(\sigma\Delta_\sigma^{1/2}(Y)) = \text{Tr}(\sigma Y) = 0$. The result follows by rescaling. $\qquad\square$

12.7.3. Poincaré inequality and exponential concentration

The following theorem is a generalization of the classical results of [Gromov and Milman, 1983] (see also the review [Milman, 2009b]). It states that the Poincaré inequality implies exponential concentration. For this, we need the following well-known chain rule (see e.g. [Carlen and Maas, 2017] Lemma 5.5):

Lemma 12.7.6. *For all $L \in \mathcal{B}(\mathcal{H})$, and all $X \in \mathcal{B}(\mathcal{H})$,*

$$[L, e^X] = \int_0^1 e^{sX}\, [L, X]\, e^{X(1-s)}\, ds\,.$$

Proof.

$$\frac{d}{ds}\,\mathrm{e}^{(1-s)X}\,L\,\mathrm{e}^{sX} = \mathrm{e}^{(1-s)X}[L,X]\,\mathrm{e}^{sX}.$$

The result follows by integrating the above equation from 0 to 1. $\qquad\square$

Theorem 12.7.7. *Let $(\mathcal{P}_t)_{t\geq 0}$ be a primitive quantum Markov semigroup on $\mathcal{B}(\mathcal{H})$ satisfying σ-DBC. Then for any $X \in \mathcal{B}_{sa}(\mathcal{H})$,*

$$\mathrm{Tr}(\sigma\mathbb{1}_{[r,\infty)}(X - \mathrm{Tr}(\sigma X))) \leq 3\,\mathrm{e}^{-r\sqrt{\lambda(\mathcal{L})}/(\|X\|_{\mathrm{Lip}}C_{X,\lambda(\mathcal{L})})}.$$

where $\|.\|_{\mathrm{Lip}}$ is defined in Equation (12.73), and $C_{X,\lambda(\mathcal{L})} \equiv \dfrac{\mathrm{e}^{2\sqrt{\lambda(\mathcal{L})}\|X\|_\infty/\|X\|_{\mathrm{Lip}}}-1}{\sqrt{2\lambda(\mathcal{L})}\|X\|_\infty/\|X\|_{\mathrm{Lip}}}.$

Proof. Assume without loss of generality that $\mathrm{Tr}(\sigma X) = 0$. For $\theta \geq 0$, and $X \neq 0$ self-adjoint, let $M_X(\theta) := \mathrm{Tr}\left(\sigma\,\mathrm{e}^{\theta X}\right)$. Since $(\mathcal{P}_t)_{t\geq 0}$ is self-adjoint with respect to $\langle .,.\rangle_{1,\sigma}$, $\lambda(\mathcal{L})$ also satisfies the following inequality:

$$\lambda(\mathcal{L})\left(M_X(\theta) - M_X(\theta/2)^2\right) \leq -\langle\mathrm{e}^{\theta X/2},\mathcal{L}\left(\mathrm{e}^{\theta X/2}\right)\rangle_{1,\sigma}. \tag{12.79}$$

Moreover, from Equation (12.1):

$$-\langle\mathrm{e}^{\theta X/2},\mathcal{L}\left(\mathrm{e}^{\theta X/2}\right)\rangle_{1,\sigma} = \sum_{j\in\mathcal{J}}\mathrm{e}^{-\omega_j/2}\mathrm{Tr}\left[\sigma\left(\nabla_{\tilde{L}_j}\mathrm{e}^{\theta X/2}\right)^*\nabla_{\tilde{L}_j}\mathrm{e}^{\theta X/2}\right]$$

$$= \frac{\theta^2}{4}\sum_{j\in\mathcal{J}}\mathrm{e}^{-\omega_j/2}\iint_{[0,1]^2}\mathrm{Tr}\left(\sigma\,\mathrm{e}^{\frac{(1-s)\theta X}{2}}(\nabla_{\tilde{L}_j}X)^*\,\mathrm{e}^{\frac{s\theta X}{2}}\,\mathrm{e}^{\frac{u\theta X}{2}}\nabla_{\tilde{L}_j}X\,\mathrm{e}^{\frac{(1-u)\theta X}{2}}\right)duds. \tag{12.80}$$

where we used Lemma 12.7.6 in the second line. Moreover, for each $u,s \in [0,1]$, the trace in Equation (12.80) is equal, by cyclicity, to

$$\mathrm{Tr}\left[\left(\mathrm{e}^{\theta X/2}\sigma\,\mathrm{e}^{\theta X/2}\right)^*\left(\mathrm{e}^{\frac{-s\theta X}{2}}(\nabla_{\tilde{L}_j}X)^*\,\mathrm{e}^{\frac{(s+u)\theta X}{2}}\nabla_{\tilde{L}_j}X\,\mathrm{e}^{\frac{-u\theta X}{2}}\right)\right] \leq \mathrm{Tr}\left(\mathrm{e}^{\theta X/2}\sigma\,\mathrm{e}^{\theta X/2}\right)\mathrm{e}^{\theta(s+u)\|X\|_\infty}\|\nabla_{\tilde{L}_j}X\|_\infty^2$$

$$= M_X(\theta)\,\mathrm{e}^{\theta(s+u)\|X\|_\infty}\|\nabla_{\tilde{L}_j}X\|_\infty^2,$$

where we used Hölder's inequality as well as the submultiplicativity of the operator norm in the second line. Substituting into (12.80), we thus get:

$$-\langle\mathrm{e}^{\theta X/2},\mathcal{L}\left(\mathrm{e}^{\theta X/2}\right)\rangle_{1,\sigma} \leq \frac{M_X(\theta)}{4\|X\|_\infty^2}\sum_{j\in\mathcal{J}}\mathrm{e}^{-\omega_j/2}\|\nabla_{\tilde{L}_j}X\|_\infty^2\left(\mathrm{e}^{\theta\|X\|_\infty}-1\right)^2$$

$$\leq \|X\|_{\mathrm{Lip}}^2\,M_X(\theta)\frac{\left(\mathrm{e}^{\theta\|X\|_\infty}-1\right)^2}{4\|X\|_\infty^2},$$

However, for any $0 \leq \theta < 2\sqrt{\lambda(\mathcal{L})}/\|X\|_{\mathrm{Lip}}$,

$$\frac{\mathrm{e}^{\theta\|X\|_\infty}-1}{\theta\|X\|_\infty} \leq \frac{\mathrm{e}^{2\sqrt{\lambda(\mathcal{L})}\|X\|_\infty/\|X\|_{\mathrm{Lip}}}-1}{\sqrt{2\lambda(\mathcal{L})}\|X\|_\infty/\|X\|_{\mathrm{Lip}}} \equiv C_{X,\lambda(\mathcal{L})} > 1.$$

Hence, substituting into Equation (12.79):

$$\lambda(\mathcal{L})\left(M_X(\theta) - M_X(\theta/2)^2\right) \leq \theta^2\|X\|_{\mathrm{Lip}}^2\,C_{X,\lambda(\mathcal{L})}^2\,M_X(\theta)/4.$$

This last inequality implies that

$$M_X(\theta) \leq \frac{1}{1 - \theta^2 \|X\|_{\mathrm{Lip}}^2 \, C_{X,\lambda(\mathcal{L})}^2/(4\lambda(\mathcal{L}))} M_X(\theta/2)^2 \,,$$

for every $\theta < 2\sqrt{\lambda(\mathcal{L})}/(C_{X,\lambda(\mathcal{L})}\|X\|_{\mathrm{Lip}})$. A simple iteration procedure yields

$$M_X(\theta) \leq \prod_{k=0}^{n-1} \left(\frac{1}{1 - \theta^2 \|X\|_{\mathrm{Lip}}^2 \, C_{X,\lambda(\mathcal{L})}^2/(4^{k+1}\lambda(\mathcal{L}))} \right)^{2^k} M_X(\theta/2^n)^{2^n} \,.$$

Note that $M_X(\theta) = 1 + \theta \operatorname{Tr}(\sigma X) + \mathcal{O}(\theta^2)$, and we have assumed that $\operatorname{Tr}(\sigma X) = 0$. Thus letting $n \to \infty$:

$$M_X(\theta) \leq \prod_{k=0}^{\infty} \left(\frac{1}{1 - \theta^2 \|X\|_{\mathrm{Lip}}^2 \, C_{X,\lambda(\mathcal{L})}^2/(4^{k+1}\lambda(\mathcal{L}))} \right)^{2^k} \,.$$

Set $\theta = \sqrt{\lambda(\mathcal{L})}/(\|X\|_{\mathrm{Lip}}C_{X,\lambda(\mathcal{L})})$, then the right hand side is a universal constant contained between e and 3. So we proved that

$$M_X\left(\sqrt{\lambda(\mathcal{L})}/(\|X\|_{\mathrm{Lip}}C_{X,\lambda(\mathcal{L})})\right) \leq 3 \,.$$

Now by functional calculus, for any $r \in \mathbb{R}$ and $\theta > 0$:

$$\mathbb{1}_{[r,\infty)}(X) = \mathbb{1}_{[\exp(\theta r),\infty)}(\exp(\theta X)) \leq \mathrm{e}^{-\theta r}\,\mathrm{e}^{\theta X} \,.$$

This leads to the following Markov-type inequality:

$$\operatorname{Tr}(\sigma \mathbb{1}_{[r,\infty)}(X)) \leq \mathrm{e}^{-r\theta} \operatorname{Tr}(\sigma \exp(\theta X)) = \mathrm{e}^{-r\theta} M_X(\theta). \tag{12.81}$$

Therefore

$$\operatorname{Tr}(\sigma \mathbb{1}_{[r,\infty)}(X)) \leq 3\,\mathrm{e}^{-r\sqrt{\lambda(\mathcal{L})}/(\|X\|_{\mathrm{Lip}}C_{X,\lambda(\mathcal{L})})} \,.$$

$\square$

12.7.4. Concentration of product states

As already mentioned in (10.3), [Temme et al., 2014] proved that the 2-logarithmic Sobolev constant satisfies

$$\frac{\min_k \lambda_k}{\ln(d_{\mathcal{H}}^4 \, \max_k \|\sigma_k^{-1}\|_\infty) + 11} \leq \alpha_2(\mathcal{K}_N) \leq \frac{\min_k \lambda_k}{2} \,. \tag{12.82}$$

Moreover, it was shown in Proposition 13 of [Kastoryano and Temme, 2013] that the generator of a primitive semigroup satisfies the following inequality:

$$\alpha_2(\mathcal{L}) \leq \alpha_1(\mathcal{L}) \tag{12.83}$$

provided it is strongly $\mathbb{L}_p$-regular, which is always the case for semigroups satisfying a detailed balance condition by the Stroock Varopoulos inequality. Therefore, by a joint use of Theorems 12.5.6, 12.7.4 and 12.7.5 as well as 12.82 and 12.83, the following holds true: for any self-adjoint operator X_N on

$\mathcal{H}^{\otimes N}$,

$$\text{Tr}(\sigma^{(N)}\mathbb{1}_{[r,\infty)}(X_N - \text{Tr}(\sigma^{(N)}X_N))) \leq \exp\left(\frac{-2\min_k \lambda_k\, r^2}{\left(\ln(d_{\mathcal{H}}^4\max_k\|\sigma_k^{-1}\|_\infty)+11\right)\|(\Delta_{\sigma^{(N)}}^{-1/2}X_N)\|_{\text{Lip}}^2}\right). \quad (12.84)$$

where $\sigma^{(N)} := \otimes_{k=1}^N \sigma_k$. Assume now that X_N has the following form:

$$X_N := \frac{1}{N}\sum_{k=1}^N \mathbb{1}^{\otimes(k-1)} \otimes X \otimes \mathbb{1}^{\otimes(N-k)}.$$

and that the generators $\mathcal{L}_k$ are all identical, with associated invariant state σ. Then, it is easy to show that

$$\|(\Delta_{\sigma^{(N)}}^{-1/2}X_N)\|_{\text{Lip}} = \frac{1}{\sqrt{N}}\|(\Delta_\sigma^{-1/2}X)\|_{\text{Lip}}. \quad (12.85)$$

In this case, (12.84) reduces to

$$\text{Tr}(\sigma^{(N)}\mathbb{1}_{[r,\infty)}(X_N - \text{Tr}(\sigma^{(N)}X_N))) \leq \exp\left(\frac{-2N\,\min_k \lambda_k\, r^2}{\left(\ln(d_{\mathcal{H}}^4\max_k\|\sigma_k^{-1}\|_\infty)+11\right)\|(\Delta_\sigma^{-1/2}X)\|_{\text{Lip}}^2}\right).$$

In the case of the generalized depolarizing semigroup, the quasi-tensorization of the MLSI constant found in Theorem 10.1.1 provides the simpler bound:

$$\text{Tr}(\sigma^{(N)}\mathbb{1}_{[r,\infty)}(X_N - \text{Tr}(\sigma^{(N)}X_N))) \leq \exp\left(\frac{-r^2}{2\|(\Delta_{\sigma^{(N)}}^{-1/2}X_N)\|_{\text{Lip}}^2}\right). \quad (12.86)$$

12.8. Excursion to infinite dimensions

Here, we aim at defining a quantum Wasserstein distance $W_{2,\mathcal{L}^{\text{qOU}}}$ associated to the quantum Ornstein Uhlenbeck semigroup $(\mathcal{P}_t^{\text{qOU}})_{t\geq 0}$ of Section 5.5.2. In the general classical setting of Theorem 4.2.1, one assumes that the curves $(\mu_t)_{t\geq 0}$ are absolutely continuous on the Wasserstein space $(\mathcal{P}_2(\mathcal{M}), W_2)$. This assumption is essential in the proof of the existence and uniqueness of a tangent vector $\nabla\varphi \in \mathbb{L}_2(\mu, T_\mu(\mathcal{P}_2(\mathcal{M})))$ satisfying the continuity equation. For a compact metric space, a similar proof can be carried out for regular enough paths of absolutely continuous measures with associated densities in $\mathbb{L}_2(\mathcal{M})$, without the need to use any property of the Wasserstein distance. In particular, one can show that the Poisson equation admits a unique weak solution by means of Poincaré estimates. This proof has to be modified in the case when $\mathcal{M} = \mathbb{R}^n$ is unbounded, since such estimates typically diverge. In fact, in this case, there exist multiple (possibly distributional) solutions to the Poisson equation, due to the degeneracy of the Laplace operator. We briefly recall how to construct a solution in the Sobolev space

$$\mathcal{W}^{1,2}(\mathbb{R}^n) := \overline{\{f \in C_c^\infty(\mathbb{R}^n)\}}^{\|\cdot\|_{\mathcal{W}^{1,2}(\mathbb{R}^n)}} \subset \mathbb{L}_2(\mathbb{R}^n),$$

that is, the closure of the space $C_c^\infty(\mathbb{R})$ of smooth, compactly supported functions on $\mathbb{R}^n$ in the norm $\|\cdot\|_{\mathcal{W}^{1,2}(\mathbb{R}^n)}$ defined as

$$\|f\|_{\mathcal{W}^{1,2}(\mathcal{H})} := \left(\sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|\leq 1}} \|\partial^\alpha f\|_{\mathbb{L}_2(\mathbb{R}^n)}^2\right)^{1/2}.$$

let $f \in \mathbb{L}_2(\mathbb{R}^n)$, and consider the Poisson equation on $\mathbb{R}^n$:

$$\Delta h = f . \tag{12.87}$$

A standard Green function argument leads to the following solution in $C_2(\mathbb{R}^n)$:

$$h(\mathbf{x}) = - \int_{\mathbb{R}^n} \nu(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) \, d\mathbf{y} , \tag{12.88}$$

where

$$\nu(\mathbf{x}, \mathbf{y}) := \begin{cases} -\dfrac{1}{2}|\mathbf{x} - \mathbf{y}| & n = 1 , \\[2mm] -\dfrac{1}{4\pi} \ln(|\mathbf{x} - \mathbf{y}|^2) & n = 2 , \\[2mm] -\dfrac{1}{(2-n) \, S_n(1)} |\mathbf{x} - \mathbf{y}|^{2-n} & n \geq 3 , \end{cases}$$

where $S_n(1)$ denotes the surface of the unit ball in $\mathbb{R}^n$. Moreover, by a Riezs transform argument (see Proposition 3 p. 59 of [Stein, 1970]), this solution is necessarily in the Sobolev space $H^2(\mathbb{R}^n)$ (i.e. its partial derivatives are in $\mathbb{L}^2(\mathbb{R}^n)$ up to second order). We use this solution together with Theorem 0.2.3 in order to find a regular solution to the quantum Poisson equation. For sake of simplicity, we reduce our analysis to the case $n = 2$, which corresponds to a 1-mode bosonic system. Let $\mathcal{A}_0 := \mathrm{span}\{|e_i\rangle\langle e_j|, \, i, j \in \mathbb{N}\}$, where $\{e_i\}_{i \in \mathbb{N}}$ denotes the eigenbasis of the number operator N. Then, given $s, p \in \mathbb{N}^*$, define the *quantum Sobolev space*

$$\mathcal{W}^{s,p} = \overline{\mathcal{A}_0}^{\|.\|_{\mathcal{W}^{s,p}(\mathcal{H})}} ,$$

where the norm $\|.\|_{\mathcal{W}^{s,p}(\mathcal{H})}$ is defined as expected:

$$\|X\|_{\mathcal{W}^{s,p}(\mathcal{H})} := \left(\sum_{t \leq s} \sum_{\beta^t \in \mathbb{Z}_2^t} \|\nabla_{\mathbf{a}}^{\beta^t} X\|_p^2 \right)^{1/2} ,$$

where $\nabla_{\mathbf{a}}^{\beta^t} X$ is a nested commutator of the form $[a_{\beta_1^t}, [a_{\beta_2^t}, [..., [a_{\beta_t^t}, X], ...]]]$, with $a_0 := a$ and $a_1 := a^*$.

Theorem 12.8.1. *Let $F \in \mathcal{T}_2(\mathcal{H})$, then the noncommutative Poisson equation*

$$\Delta(X) := [a^*, [a, X]] + [a, [a^*, X]] = F$$

admits a solution X in the quantum Sobolev space $\mathcal{W}^{2,2}(\mathcal{H})$.

Proof. Since $F \in \mathcal{T}_2(\mathcal{H})$, we can define the function $f := \mathcal{F}^{-1}\hat{f} = \mathcal{F}^{-1}(\mathcal{F}_F^{\mathrm{q}}) \in \mathbb{L}_2(\mathbb{R}^2)$, where $\mathcal{F}_F^{\mathrm{q}}$ is the quantum Fourier transform of F and $\mathcal{F}^{-1}$ the inverse classical Fourier transform. Then, the Poisson equation (12.87) admits a solution $h \in \mathcal{W}^{2,2}(\mathbb{R}^2)$. Its Fourier transform $\mathcal{F}(h)$ is in $\mathbb{L}_2(\mathbb{R}^2)$, and the following holds: $\mathbb{R}^2 \ni (x, y) \mapsto x\mathcal{F}(h)(z), \, y\mathcal{F}(h)(z), \, |z|^2\mathcal{F}(h)(z) \in \mathbb{L}_2(\mathbb{R}^2)$ and

$$|z|^2 \mathcal{F}(h)(z) = \hat{f}(z) .$$

The result follows by an argument similar to the one in the proof of Theorem 7.5.2, and choosing X as the operator whose quantum Fourier transform $\mathcal{F}_X$ is equal to $\mathcal{F}(h)$. $\qquad\square$

Back to our problem of defining the quantum Wasserstein distance $W_{2, \mathcal{L}^{\mathrm{qOU}}}$ associated to the quantum Ornstein Uhlenbeck semigroup $(\mathcal{P}_t^{\mathrm{qOU}})$, we let $\gamma : [0,1] \mapsto \mathcal{D}_+(\mathcal{H})$ be a path on the space of

faithful states on $\mathcal{H}$, and assume that for almost every $t \in [0,1]$, there exists an operator $\dot{\gamma}(t) \in \mathcal{T}_2(\mathcal{H})$ that is the weak derivative of γ at t: for any $X \in \mathcal{A}_0$,

$$\lim_{s \to t} \operatorname{Tr}\left(X \left[\frac{\gamma(s) - \gamma(t)}{s - t} - \dot{\gamma}(t) \right] \right) = 0. \tag{12.89}$$

This condition replaces the one of absolute continuity of the paths in the setting of Theorem 4.2.1. Last theorem implies the existence almost everywhere of $X(t) \in \mathcal{W}^{2,2}(\mathcal{H})$ such that for all $Y \in \mathcal{A}_0$,

$$\langle \nabla_a Y, \nabla_a X(t) \rangle_{\mathrm{HS}} + \langle \nabla_{a^*} Y, \nabla_{a^*} X(t) \rangle_{\mathrm{HS}} = \operatorname{Tr}(X \dot{\gamma}(t)) \qquad \text{a.e.}.$$

We aim at defining a quantum Wasserstein distance associated to the quantum Ornstein Uhlenbeck semigroup $(\mathcal{P}_t^{\mathrm{qOU}})_{t \geq 0}$, introduced in Section 5.5.2, and whose generator takes the following form on the subalgebra $\mathcal{A}_0$:

$$\mathcal{L}^{\mathrm{qOU}}(X) := -\frac{\mu^2}{2}(a^* a X - 2a^* X a + X a^* a) - \frac{\lambda^2}{2}(a a^* X - 2a X a^* + X a a^*).$$

The above generator is of the form of Equation (12.1) when choosing $\mathrm{e}^{-\omega_0/2} = \mu^2/2$ and $\mathrm{e}^{-\omega_1/2} = \lambda^2/2$.

Given $\rho \in \mathcal{D}_+(\mathcal{H})$ and $\varphi_\omega^{\log}$ defined as in Equation (12.6), define $\mathbb{L}_{2,\log}(\rho)$ as the completion of $\mathcal{A}_0 \times \mathcal{A}_0$ in the following norm: given $\mathbf{X} = (X,Y) \in \mathcal{A}_0 \times \mathcal{A}_0$ and $\boldsymbol{\omega} := (\omega_0, \omega_1) \in \mathbb{R}^2$:

$$\|\mathbf{X}\|_{\mathbb{L}_{2,\log}(\rho)} = \sqrt{\langle X, [\rho]_{\varphi^{\log}, \omega_0}(X) \rangle_{\mathrm{HS}} + \langle Y, [\rho]_{\varphi^{\log}, \omega_1}(Y) \rangle_{\mathrm{HS}}}.$$

Next, set $\mathbf{V}(t) := [\gamma]_{\varphi^{\log}, \boldsymbol{\omega}}^{-1}(\nabla_{\mathbf{a}} X(t))$, which is a well-defined operator in $\mathbb{L}_{2,\log}(\gamma(t))$ and weakly satisfies:

$$\operatorname{div}_{\mathbf{a}}[\gamma(t)]_{\varphi^{\log}, \boldsymbol{\omega}}(\mathbf{V}(t)) = \dot{\gamma}(t) \quad \text{a.e..} \qquad \text{(CCR-continuity equation)}$$

Now, the set $\{\mathbf{V}(t) + \mathbf{W}(t)\}$, where $\mathbf{W}(t) \in \mathbb{L}_{2,\log}(\gamma(t))$ is such that $[\gamma(t)]_{\varphi^{\log}, \boldsymbol{\omega}}(\mathbf{W}(t))$ is divergence-free in the weak sense, is a closed affine subspace of solutions of the CCR-continuity equation, in which there is a unique solution $\mathbf{V}_0(t)$ of minimal $\|.\|_{\mathbb{L}_{2,\log}(\gamma(t))}$ norm. Obviously, $\mathbf{V}_0(t)$ is orthogonal in $\mathbb{L}_{2,\log}(\gamma(t))$ to the set of such vector $\mathbf{W}(t)$. In analogy with in the classical case (4.3), the set of minimal solutions is given by

$$\overline{\{\nabla_{\mathbf{a}} \Phi : \ \Phi \in \mathcal{A}_0\}}^{\mathbb{L}_{2,\log}(\gamma(t))}$$
$$= \left\{ \mathbf{V} \in \mathbb{L}_{2,\log}(\gamma(t)) : \ \langle \mathbf{V}, \mathbf{W} \rangle_{\mathbb{L}_{2,\log}(\gamma(t))} = 0, \ \forall \mathbf{W} \in \mathbb{L}_{2,\log}(\gamma(t)) \text{ s.t. } \nabla_{\mathbf{a}} \cdot ([\gamma(t)]_{\varphi^{\log}, \boldsymbol{\omega}} \mathbf{W}) = 0 \right\}. \tag{12.90}$$

and we refer to it as the tangent space $T_{\gamma(t), \mathrm{CCR}} \mathcal{D}_+(\mathcal{H})$ at $\gamma(t) \in \mathcal{D}_+(\mathcal{H})$.

We are finally in a position of defining the quantum Wasserstein distance associated to the quantum Ornstein Uhlenbeck semigroup. We will call it the CCR Wasserstein distance since it simply involves the noncommutative derivations associated to the quantum phase space annihilation and creation operators a and a^*.

Definition 12.8.2. For any two density operators $\gamma_0, \gamma_1 \in \mathcal{D}_+(\mathcal{H})$, the CCR *Wasserstein distance of order* 2 between γ_0 and γ_1 is defined as

$$W_{2,\mathrm{CCR}}(\gamma_0, \gamma_1) := \inf_\gamma \left\{ \int_0^1 \|\mathbf{V}(s)\|_{\mathbb{L}_{2,\log}(\gamma(s))}^2 \, ds \right\}^{\frac{1}{2}},$$

where the infimum is taken over the paths $(\gamma(t), \mathbf{V}(t))_{t \epsilon [0,1]}$, where $\gamma : [0,1] \to \mathcal{D}_+(\mathcal{H})$ has a.e. weak derivative in the sense of Equation (12.89), $\gamma_0 = \gamma(0)$, $\gamma_1 = \gamma(1)$ and $\mathbf{V}(t)$ is the only vector in $T_{\gamma(t),\mathrm{CCR}}\mathcal{D}_+(\mathcal{H})$ so that the CCR-continuity equation is satisfied a.e..

Part V.

Applications

Chapter 13.

Quantum statistics

13.1. Quantum Hypothesis testing

Quantum hypothesis testing concerns the problem of discriminating between two different quantum states[1]. This task is of paramount importance in quantum information theory, since many other tasks can be reduced to it. In the language of hypothesis testing, one considers two hypotheses – the *null hypothesis* $H_0 : \rho$ and the *alternative hypothesis* $H_1 : \sigma$, where ρ and σ are two quantum states. In an operational setting, say Bob receives a state ω with the knowledge that either $\omega = \rho$ or $\omega = \sigma$. His goal is then to infer which hypothesis is true, i.e., which state he has been given, by means of a measurement on the state he receives. The measurement is given most generally by a POVM $\{T, \mathbb{1} - T\}$ where $0 \leq T \leq \mathbb{1}$. Adopting the nomenclature from classical hypothesis testing, we refer to T as a *test*. The probability that Bob correctly guesses the state to be ρ is then equal to $\mathrm{Tr}(T\rho)$, whereas his probability of correctly guessing the state to be σ is $\mathrm{Tr}((\mathbb{1} - T)\sigma)$. Bob can erroneously infer the state to be σ when it is actually ρ or vice versa. The corresponding error probabilities are referred to as the *type I error* and *type II error* respectively. They are denoted as follows:

$$\alpha(T) \coloneqq \mathrm{Tr}\left((\mathbb{1} - T)\rho\right), \quad \beta(T) \coloneqq \mathrm{Tr}\left(T\sigma\right), \tag{13.1}$$

where $\alpha(T)$ is the probability of accepting H_1 when H_0 is true, while $\beta(T)$ is the probability of accepting H_0 when H_1 is true. Obviously, there is a trade-off between the two error probabilities, and there are various ways to jointly optimize them, depending on whether or not the two types of errors are treated on an equal footing. In the setting of *symmetric hypothesis testing*, one minimizes the total probability of error $\alpha(T) + \beta(T)$, whereas in *asymmetric hypothesis testing* one minimizes the type II error under a suitable constraint on the type I error.

13.1.1. Summary of known results

Quantum hypothesis testing was originally studied in the *asymptotic i.i.d. setting* in which Bob is provided not with just a single copy of the state but with multiple (say n) identical copies of the state, say $\rho^{\otimes n}$ or $\sigma^{\otimes n}$, where ρ and σ are states on a finite dimensional Hilbert space $\mathcal{H}$, and he is allowed to do a joint measurement on all these copies. The optimal asymptotic performance in the different settings is quantified by the following exponential decay rates, evaluated in the limit $n \to \infty$:

---

[1]It is often referred to as *binary* quantum hypothesis testing, to distinguish it from the case in which more than two states are being tested.

The Chernoff bound in symmetric QHT   the optimal exponential decay rate of the sum of type I and type II errors. This corresponds to the *symmetric setting* and is given by the *quantum Chernoff distance* [Audenaert et al., 2007, Nussbaum and Szkoła, 2009]: for any $\lambda, \kappa > 0$,

$$-\lim_{n \to \infty} \frac{1}{n} \ln \min_{0 \le T_n \le \mathbb{1}_n} \{\kappa \alpha(T_n) + \lambda \beta(T_n)\} = -\lim_{n \to \infty} \frac{1}{n} \ln \min_{0 \le T_n \le \mathbb{1}_n} \{\alpha(T_n) + \beta(T_n)\} = -\inf_{0 \le s \le 1} (s-1) D_s(\rho \| \sigma).$$

Here $\alpha(T_n) = \text{Tr}[(\mathbb{1}_n - T_n)\rho^{\otimes n}]$ and $\beta(T_n) = \text{Tr}[T_n \sigma^{\otimes n}]$, with $\mathbb{1}_n$ being the identity operator acting on $\mathcal{H}^{\otimes n}$.

Stein's lemma and its refinements   the optimal exponential decay rate of the type II error under the assumption that the type I error remains bounded. This is given by *Stein*'s lemma and its refinements [Hiai and Petz, 1991, Ogawa and Nagaoka, 2000]: for any $\varepsilon \in (0, 1)$,

$$-\lim_{n \to \infty} \frac{1}{n} \ln \min_{0 \le T_n \le \mathbb{1}_n} \{\beta(T_n) : \alpha(T_n) \le \varepsilon\} = D(\rho \| \sigma).$$

Hence, the minimal asymptotic type I error jumps discontinuously from 0 to 1 as the asymptotic type II error exponent crosses the value $D(\rho \| \sigma)$ from below. However, this discontinuous dependence of the minimal asymptotic type I error on the asymptotic type II error exponent is a manifestation of the coarse-grained analysis underlying the Quantum Stein's lemma, in which only the linear term (in n) of the *type II error exponent* $(-\ln \beta(T_n))$ is considered. More recently, [Li, 2014] and [Tomamichel and Hayashi, 2012] independently showed that this discontinuity vanishes under a more refined analysis of the type II error exponent, in which its second order (i.e. order $\sqrt{n}$) term is retained, in addition to the linear term. This analysis is referred to as the *second order asymptotics* for (asymmetric) quantum hypothesis testing, since it involves the evaluation of $(-\ln \beta_n(\varepsilon))$ up to second order. It was proved in [Tomamichel and Hayashi, 2012, Li, 2014] to be given by:

$$-\ln \min_{0 \le T_n \le \mathbb{1}_n} \{\beta(T_n) : \alpha(T_n) \le \varepsilon\} = n D(\rho \| \sigma) + \sqrt{n V(\rho \| \sigma)} \, \Phi^{-1}(\varepsilon) + \mathcal{O}(\ln n), \tag{13.2}$$

where Φ denotes the cumulative distribution function (c.d.f.) of a standard normal distribution, and $V(\rho \| \sigma) := \text{Tr}[\rho (\ln \rho - \ln \sigma)^2] - D(\rho \| \sigma)^2$ is the so-called *quantum information variance*. The Gaussian c.d.f. Φ arises from the central limit theorem, or rather from its refinement, the Berry-Esseen theorem (see e.g. [Feller, 2008]), which gives the rate of convergence of the distribution of the scaled sum of i.i.d. random variables to a normal distribution.

The Hoeffding bound   On the other hand if we require the type II error probabilities to vanish with an exponent below the relative entropy, the type I error is given by the *Hoeffding distance* [Hayashi, 2007, Nagaoka, 2006, Ogawa and Hayashi, 2004]

$$\sup_{0 \le T_n \le \mathbb{1}_n} \left\{ -\limsup_{n \to \infty} \frac{1}{n} \ln \alpha(T_n) : \limsup_{n \to \infty} \frac{1}{n} \ln \beta(T_n) \le -r \right\} = \sup_{0 < s < 1} \frac{s-1}{s} \left[r - D_s(\rho \| \sigma) \right],$$

which provides Rényi divergences with an operational interpretation.

The converse Hoeffding bound   [Ogawa and Nagaoka, 2000] actually showed a strong converse result according to which, if the type II error is constrained to vanish exponentially fast with a rate that is higher than Stein's exponent $(r > D(\rho \| \sigma))$, the type I error diverges to 1 exponentially fast. More recently, [Mosonyi and Ogawa, 2015] quantified this claim by proving what is the following

strong converse exponent:

$$\inf_{0 \leq T_n \leq \mathbb{1}_n} \left\{ -\liminf_{n \to \infty} \frac{1}{n} \ln(1 - \alpha(T_n)) : \limsup_{n \to \infty} \frac{1}{n} \ln \beta(T_n) \leq -r \right\} = \sup_{1 < s} \frac{s-1}{s} \left[r - \tilde{D}_s(\rho \| \sigma) \right],$$

hence providing sandwiched Rényi divergences with an operational interpretation.

13.1.2. Finite sample size strong converse bounds

All the results that we listed above are asymptotic in the sense that they hold when the number of copies of the unknown state that are being tested goes to infinity. However, in a more practical situation, one might be interested in getting estimates on the errors made when a finite number n of copies are available. This is the so-called *finite blocklength regime*. In this section we are interested in obtaining a finite sample size strong converse bound on the rate of convergence of $\alpha(T_n)$ as a function of n. Inspired by a recent article of [Liu et al., 2017], we use reverse hypercontractivity in order to obtain our bound. Before stating and proving the main theorem of this section, we establish an easy inequality that will be used in the proof.

Lemma 13.1.1. *Let $(\mathcal{P}_t)_{t \geq 0}$ be a a primitive QMS that satisfies σ-DBC. Let $X, Y > 0$ and $-\infty \leq q, p \leq 1$. Then, for any $t \geq 0$ such that $(1-p)(1-q) \geq e^{-4\alpha_1(\mathcal{L})t}$:*

$$\langle X, \mathcal{P}_t(Y) \rangle_\sigma \geq \|X\|_{\mathbb{L}_p(\sigma)} \|Y\|_{\mathbb{L}_q(\sigma)}.$$

Proof. The result follows by a direct application of Lemma 7.4.1 together with the reverse hypercontractivity inequality in Corollary 7.4.10.

The i.i.d. setting Our main result, from which a bound for the finite blocklength strong converse rate follows directly as a corollary, is the following second order upper bound on the type 2 error exponent in the i.i.d. setting.

Theorem 13.1.2. *Let $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ full-rank density matrices.[2] Then for any test $0 \leq T_n \leq \mathbb{1}_n$:*

$$-\frac{1}{n} \ln \operatorname{Tr}(\sigma^{\otimes n} T_n) \leq D(\rho \| \sigma) + \frac{2}{\sqrt{n}} \sqrt{\left\| \sigma^{-\frac{1}{2}} \rho \sigma^{-\frac{1}{2}} \right\|_\infty \ln \frac{1}{\operatorname{Tr}(\rho^{\otimes n} T_n)}} - \frac{1}{n} \ln \operatorname{Tr}(\rho^{\otimes n} T_n). \tag{13.3}$$

Proof. To simplify notations we use $\sigma_n := \sigma^{\otimes n}$ and $\rho_n := \rho^{\otimes n}$. Let $0 \leq p, q \leq 1$ and let $t \geq 0$ be such that

$$(1-p)(1-q) = e^{-t}. \tag{13.4}$$

Let $\mathcal{L}_\rho$ denote the generator of a generalized depolarizing semigroup $(\mathcal{P}_t^{\mathrm{depol},\rho})_{t \geq 0}$ with invariant state ρ, i.e., $\mathcal{P}_t^{\mathrm{depol},\rho}(X) = e^{-t} X + (1 - e^{-t}) \operatorname{Tr}(\rho X) \mathbb{1}$. By Theorem 10.1.1 the MLSI constants of this QMS and its tensor powers are lower bounded by $1/4$. Then using Lemma 13.1.1 for $Y = T_n$ and $X = \Gamma_{\rho_n}^{-1}(\sigma_n)$ we obtain

$$\operatorname{Tr}((\mathcal{P}_{t*}^{\mathrm{depol},\rho})^{\otimes n}(\sigma_n) T_n) = \operatorname{Tr}\left(\sigma_n (\mathcal{P}_t^{\mathrm{depol},\rho})^{\otimes n}(T_n) \right) \geq \left\| \Gamma_{\rho_n}^{-1}(\sigma_n) \right\|_{\mathbb{L}_p(\rho_n)} \|T_n\|_{\mathbb{L}_q(\rho_n)}. \tag{13.5}$$

---

[2]What we really need is that the supports of ρ and σ are the same (and not being the whole $\mathcal{H}$) since in this case we may restrict everything to this support.

An application of the Araki-Lieb-Thirring inequality, Lemma 1.2.8, with $A = \sigma_n$, $B = \rho_n^{(1-p)/p}$ and $r = p \in [0, 1]$ leads to

$$\left\|\Gamma_{\rho_n}^{-1}(\sigma_n)\right\|_{\mathbb{L}_p(\rho_n)} = \left[\mathrm{Tr}\left(\rho_n^{(1-p)/2p}\sigma_n\rho_n^{(1-p)/2p}\right)^p\right]^{1/p} \geq \left[\mathrm{Tr}\left(\rho_n^{1-p}\sigma_n^p\right)\right]^{1/p} = \exp\left(-D_{1-p}(\rho_n\|\sigma_n)\right),$$

where

$$D_{1-p}(\rho\|\sigma) := \frac{-1}{p}\ln\mathrm{Tr}\left(\sigma^p\rho^{1-p}\right),$$

denotes the quantum Rényi divergence of order p between ρ and σ. A very similar application of Lemma 1.2.8 for $A = T_n$ and $B = \rho_n^{1/q}$ and $r = q \in [0, 1]$ yields

$$\|T_n\|_{\mathbb{L}_q(\rho_n)} = \left[\mathrm{Tr}\left(\rho_n^{1/2q}T_n\rho_n^{1/2q}\right)^q\right]^{1/q} \geq \left[\mathrm{Tr}\left(\rho_n T_n^q\right)\right]^{1/q} \geq \left[\mathrm{Tr}\left(\rho_n T_n\right)\right]^{1/q},$$

where in the last inequality, we used that $0 \leq T_n \leq \mathbb{1}_n$, so that $T_n^q \geq T_n$. Using the last two bounds in (13.5), we get

$$\mathrm{Tr}(\sigma_n(\mathcal{P}_t^{\mathrm{depol},\rho})^{\otimes n}(T_n)) \geq \left[\mathrm{Tr}(\rho_n T_n)\right]^{1/q}\exp\left(-D_{1-p}(\rho_n\|\sigma_n)\right).$$

Taking the limit $p \to 0$ (and $q \to 1 - e^{-t}$) on both sides of the above inequality yields

$$\mathrm{Tr}(\sigma_n(\mathcal{P}_t^{\mathrm{depol},\rho})^{\otimes n}(T_n)) \geq \left[\mathrm{Tr}(\rho_n T_n)\right]^{1/(1-e^{-t})}\exp\left(-D(\rho_n\|\sigma_n)\right). \tag{13.6}$$

Let $\gamma := \|\sigma^{-\frac{1}{2}}\rho\sigma^{-\frac{1}{2}}\|_\infty \geq 1$. Then

$$\mathcal{P}_{t*}^{\mathrm{depol},\rho}(\sigma) = e^{-t}\sigma + (1 - e^{-t})\rho \leq (e^{-t} + \gamma(1 - e^{-t}))\sigma$$

Therefore,

$$\mathrm{Tr}((\mathcal{P}_{t*}^{\mathrm{depol},\rho})^{\otimes n}(\sigma_n)T_n) \leq (e^{-t} + \gamma(1 - e^{-t}))^n\mathrm{Tr}(T_n\sigma_n). \tag{13.7}$$

Next using $\gamma \geq 1$, the convexity of $h(x) = x^\gamma$ implies $(h(x) - h(1))/(x - 1) \geq h'(1)$ for every $x \geq 1$. Therefore, $e^{\gamma t} - 1 \geq \gamma(e^t - 1)$ for every $t \geq 0$, and $e^{-t} + \gamma(1 - e^{-t}) \leq e^{(\gamma-1)t}$. As a result

$$\mathrm{Tr}\left(\sigma_n(\mathcal{P}_t^{\mathrm{depol},\rho})^{\otimes n}(T_n)\right) \leq e^{(\gamma-1)nt}\mathrm{Tr}(\sigma_n T_n). \tag{13.8}$$

Then from (13.15) and (13.8) we get

$$\left[\mathrm{Tr}(\rho_n T_n)\right]^{1/(1-e^{-t})}\exp\left(-D(\rho_n\|\sigma_n)\right) \leq e^{(\gamma-1)nt}\mathrm{Tr}(\sigma_n T_n).$$

Taking the logarithm of both sides yields

$$\begin{aligned}
\ln\mathrm{Tr}(\sigma_n T_n) &\geq -D(\rho_n\|\sigma_n) - (\gamma - 1)nt + \frac{1}{1 - e^{-t}}\ln\mathrm{Tr}(\rho_n T_n) \\
&\geq -D(\rho_n\|\sigma_n) - \gamma nt + \left(1 + \frac{1}{t}\right)\ln\mathrm{Tr}(\rho_n T_n),
\end{aligned} \tag{13.9}$$

where the second inequality follows from $e^t \geq 1 + t$ and

$$\frac{1}{1 - e^{-t}} = 1 + \frac{1}{e^t - 1} \leq 1 + \frac{1}{t}.$$

Optimizing (13.9) over the choice of t yields $t = \left(\frac{-\ln \operatorname{Tr}(\rho_n T_n)}{\gamma n} \right)^{1/2}$, and we obtain the desired inequality.

□

Corollary 13.1.3 (Finite-blocklength strong converse bound for quantum hypothesis testing). *Let* $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ *full-rank density and* $\gamma = \|\sigma^{-\frac{1}{2}} \rho \sigma^{-\frac{1}{2}}\|_\infty$. *Then for any test* $0 \le T_n \le \mathbb{1}_n$, *where* $T_n \in \mathcal{B}(\mathcal{H}^{\otimes n})$, *if the Type II error satisfies the inequality* $\beta(T_n) \le \mathrm{e}^{-nr}$ *for* $r > D(\rho\|\sigma)$, *then the Type I error satisfies*

$$\alpha(T_n) \ge 1 - \mathrm{e}^{-nf}, \tag{13.10}$$

where

$$f = \left(\sqrt{\gamma + (r - D(\rho\|\sigma))} - \sqrt{\gamma} \right)^2,$$

and hence tends to zero in the limit of $r \to D(\rho\|\sigma)$.

Proof. Fix $r > D(\rho\|\sigma)$ and consider a sequence of tests T_n such that $\beta(T_n) \le \mathrm{e}^{-nr}$. Then, from Theorem 13.1.2 we have

$$-nr \ge -nD(\rho\|\sigma) - 2\sqrt{n\gamma \ln \frac{1}{1 - \alpha(T_n)}} - \ln \frac{1}{1 - \alpha(T_n)}.$$

Defining $x_n^2 := \ln \frac{1}{1-\alpha(T_n)}$ this is equivalent to

$$x_n^2 + 2\sqrt{n\gamma}\, x_n - n\left(r - D(\rho\|\sigma) \right) \ge 0,$$

solving which directly leads to the statement of the corollary.

Theorem 13.1.2 also leads to the following finite blocklength second order lower bound on the Type II error when the Type I error is less than a threshold value.

Corollary 13.1.4. *Let* $\rho, \sigma \in \mathcal{D}_+(\mathcal{H})$. *Then for any* $n \in \mathbb{N}$ *and* $\varepsilon > 0$ *the minimal Type II error satisfies*

$$-\ln \beta_n(\varepsilon) := -\ln \min_{0 \le T_n \le \mathbb{1}_n} \{ \beta(T_n) : \alpha(T_n) \le \varepsilon \} \le nD(\rho\|\sigma) + 2\sqrt{n\gamma \ln\left(\frac{1}{1-\varepsilon} \right)} - \ln(1-\varepsilon),$$

where $\gamma = \|\sigma^{-\frac{1}{2}} \rho \sigma^{-\frac{1}{2}}\|_\infty$.

Remark 13.1.5. The bound found in the above corollary can be compared to different bounds previously known: first, the following estimations were obtained in Theorem 3.3 of [Audenaert et al., 2012]:

$$nD(\rho\|\sigma) - 4\sqrt{2n} \ln(\varepsilon^{-1}) \ln \eta + 2\ln 2 \le -\ln \beta_n(\varepsilon) \le nD(\rho\|\sigma) + 4\sqrt{2n} \ln\left(\frac{1}{1-\varepsilon} \right) \ln \eta, \tag{13.11}$$

where $\eta := 1 + \mathrm{e}^{\frac{1}{2}D_{3/2}(\rho\|\sigma)} + \mathrm{e}^{-\frac{1}{2}D_{1/2}(\rho\|\sigma)}$. The upper bound in (13.11) is a simple consequence of the monotonicity of the Rényi divergences. Our bound is tighter for small values of ε. We also mention the following lower bound we recently found in [Rouzé and Datta, 2018] by means of classical martingale concentration inequalities in the spirit of [Sason, 2012]:

$$nD(\rho\|\sigma) - \sqrt{2n \ln \varepsilon^{-1}} \, \| \ln \Delta_{\sigma|\rho} + D(\rho\|\sigma) \operatorname{id} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}) \| \le -\ln \beta_n(\varepsilon),$$

which also constitutes an improvement of the lower bound in (13.11) for small ε.

Beyond the i.i.d. setting In Theorem 13.1.2, we established an upper bound on the type II error exponent in terms of the type I error and the Stein exponent in the i.i.d. setting. The technique

used to prove this result, namely reverse hypercontractivity of the generalized depolarizing semigroup converging towards the state ρ, can be extended as follows: assume given two sequences $\{\rho_n\}_{n\in\mathbb{N}}$ and $\{\sigma_n\}_{n\in\mathbb{N}}$, where for each n, $\rho_n, \sigma_n \in \mathcal{D}(\mathcal{H}^{\otimes n})$ are full-rank. Filtering out the two main inequalities at the core of the proof of Theorem 13.1.2, we need:

(i) a family of primitive semigroups $(\mathcal{P}_t^n = e^{t\mathcal{L}_n})_{t\geq 0}$, each of which satisfying the ρ_n-DBC, with MLSI constant $\alpha_1(\mathcal{L}_n)$ uniformly bounded away from 0 (cf. (13.5));

(ii) a constant $\gamma < \infty$ such that for all $n \in \mathbb{N}$, the state $\mathcal{P}_{t*}^n(\sigma_n) \leq e^{nt\gamma} \sigma_n$ (cf. (13.8)). In other words, we want for all $t \geq 0$ and any $n \in \mathbb{N}^*$:

$$\frac{1}{nt} D_{\max}(\mathcal{P}_{t*}^n(\sigma_n) \| \sigma_n) < \gamma < \infty\,, \qquad (\star)$$

where $D_{\max}$ is the max-relative entropy defined in Equation (1.63).

From these two conditions, we can state a meta strong converse bound result as follows:

Theorem 13.1.6. *Let $\{\rho_n\}_{n\in\mathbb{N}^*}$ be a sequence of full-rank states, where for each $n \in \mathbb{N}^*$, $\rho_n \in \mathcal{D}(\mathcal{H}^{\otimes n})$ is full-rank. Assume, moreover, that there exists a family of primitive QMS $(\mathcal{P}_t^n = e^{t\mathcal{L}_n})_{t\geq 0}$ such that, for each $n \in \mathbb{N}^*$, $(\mathcal{P}_t^n)_{t\geq 0}$ satisfies ρ_n-DBC and $\alpha_1(\mathcal{L}_n) \geq \alpha_1 > 0$. Then, for any other family $\{\sigma_n\}_{n\in\mathbb{N}^*}$ of full-rank states $\sigma_n \in \mathcal{D}(\mathcal{H}^{\otimes n})$ such that $(\star)$ holds, for any $n \in \mathbb{N}^*$ and any test $0 \leq T_n \leq \mathbb{1}_n$:*

$$-\frac{1}{n} \ln \operatorname{Tr}(\sigma_n T_n) \leq \frac{1}{n} D(\rho_n \| \sigma_n) + \frac{2}{\sqrt{n}} \sqrt{\frac{\gamma}{4\alpha_1} \ln \frac{1}{\operatorname{Tr}(\rho_n T_n)}} - \frac{1}{4\alpha_1 n} \ln \operatorname{Tr}(\rho_n T_n)\,. \qquad (13.12)$$

The following corollary constitutes a first simple non i.i.d. instance of Theorem 13.1.6:

Corollary 13.1.7. *For each $n \in \mathbb{N}^*$, let $K_n \in \mathbb{N}^*$, $p_n = \{p_n(i)\}$ a probability mass function on $\{1, ..., K_n\}$, and $\tilde{\rho}_i, \tilde{\sigma}_{ij} \in \mathcal{D}_+(\mathcal{H})$ such that $\gamma := \sup_{ij} D_{\max}(\tilde{\rho}_i \| \tilde{\sigma}_{ij}) < \infty$. Next, define*

$$\rho_n = \bigotimes_{i=1}^n \tilde{\rho}_i \qquad and \qquad \sigma_n = \sum_{i=1}^{K_n} p_n(i) \bigotimes_{j=1}^n \tilde{\sigma}_{ij}$$

Then, for any test $0 \leq T_n \leq \mathbb{1}_n$,

$$-\frac{1}{n} \ln \operatorname{Tr}(\sigma_n T_n) \leq \frac{1}{n} D(\rho_n \| \sigma_n) + \frac{2}{\sqrt{n}} \sqrt{\gamma \ln \frac{1}{\operatorname{Tr}(\rho_n T_n)}} - \frac{1}{n} \ln \operatorname{Tr}(\rho_n T_n)\,. \qquad (13.13)$$

Proof. For each $n \in \mathbb{N}^*$, use the tensor product $\mathcal{P}^n := (\otimes_{i=1}^n \mathcal{P}_t^{\mathrm{depol},\tilde{\rho}_i})_{t\geq 0}$ of generalized depolarizing semigroups $(\mathcal{P}_t^{\mathrm{depol},\tilde{\rho}_i})_{t\geq 0}$, each converging to the state $\tilde{\rho}_i$. We know from Theorem 10.1.1 that $\mathcal{P}^n$ satisfies MLSI(1/4). It remains to prove that $\frac{1}{nt} D_{\max}(\mathcal{P}_{t*}^n(\sigma_n) \| \sigma_n) \leq \gamma$. This is done as follows:

$$\begin{aligned}
\frac{1}{nt} D_{\max}(\mathcal{P}_{t*}^n(\sigma_n) \| \sigma_n) &\leq \frac{1}{nt} \max_{i\in\{1,...,K_n\}} D_{\max}\left(\bigotimes_{j=1}^n \mathcal{P}_{t*}^{\mathrm{depol},\tilde{\rho}_j}(\tilde{\sigma}_{ij}) \Big\| \bigotimes_{j=1}^n \tilde{\sigma}_{ij}\right) \\
&\leq \frac{1}{t} \max_{i\in\{1,...,K_n\}} \max_{j\in\{1,...,n\}} D_{\max}(e^{-t}\tilde{\sigma}_{ij} + (1-e^{-t})\tilde{\rho}_j \| \tilde{\sigma}_{ij}) \\
&\leq \frac{1-e^{-t}}{t} \max_{i\in\{1,...,K_n\}} \max_{j\in\{1,...,n\}} D_{\max}(\tilde{\rho}_j \| \tilde{\sigma}_{ij}) \\
&\leq \sup_{ij} D_{\max}(\tilde{\rho}_j \| \tilde{\sigma}_{ij})\,.
\end{aligned}$$

where the first and third inequalities follow by quasi-convexity of $D_{\max}$, and where the second follows by additivity of $D_{\max}$.

□

13.2. Quantum parameter estimation

Here, we apply the concentration inequality (12.86) to the problem of parameter estimation of quantum states. Assume that n independent physical systems are prepared in the same state ρ_θ, where θ is an unknown parameter belonging to a set Θ. Here, we assume that $\Theta = \mathbb{R}$. In order to estimate θ, an estimator is described by a sequence of positive operator valued measurements (POVMs in short) $\mathbf{M} \coloneqq \{M^{(N)}\}_{N \in \mathbb{N}}$, where, for each N, $M^{(N)} : \mathcal{B}(\mathbb{R}) \mapsto \mathcal{P}(\mathcal{H}^{\otimes N})$ is a POVM on the Hilbert space $\mathcal{H}^{\otimes N}$ associated to the N systems, where $\mathcal{B}(\mathbb{R})$ stands for the Borel algebra associated to $\mathbb{R}$. The merit of such a sequence can be quantified in terms of the following error exponent (see [Hayashi, 2002, Nagaoka, 2005, Masahito, 2005]):

$$\beta(\mathbf{M}, \theta, \varepsilon, N) \coloneqq -\frac{1}{N\varepsilon^2} \log \mathbb{P}_{M^{(N)}}(\hat{\theta}_N \in [\theta - \varepsilon, \theta + \varepsilon]^c),$$

where

$$\mathbb{P}_{M^{(N)}}(\hat{\theta}_N \in [\theta - \varepsilon, \theta + \varepsilon]^c) \coloneqq \mathrm{Tr}(M^{(N)}([\theta - \varepsilon, \theta + \varepsilon]^c)\rho_\theta^{\otimes N})$$

is the probability that the estimated value $\hat{\theta}_N$ is at least ε away from the true parameter θ. In the asymptotic setting $N \to \infty$, it was shown in Lemma 14 of [Hayashi, 2002] that, under some technical assumptions, any POVM $\mathbf{M}$ satisfies

$$\limsup_{\varepsilon \to 0} \limsup_{N \to \infty} \beta(\mathbf{M}, \theta, \varepsilon, N) \leq \frac{\mathrm{I}_{\mathrm{SLD}}(\theta)}{2}, \tag{13.14}$$

where $\mathrm{I}_{\mathrm{SLD}}(\theta) \coloneqq \mathrm{Tr}(\rho_\theta(L_\theta^{\mathrm{SLD}})^2)$ is the quantum symmetric logarithmic derivative (SLD for short) Fisher information defined in Section 1.3. For sake of simplicity, we assume that for any $\theta \in \mathbb{R}$, ρ_θ is full-rank. Moreover, the bound in Equation (13.14) was proved to be saturated for a sequence of projection-valued measurements $\mathbf{M}_\theta$ associated to the self-adjoint operator

$$X_\theta^{(N)} \coloneqq \frac{1}{N} \sum_{k=1}^{N} \mathbb{1}^{\otimes(k-1)} \otimes \left(\frac{L_\theta^{\mathrm{SLD}}}{\mathrm{I}_{\mathrm{SLD}}(\theta)} + \theta \mathbb{1} \right) \otimes \mathbb{1}^{\otimes(N-k)}, \tag{13.15}$$

where the estimated value $\hat{\theta}_N$ is determined to be the outcome of the measurement $M_\theta^{(N)}$. This means that, for N large enough and $\varepsilon > 0$ small enough, the error probability satisfies the following lower bound

$$\mathbb{P}_{M_\theta^{(N)}}(\hat{\theta}_N \in [\theta - \varepsilon, \theta + \varepsilon]^c) \gtrsim \mathrm{e}^{-\varepsilon^2 N \mathrm{I}_{\mathrm{SLD}}(\theta)/2}.$$

The family $\mathbf{M}_\theta$ corresponds to a sequence of unbiased estimators in the sense that for all $N \in \mathbb{N}$ and any $\theta \in \mathbb{R}$:

$$\mathrm{Tr}\,\rho_\theta^{\otimes N} X_\theta^{(N)} = \theta. \tag{13.16}$$

The following result provides a finite N upper bound on the error probability.

Proposition 13.2.1. *Let $\mathcal{H}$ be a finite dimensional Hilbert space. For any $\theta \in \mathbb{R}$, let $\rho_\theta \in \mathcal{D}(\mathcal{H})$ be full-rank. Then, for any sequence of unbiased projective estimators $\mathbf{M} \coloneqq \{M^{(N)}\}_{N \in \mathbb{N}}$ associated to the*

self-adjoint operators defined by

$$X_N := \frac{1}{N} \sum_{k=1}^{N} \mathbb{1}^{\otimes(k-1)} \otimes X \otimes \mathbb{1}^{\otimes(N-k)}, \tag{13.17}$$

and any $N \in \mathbb{N}$, the probability that the associated estimated value $\hat{\theta}_N$ lies at least ε away from the true parameter θ is given by

$$\beta(\mathbf{M}, \theta, \varepsilon, N) \geq \frac{1}{2 \, \|(\Delta_{\rho_\theta}^{-1/2} X)\|_{\mathrm{Lip}}^2}, \tag{13.18}$$

where the Lipschitz constant in (13.18) is associated with the generalized depolarizing semigroup (cf. Equation (12.73)).

Proof. This is a straightforward application of (12.86) and (12.85). □

Chapter 14.

Markovianity and Entanglement loss

Distributing entangled quantum states using noisy quantum channels is one of the fundamental challenges in quantum information theory. Entanglement breaking channels, i.e. quantum channels that only output separable states when acting on one half of a bipartite quantum state, are useless for such non-classical communication protocols. In this chapter, we prove several upper and lower bounds on how often we have to apply a quantum channel to one half of a bipartite quantum system to ensure that its output is not entangled regardless of the input, i.e., how long does it take for it to become entanglement breaking. In the context of quantum repeaters, the successive application of quantum channels arises naturally [Bäuml et al., 2015, Christandl and Ferrara, 2017] and such bounds, thus, limit their applicability to implement non-classical communication protocols.

We will focus on estimating these times for primitive quantum channels and show that they always become entanglement breaking after a finite number of iterations. It then follows that the set of quantum channels that eventually become entanglement breaking is dense in the set of quantum channels. This restriction is justified by structural results we obtain which show that the situation is much more subtle if we drop these assumptions. In the case of quantum dynamical semigroups in continuous time, we obtain a complete characterization and show that only primitive semigroups become entanglement breaking in finite time. We obtain concrete bounds on the necessary number of iterations based on spectral data of the underlying quantum channel, in particular the spectral gap of the channel. The bounds scale logarithmically in the underlying dimension and inverse linear in the logarithm of the spectral gap. We also consider quantum channels in continuous time and observe a similar behaviour. These techniques allow us to easily generalize some of the results of [Rahaman et al., 2018].

The problem of classifying quantum channels that become entanglement breaking after repeated applications has recently received considerable attention. The authors of [Lami and Giovannetti, 2016, Kennedy et al., 2017, Rahaman et al., 2018] took a more qualitative and asymptotic point of view, characterizing classes of channels that never become entanglement breaking and showing that certain classes of channels eventually become entanglement breaking. In [Christandl et al., 2018], the authors take a more quantitative point of view which is closer to ours, obtaining upper bounds on the number of iterations in terms of the Schmidt number of a channel.

Moreover, we use similar techniques to consider the related problem of when a pair of quantum channels becomes entanglement annihilating [Moravčíková and Ziman, 2010]. That is, how often we have to apply each channel to each half of the system until the output is separable independently of the input.

Most of our results to derive upper bounds are based on the observation that full rank product quantum states lie in the relative interior of the set of quantum states, as already proved in [Lami

and Giovannetti, 2016]. We then use techniques similar to those of [Gurvits and Barnum, 2002] to obtain estimates on the radius of the separable ball around such states in different metrics. Combining these with tools to estimate the convergence and mixing of quantum channels, we manage to obtain estimates on how long it takes for all outputs to be in the separable ball. Furthermore, we obtain a complete characterization of semigroups that become entanglement breaking in finite time in the continuous time setting.

To derive lower bounds, we exploit the fact that quantum channels that only output separable states remain positive maps when composed with the partial transposition. Thus, if we can show that the output of a state under the channel has a negative partial transposition, the channel still preserves some entanglement. Applying this reasoning with the maximally entangled state as an input, we are then able to obtain criteria based on the spectrum of the quantum channel to certify it still preserves some entanglement. Unlike our upper bounds, we do not make any assumptions on the structure of the quantum channels to prove these lower bounds, although we derive specialized versions for quantum channels of particular interest, such as quantum Markov semigroups in continuous time.

14.1. Preliminary definitions

Given a bipartite system $\mathcal{H} \otimes \mathcal{H}$, we denote by Υ the *maximally entangled state* on $\mathcal{H} \otimes \mathcal{H}$: given any orthonormal basis $\{e_i\}$ of $\mathcal{H}$,

$$\Upsilon := \frac{1}{\sqrt{d_\mathcal{H}}} \sum_{i=1}^{d_\mathcal{H}} e_i \otimes e_i \,.$$

Next, we recall that a state $\rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H})$ is said to be *separable* if it can be written in the following form

$$\rho = \sum_{i=1}^{K} p_i \, \rho_i \otimes \tau_i \,,$$

where $\{\rho_i\}$ and $\{\tau_i\}$ are families of single-partite states in $\mathcal{D}(\mathcal{H})$, and $\{p_i\}$ forms a distribution. More generally, a positive semidefinite operator is called separable if it can be written as the sum of tensor products of positive semidefinite operators. A state that is not separable is then said to be *entangled*. We denote by $\mathrm{SEP}(\mathcal{H} : \mathcal{K})$ the convex subset of separable states in $\mathcal{D}(\mathcal{H} \otimes \mathcal{K})$. When the Hilbert spaces are labeled as $\mathcal{H} = \mathcal{H}_A$ and $\mathcal{K} = \mathcal{H}_B$, we also use the shortcut $\mathrm{SEP}(A : B)$.

A quantum channel $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is called *entanglement breaking* if $\Phi \otimes \mathrm{id}_\mathcal{H}(\rho)$ is a separable state for all input states $\rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H})$. This is equivalent to the Choi matrix of the channel being separable [Horodecki et al., 2003], where the (unnormalized) Choi matrix of Φ is defined as

$$J(\Phi) := d_\mathcal{H} \, (\Phi \otimes \mathrm{id}_\mathcal{H})(|\Upsilon\rangle\langle\Upsilon|) \,. \tag{14.1}$$

The class of entanglement breaking channels on $\mathcal{H}$ is denoted by $\mathrm{EB}(\mathcal{H})$.

The map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is of *positive partial transpose* (PPT) if $(\mathcal{T} \circ \Phi) \otimes \mathrm{id}_\mathcal{H}$ is a positive operator, where $\mathcal{T}$ is the partial transpose w.r.t. to some basis. The class of PPT channels on $\mathcal{H}$ is called $\mathrm{PPT}(\mathcal{H})$. It is well-known that $\mathrm{EB}(\mathcal{H}) \subset \mathrm{PPT}(\mathcal{H})$.

More generally, we also consider quantum Markovian evolutions on a bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. We call a bipartite quantum channel $\mathcal{T} : \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ *entanglement annihilating* if its output is separable (on any density matrix input). Similarly, a quantum channel on a single-partite Hilbert space $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is called 2-*locally entanglement annihilating* if $\Phi \otimes \Phi$ is entanglement annihilating. The class of bipartite, entanglement annihilating channels on $\mathcal{H}_A \otimes \mathcal{H}_B$

is denoted by $\mathrm{EA}(\mathcal{H}_A, \mathcal{H}_B)$, whereas the class of 2-locally entanglement annihilating channels on $\mathcal{H}$ is denoted by $\mathrm{LEA}_2(\mathcal{H})$. Note that any quantum channel that is entanglement breaking is also 2-locally entanglement annihilating: $\mathrm{EB}(\mathcal{H}) \subset \mathrm{LEA}_2(\mathcal{H})$.

In this chapter, we exclusively study the entanglement properties of quantum Markovian evolutions in discrete and continuous time. Given a quantum channel $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, the sequence $\{\Phi^n\}_{n\in\mathbb{N}}$ is called a (homogeneous) *quantum Markov chain* (QMC), by analogy with the classical case. The channel Φ can then be interpreted as the transition map associated to the chain $\{\Phi^n\}_{n\in\mathbb{N}}$. A quantum Markov chain $\{\Phi^n\}_{n\in\mathbb{N}}$, resp. a quantum Markov semigroup $(\mathcal{P}_t)_{t\geq 0}$, is said to be *eventually entanglement breaking* (EEB) if there exists $n_0 \in \mathbb{N}$, resp. $t_0 \geq 0$, such that for any $n \geq n_0$, resp. $t \geq t_0$, Φ^n, resp. $\mathcal{P}_{*t}$, is entanglement breaking. The class of eventually entanglement breaking Markovian evolutions in discrete or continuous time is denoted by $\mathrm{EEB}(\mathcal{H})$. On the other hand, Markovian evolutions in discrete or continuous time which are not entanglement breaking at any finite time are called *entanglement saving*, using language introduced by [Lami and Giovannetti, 2016]; the class of entanglement saving Markovian evolutions in discrete or continuous time is denoted by $\mathrm{ES}(\mathcal{H})$. Thus, the set of all Markovian evolutions (in either discrete or continuous time) decomposes into two disjoint classes:

$$\mathrm{EEB}(\mathcal{H}) \sqcup \mathrm{ES}(\mathcal{H}). \tag{14.2}$$

[Lami and Giovannetti, 2016] also introduce the notion of *asymptotically entanglement saving* evolutions in the discrete-time case. They showed that every QMC has at least one limit point, and either all of the limit points of a QMC $\{\Phi^n\}_{n=1}^{\infty}$ are entanglement breaking, or none of them are. They term the latter case as asymptotically entanglement saving, and we denote the set of asymptotically entanglement saving evolutions on $\mathcal{H}$ by $\mathrm{AES}(\mathcal{H})$. In analogy, we call the former case by *asymptotically entanglement breaking*, denoted $\mathrm{AES}(\mathcal{H})$. Thus, the set of QMC on $\mathcal{H}$ decomposes into the disjoint classes:

$$\mathrm{AES}(\mathcal{H}) \sqcup \mathrm{AEB}(\mathcal{H}). \tag{14.3}$$

It is interesting to compare (14.2) and (14.3). A QMC $\{\Phi^n\}_{n=1}^{\infty}$ is AES if all the limit points of the sequence $\{J(\Phi^n)\}_{n=1}^{\infty}$ are all entangled. Since $J(\Phi^{n+1}) = \Phi \otimes \mathrm{id}(J(\Phi^n))$, if $J(\Phi^{n+1})$ is entangled, $J(\Phi^n)$ must be as well. In particular, if $\{\Phi^n\}_{n=1}^{\infty}$ is asymptotically entanglement saving, then $J(\Phi^n)$ is entangled for every n, and the QMC is entanglement saving. So we see $\mathrm{AES}(\mathcal{H}) \subset \mathrm{ES}(\mathcal{H})$. However, a priori, an entanglement saving channel could be asymptotically entanglement breaking: at any finite n, $J(\Phi^n)$ could be entangled, but in the limit, $J(\Phi^n)$ could be in the set of separable states (though necessarily on the boundary). We therefore define $\mathrm{EB}_{\infty}(\mathcal{H}) = \mathrm{AEB}(\mathcal{H}) \cap \mathrm{ES}(\mathcal{H})$, the set of QMC which are asymptotically entanglement breaking, but not entanglement breaking for any finite n. With this notion, we may relate (14.2) and (14.3). We have the disjoint decomposition of the set of all QMC,

$$\mathrm{QMC}(\mathcal{H}) = \mathrm{EEB}(\mathcal{H}) \sqcup \overbrace{\mathrm{EB}_{\infty}(\mathcal{H}) \sqcup \mathrm{AES}(\mathcal{H})}^{\mathrm{AEB}(\mathcal{H})}. \tag{14.4}$$

where $\mathrm{EB}_{\infty}(\mathcal{H}) \sqcup \mathrm{AES}(\mathcal{H})$ is underbraced as $\mathrm{ES}(\mathcal{H})$.

Theorem 14.1.1 (Eventually entanglement breaking channels are dense). *For any finite dimensional Hilbert space $\mathcal{H}$, the set of transition maps of eventually entanglement breaking quantum Markov chains is dense in the set of quantum channels.*

Proof. First note that the set of quantum channels that have a full rank stationary state is dense in the set of quantum channels. To see this, let $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be a quantum channel and for a state

$\sigma \in \mathcal{D}(\mathcal{H})$ define $\Phi_\sigma : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be the quantum channel that acts as

$$\Phi_\sigma(X) = \text{Tr}(X)\sigma\,.$$

Now note that for $\epsilon \in (0,1]$, $\Phi' = (1-\epsilon)\Phi + \epsilon\,\Phi_{\frac{1}{d_\mathcal{H}}}$ maps quantum states to strictly positive operators and, thus, has a stationary state $\tau \in \mathcal{D}_+(\mathcal{H})$ of full rank [Wolf, 2012, Theorem 6.2]. We may then consider the quantum channel $\Phi'' = (1-\epsilon)\Phi' + \epsilon\,\Phi_\tau$. As Φ' has τ as a stationary state, it follows that $\Phi'\Phi_\tau = \Phi_\tau\Phi'$ and, thus,

$$\left(\Phi''\right)^n = (1-\epsilon)^n \left(\Phi'\right)^n + \left(1 - (1-\epsilon)^n\right)\Phi_\tau\,.$$

Clearly, $\lim_{n\to\infty} \left(\Phi''\right)^n = \Phi_\tau$. We have

$$\text{id} \otimes \Phi_\tau(|\Upsilon\rangle\langle\Upsilon|) = \frac{\mathbb{1}}{d_\mathcal{H}} \otimes \tau\,.$$

As was already observed in [Lami and Giovannetti, 2016, Proposition 3], separable states of the form $\sigma_1 \otimes \sigma_2$ with $\sigma_1, \sigma_2 \in \mathcal{D}(\mathcal{H})$ full rank are in the interior of the set of separable states. Thus, the Choi matrix of $\left(\Phi''\right)^n$ converges to a separable state in the interior of the set of separable states and will be separable for some finite n_0. $\qquad\square$

14.2. Asymptotic entanglement properties of Markovian evolutions

As mentioned before, our goal in this chapter is to estimate the time after which a quantum system undergoing a quantum Markovian evolution has lost all its entanglement. In order to better characterize evolutions for which asking this question makes sense, we first need to leave aside evolutions for which the phenomenon does not occur, that is, evolutions that either destroy entanglement after an infinite amount of time (entanglement saving evolutions), or even those of never-vanishing entanglement.

A big part of this question was already answered in the case of a quantum Markov chain by [Lami and Giovannetti, 2016]. In this paper, the authors showed that, given a quantum channel $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ with $\dim(\ker \Phi) < 2(d_\mathcal{H} - 1)$, Φ is entanglement saving if and only if either it has a positive semidefinite fixed point, or the number of peripheral eigenvalues is strictly greater than 1, which itself is equivalent to the existence of $1 \le n \le d_\mathcal{H}$ such that Φ^n has a positive semidefinite fixed point (see Theorem 21). In the same paper, the authors showed that if Φ has strictly more than $d_\mathcal{H}$ peripheral eigenvalues, then $\{\Phi^n\}_{n\in\mathbb{N}}$ is asymptotically entanglement saving. This last result is a simple consequence of the following technical lemma, whose proof can also be found in [Lami and Giovannetti, 2016].

Lemma 14.2.1. *A quantum channel $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ s.t.[1]*

$$\|\Phi\|_1 > d_\mathcal{H}$$

is not entanglement breaking.

This result has another simple consequence.

---

[1]Here, $\|\Phi\|_1$ simply stands for the sum of the singular values of Φ.

Proposition 14.2.2. *Let Φ be a quantum channel on $\mathcal{H}$ with $d_{\mathcal{H}}$ peripheral eigenvalues and at least one non-peripheral eigenvalue nonzero. Then $\{\Phi^n\}_{n\in\mathbb{N}} \in \mathrm{ES}(\mathcal{H})$, i.e. $\Phi^n \notin \mathrm{EB}(\mathcal{H})$ for any $n \in \mathbb{N}$.*

Proof. For any n, Φ^n has $d_{\mathcal{H}}$ peripheral eigenvalues and at least one non-peripheral eigenvalue non-zero. Thus, if $\{\lambda_k\}_{k=1}^{d_{\mathcal{H}}^2}$ are the eigenvalues of Φ counted with multiplicity, we have

$$\|\Phi^n\|_1 = \sum_{k=1}^{d_{\mathcal{H}}^2} |\lambda_k^n| = d_{\mathcal{H}} + \sum_{\lambda_k:|\lambda_k|<1} |\lambda_k^n| > d_{\mathcal{H}}$$

and thus, by Lemma 14.2.1, Φ^n is not EB. $\qquad\square$

14.2.1. Irreducibility and primitivity

A positive linear map $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is said to be *irreducible* if, for any orthogonal projection $P \in \mathcal{B}(\mathcal{H})$, $\Phi(P\mathcal{B}(\mathcal{H})P) \subset P\mathcal{B}(\mathcal{H})P$ implies that either $P = 0$ or $P = \mathbb{1}$. In the case of an irreducible, completely positive (in fact, for a Schwarz) map, it is known from Perron-Frobenius theory that the peripheral eigenvalues λ_k in (6.15) are non-degenerate and equal to ϕ^k, where $\phi := \exp(2\mathrm{i}\pi/z)$, for some fixed $z \in \mathbb{N}$.

For a positive trace-preserving map, this property is equivalent to the existence of a unique state $\sigma > 0$ such that for every $\omega \in \mathcal{D}(\mathcal{H})$, we have (see [Wolf, 2012, Corollary 6.3])

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \Phi^n(\omega) = \sigma\,. \tag{14.5}$$

Irreducible maps have many useful algebraic properties. The following proposition establishes some of these, along with a minimal set of quantities needed to construct such a map.

Proposition 14.2.3. *Consider*

1. *A number $z \in \{1,\ldots,d_{\mathcal{H}}\}$,*

2. *An orthogonal resolution of the identity $\{p_n\}_{n=0}^{z-1}$,*

3. *A faithful state ρ^{inv} such that $[\rho^{\mathrm{inv}}, p_n] = 0$ and $\mathrm{Tr}[\rho^{\mathrm{inv}} p_n] = \frac{1}{z}$, for each $n = 0,\ldots,z-1$,*

4. *A linear map Φ_Q such that:*

   a) *$\mathrm{spr}(\Phi_Q) < 1$*

   b) *$J(\Phi_Q) \geq -z\left(\rho^{\mathrm{inv}} \otimes \mathbb{1}\right) L_1$, where for $k = 0,\ldots,z-1$, we define $L_k := \sum_{n=0}^{z-1} p_{n-k} \otimes p_n$ where the subscripts are taken modulo z.*

   c) *We have*
   $$\Phi_Q \circ P_j = P_j \circ \Phi_Q = 0\,, \qquad \forall\, j = 0,\ldots,z-1\,,$$
   where $P_n(\cdot) = \mathrm{Tr}[u^{-n}\,\cdot\,]u^n\rho^{\mathrm{inv}}$ for $u := \sum_{n=0}^{z-1}\theta^n p_n$ and $\theta := \exp(2\mathrm{i}\pi/z)$.

Let

$$\Phi := \sum_{n=0}^{z-1} \theta^n P_n + \Phi_Q\,. \tag{14.6}$$

Then Φ is an irreducible quantum channel. On the other hand, any irreducible quantum channel Φ can be decomposed as (14.6) for some choices of z, $\{p_n\}_{n=0}^{z-1}$, ρ^{inv}, and Φ_Q as in (1)–(4). Moreover, in either case, ρ^{inv} is the unique invariant[2] of Φ; $P_n(\cdot)$ are its peripheral eigenprojections, associated

---

[2] up to a multiplicative constant

to eigenvalues θ^n and eigenvectors $u^n \rho^{\text{inv}}$; and, for any $j, k = 0, \ldots, z-1$, we have the intertwining relations

$$\Phi(p_j X p_k) = p_{j-1} \Phi(X) p_{k-1}, \quad \text{and} \quad \Phi^*(p_j X p_k) = p_{j+1} \Phi^*(X) p_{k+1}, \quad \forall X \in \mathcal{B}(\mathcal{H}) \tag{14.7}$$

where the subscripts are interpreted modulo z. Additionally, for $\Phi_P := \sum_{n=0}^{z-1} \theta^n P_n$,

$$J(\Phi_P^k) = \hat{J}_k := z\left(\rho^{\text{inv}} \otimes \mathbb{1}\right) \cdot L_k = \sum_{m=0}^{z-1} \text{Tr}[p_k] \frac{p_{m-k} \rho^{\text{inv}} p_{m-k}}{\text{Tr}[p_{m-k} \rho^{\text{inv}}]} \otimes \frac{p_k}{\text{Tr}[p_k]}. \tag{14.8}$$

Proof. That any irreducible CPTP map can be decomposed as (14.6) for some choices of z, $\{p_n\}_{n=0}^{z-1}$, ρ^{inv}, and Φ_Q as in (1)–(4) are some of the main results of the Perron-Frobenius theory for irreducible completely positive maps; see [Evans and Høegh-Krohn, 1978], or Section 6.2 of [Wolf, 2012]. See also Appendix A of [Hanson et al., 2018] for a summary of this theory and extensions to deformations of irreducible CPTP maps. Equation (14.7) follows from Theorem 5.4 of [Fagnola and Pellicer, 2009]. Let us show that given z, $\{p_n\}_{n=0}^{z-1}$, ρ^{inv}, and Φ_Q, the decompositon (14.6) gives an irreducible quantum channel. Note, by the definition of P_n, that for any $X \in \mathcal{B}(\mathcal{H})$,

$$P_j \circ P_k(X) = \text{Tr}[u^{-k} X]\, \text{Tr}[u^{k-j} \rho^{\text{inv}}] u^j \rho^{\text{inv}} = \delta_{jk} P_j(X)$$

using $\text{Tr}[\rho^{\text{inv}} p_n] = \frac{1}{z}$ and the formula

$$\sum_{n=0}^{z-1} \theta^{nm} = \begin{cases} z & m = zk \text{ for some } k \in \mathbb{Z} \\ 0 & \text{otherwise}. \end{cases} \tag{14.9}$$

Since $P_0(X) = \text{Tr}[X] \rho^{\text{inv}}$, we have for $j \neq 0$, $0 = P_0 \circ P_j(X) = \text{Tr}[P_j(X)] \rho^{\text{inv}}$, yielding that P_j is trace-annihilating: $\text{Tr}[P_j(X)] = 0$ for all $X \in \mathcal{B}(\mathcal{H})$. In the same way, using assumption (4c), Φ_Q is trace-annihilating. Thus, $\Phi = P_0 + \sum_{n=1}^{z-1} \theta^n P_n + \Phi_Q$ is trace-preserving.

Next, we prove (14.8), which will prove Φ is CP via assumption (4b). For $\Phi_P := \sum_{m=0}^{z-1} \theta^m P_m$, we have $\Phi_P^k = \sum_{m=0}^{z-1} \theta^{km} P_m$. Then, for any $X \in \mathcal{B}(\mathcal{H})$, we have the discrete Fourier-type computation,

$$\Phi_P^k(X) = \sum_{m=0}^{z-1} \theta^{km} \text{Tr}[u^{-m} X] u^m \rho^{\text{inv}} = \sum_{m,n,\ell=0}^{z-1} \theta^{km} \text{Tr}[\theta^{-mn} p_n X] \theta^{\ell m} p_\ell \rho^{\text{inv}}$$

$$= \sum_{m,n,\ell=0}^{z-1} \theta^{m(k-n+\ell)} \text{Tr}[p_n X] p_\ell \rho^{\text{inv}} = \sum_{n,\ell=0}^{z-1} z \delta_{\ell,n-k} \text{Tr}[p_n X] p_\ell \rho^{\text{inv}}$$

$$= z \sum_{n=0}^{z-1} \text{Tr}[p_n X] p_{n-k} \rho^{\text{inv}}$$

using (14.9). Next, let $\{e_i\}_{i=0}^{d_{\mathcal{H}}-1}$ be an orthonormal basis of $\mathcal{H}$ such that the first $\text{rank}(p_0)$ elements are a basis for $p_0 \mathcal{H}$, the next $\text{rank}(p_1)$ elements are a basis for $p_1 \mathcal{H}$, and so on. We have $p_0 = \sum_{i=0}^{\text{rank}(p_0)} |e_i\rangle \langle e_i|$, and so forth. Thus,

$$J(\Phi_P^k) = \sum_{i,j=0}^{d_{\mathcal{H}}-1} \Phi_P^k(|e_i\rangle \langle e_j|) \otimes |e_i\rangle \langle e_j| = z \sum_{i,j} \sum_{n=0}^{z-1} \text{Tr}[p_n |e_i\rangle \langle e_j|] p_{n-k} \rho^{\text{inv}} \otimes |e_i\rangle \langle e_j|$$

$$= z \sum_{n=0}^{z-1} \sum_{i=0}^{d_{\mathcal{H}}-1} \langle e_i, p_n e_i\rangle p_{n-k} \rho^{\text{inv}} \otimes |e_i\rangle \langle e_i| = z \sum_{n=0}^{z-1} \rho^{\text{inv}} p_{n-k} \otimes p_n$$

$$= z\left(\rho^{\text{inv}} \otimes \mathbb{1}\right) L_k.$$

In particular, $J(\Phi_P) = z(\rho^{\mathrm{inv}} \otimes \mathbb{1})L_1$. Thus, by assumption (4b),

$$J(\Phi) = J(\Phi_P) + J(\Phi_Q) \geq J(\Phi_P) - z(\rho^{\mathrm{inv}} \otimes \mathbb{1})L_1 = 0$$

and hence Φ is CP. Since Φ is CPTP, we can use (14.5) to prove Φ is irreducible. We have

$$\frac{1}{M} \sum_{n=0}^{M-1} \Phi^n = P_0 + \frac{1}{M} \sum_{m=1}^{z-1} \frac{1 - \theta^{Mm}}{1 - \theta^m} P_m + \frac{1}{M} \sum_{n=0}^{M-1} \Phi_Q^n$$

using the geometric series $\sum_{n=0}^{M-1} \theta^{mM} = \frac{1 - \theta^{mM}}{1 - \theta^m}$ for $m \neq 0$, which is valid since $\theta^m \neq 1$. Since $P_0[X] = \mathrm{Tr}[X]\rho^{\mathrm{inv}}$, it remains to show that the latter two terms vanish in the limit $M \to \infty$. In fact, since $\sum_{m=1}^{z-1} \frac{1 - \theta^{Mm}}{1 - \theta^m} P_m$ is bounded in norm uniformly in M, the second term vanishes asymptotically. Next, since $\ell := \mathrm{spr}(\Phi_Q) < 1$ by assumption (4a), for $\varepsilon = \frac{1 - \ell}{2} > 0$, Gelfand's theorem gives that there is $n_0 > 0$ such that for all $n \geq n_0$,

$$\|\Phi_Q^n\| \leq (\ell + \varepsilon)^n < 1.$$

We may write

$$\frac{1}{M} \sum_{n=0}^{M-1} \Phi_Q^n = \frac{1}{M} \sum_{n=0}^{n_0} \Phi_Q^n + \frac{1}{M} \sum_{n=n_0+1}^{M-1} \Phi_Q^n.$$

Since $\sum_{n=0}^{n_0} \Phi_Q^n$ is bounded in norm independently of M, the first term vanishes asymptotically; the second term is bounded in norm by the triangle inequality and the geometric series $\sum_{n=0}^{\infty}(\ell+\varepsilon)^n = \frac{1}{1-(\ell+\varepsilon)}$. Thus, the limit

$$\lim_{M \to \infty} \frac{1}{M} \sum_{n=0}^{M-1} \Phi^n = P_0$$

holds in (any) norm. In particular, we have (14.5), so Φ is irreducible.

$\square$

A positive map Φ is called *primitive* if the maps $\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)} := \sum_{k:|\lambda_k|=1} P_k$ and Φ^n, $n \in \mathbb{N}$, are irreducible. This turns out to be equivalent to the existence of a full rank state σ such that, for any $\rho \in \mathcal{D}(\mathcal{H})$, $\Phi^n(\rho) \to \sigma$ as $n \to \infty$. We refer to [Burgarth et al., 2013, Wolf, 2012] for other characterizations of primitive channels and sufficient conditions for primitivity. A quantum Markov chain $\{\Phi^n\}_{n\in\mathbb{N}}$ is called *primitive* if the channel Φ is primitive.

With the notations of Proposition 14.2.3, we can make the following remarks:

(i) Note that $\{L_k\}_{k=0}^{z-1}$ is an orthogonal resolution of $\mathbb{1} \otimes \mathbb{1}$, such that each L_k commutes with $\rho^{\mathrm{inv}} \otimes \frac{\mathbb{1}}{d_{\mathcal{H}}}$ and $\mathrm{Tr}[L_k \rho^{\mathrm{inv}} \otimes \frac{\mathbb{1}}{d_{\mathcal{H}}}] = \frac{1}{z}$.

(ii) The matrix $\hat{J}_k$ is separable, and thus Φ_P^k is entanglement-breaking, for any $k \geq 1$.

(iii) Proposition 14.2.3 shows that the peripheral eigenvectors of irreducible channels Φ commute. By characterization of asymptotically entanglement saving channels given in [Lami and Giovannetti, 2016], this implies $\{\Phi^n\}_{n\in\mathbb{N}} \in \mathrm{AEB}(\mathcal{H})$. Thus, Lemma 14.2.1 shows that Quantum Markov chains with irreducible generators, $d_{\mathcal{H}}$ peripheral eigenavalues (the maximum possible for an irreducible QMC), and at least one non-zero non-peripheral eigenvalue are all in $\mathrm{EB}_{\infty}(\mathcal{H})$. Such chains exist by virtue of next corollary.

Corollary 14.2.4. *There are irreducible quantum channels $\Phi : \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathbb{C}^d)$ s.t. $\{\Phi^n\}_{n\in\mathbb{N}} \in \mathrm{EB}_{\infty}(\mathbb{C}^d)$.*

Proof. We will follow the recipe outlined in proposition 14.2.3 to construct such a quanntum channel. We will also follow the notation of that proposition. We will first fix $\rho_{\mathrm{inv}} = \mathbb{1}/d$, $z = d$ and

$p_i = |i\rangle\langle i|$. This choice of operators clearly satisfies all assumptions made on the p_i in the statement of proposition 14.2.3. It remains to construct the map Φ_Q. To this end, first note that

$$L_1 = \sum_{i=0}^{d-1} |i\rangle\langle i| \otimes |i+1\rangle\langle i+1|$$

and that with this choice of p_i, the P_i are just projections onto diagonal matrices. For some $\epsilon \in (0, 1/d)$, we define Φ_Q as

$$\Phi_Q\left(|i\rangle\langle j|\right) = (1 - \delta_{i,j})\,\epsilon\,|i+1\rangle\langle j+1|\,.$$

The Choi matrix of the map defined above is given by

$$J\left(\Phi_Q\right) = \epsilon \sum_{i,j=0, i \neq j}^{d-1} |i\rangle\langle j| \otimes |i+1\rangle\langle j+1|\,.$$

To see that

$$L_1 + J\left(\Phi_Q\right) \geq 0\,,$$

note that for a fixed i, the $i(i+1)$–th line of $L_1 + J\left(\Phi_Q\right)$ has a 1 on its diagonal and $d-1$ off-diagonal entries ϵ. All lines whose number cannot be written as $i(i+1)$ are identically 0. By our choice of ϵ, we see by Gershgorin's circle theorem that the matrix is indeed positive. Moreover, as the outputs of Φ_Q only have off-diagonal elements and those of P_n only diagonal, it follows that $\Phi_Q \circ P_n = P_n \circ \Phi_Q = 0$. Finally, note that as $\Phi_Q^d = \epsilon^d\,\Phi_{\mathrm{off}}$, where Φ_{off} is the projection onto the off-diagonal elements in the computational basis, it follows that $\mathrm{spr}(\Phi_Q) = \epsilon$. We then have by propostion 14.2.3 that

$$\Phi = \sum_{i=0}^{d-1} \theta^i P_i + \Phi_Q$$

is an irreducible quantum channel with $\|\Phi^k\|_1 > d$ and, thus, not entanglement breaking, but with $\lim_{k\to\infty} \Phi^k$ entanglement breaking. $\square$

14.2.2. Convergence of Markovian evolutions and characterization of $\mathrm{AEB}(\mathcal{H})$

Next theorem is a straightforward extension of Theorem 24 of [Lami and Giovannetti, 2016] to the case of a quantum Markov semigroup.

Proposition 14.2.5. *Let $(\mathcal{P}_t)_{t\geq 0}$ be a faithful quantum Markov semigroup. Then $(\mathcal{P}_t)_{t\geq 0} \in \mathrm{AES}(\mathcal{H})$ if and only if its decoherence-free subalgebra $\mathcal{N}$ is noncommutative, which means that there exists $i \in \mathcal{J}$ such that $d_{\mathcal{H}_i} > 1$ in the decomposition (6.4).*

Proof. We simply need to show that $E_{\mathcal{N}*} \in \mathrm{EB}(\mathcal{H})$ if and only if $d_{\mathcal{H}_i} = 1$ for all i. If $d_{\mathcal{H}_i} = 1$ for all i, the result follows directly from the characterization of entanglement breaking channels of [Horodecki et al., 2003]. If now there exists i such that $d_{\mathcal{H}_i} > 1$, choose as input state $\rho = |\Upsilon\rangle\langle\Upsilon| \otimes \tau_i$, where $\Upsilon = (d_{\mathcal{H}_i})^{-1/2} \sum_{j=1}^{d_{\mathcal{H}_i}} e_j \otimes e_j$ is a maximally entangled state on $\mathcal{H}_i \otimes \mathcal{H}_i$, and the result follows from the fact that $\mathrm{id}_{\mathcal{H}} \otimes E_{\mathcal{N}*}(\rho)$ is entangled. $\square$

14.2.3. Primitive Markovian evolutions and entanglement loss times

We just showed that non-primitive, irreducible maps can lead to Markov chains in $\mathrm{EB}_\infty(\mathcal{H})$. In the more constraint case of primitive quantum Markov chains, and their time-continuous counterparts, the situation becomes more simple. First, we need the following straightforward extension of Theorem 1 of [Gurvits and Barnum, 2002], whose proof is postponed to Proposition 14.3.2:

Lemma 14.2.6. *Let $\sigma, \omega \in \mathcal{D}(\mathcal{H})$ be full-rank. Then $\omega \otimes \sigma + \Delta$ is separable for any Hermitian operator Δ such that $\|\Delta\|_2 \leq \sigma_{\min} \omega_{\min}$, where $\sigma_{\min}$, resp. $\omega_{\min}$, stands for the minimum eigenvalue of σ, resp. ω.*

The first part of the following lemma is taken from [Lami and Giovannetti, 2016]:

Lemma 14.2.7. *We have that*

- *A quantum Markov chain $\{\Phi^n\}_{n \in \mathbb{N}}$ with $\dim \ker \Phi < 2(d_\mathcal{H} - 1)$ is eventually entanglement breaking if and only if it is primitive.*

- *A quantum Markov semigroup $(\mathcal{P}_t)_{t \geq 0}$ is eventually entanglement breaking if and only if it is primitive.*

Proof. The first part is proved in [Lami and Giovannetti, 2016]. We focus on the second part: first assume that $(\mathcal{P}_t)_{t \geq 0}$ is primitive. Then there exists a full-rank state $\sigma \in \mathcal{D}(\mathcal{H})$ such that, for any $\rho \in \mathcal{D}(\mathcal{H})$, $\mathcal{P}_{t*}(\rho) \to \sigma$ as $t \to \infty$. Therefore,

$$(\mathcal{P}_{t*} \otimes \mathrm{id})(|\Upsilon\rangle\langle\Upsilon|) \to \sigma \otimes d_\mathcal{H}^{-1} \mathbb{1},$$

since $\mathcal{P}_{t*}$ is trace-preserving for each t. The result follows by Lemma 14.2.6, which implies the existence of $t_0 > 0$ such that for any $t \geq t_0$, $(\mathcal{P}_{t*} \otimes \mathrm{id})(|\Upsilon\rangle\langle\Upsilon|)$ is in a nonempty ball around $\sigma \otimes d_\mathcal{H}^{-1} \mathbb{1}$ consisting of separable states.

Conversely, assume that $(\mathcal{P}_t)_{t \geq 0} \in \mathrm{EEB}(\mathcal{H})$. Therefore there exists $n \in \mathbb{N}$ such that $\Phi_n = \mathcal{P}_{1*}^n$ is entanglement breaking. Since for any $t \geq 0$ the map $\mathcal{P}_{t*}$ is invertible, the first part of the lemma implies that $\mathcal{P}_{1*}$ is primitive. We conclude by a use of Proposition 7.5 of [Wolf, 2012]. $\qquad\square$

Lemma 14.2.7 justifies the introduction of the following characteristic times, at least in the case of a primitive quantum Markovian evolution:

Let $\{\Phi^n\}_{n \in \mathbb{N}}$ be a quantum Markov chain on $\mathcal{B}(\mathcal{H})$, with invariant state $\sigma \in \mathcal{D}_+(\mathcal{H})$. The *entanglement breaking time* $t_{\mathrm{EB}}(\Phi)$ of $\{\Phi^n\}_{n \in \mathbb{N}}$ is defined as follows:

$$t_{\mathrm{EB}}(\Phi) \overset{\mathrm{def}}{=} \min\{n_0 \in \mathbb{N}|\ \forall n \geq n_0,\ \Phi^n \in \mathrm{EB}(\mathcal{H})\}\ .$$

Similarly, a quantum Markov chain $(\Gamma^n)_{n \in \mathbb{N}}$ over a bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, we define the *entanglement annihilation time* $t_{\mathrm{EA}}(\Gamma)$ as follows

$$t_{\mathrm{EA}}(\Gamma) \overset{\mathrm{def}}{=} \min\{n_0 \in \mathbb{N}|\ \forall n \geq n_0,\ \Gamma^n \in \mathrm{EA}(\mathcal{H}_A, \mathcal{H}_B)\}\ .$$

In the case when $\Gamma = \Phi \otimes \Phi$, $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, this time is called the 2-*local entanglement annihilation time*, and is denoted by

$$t_{\mathrm{LEA}_2}(\Phi) \overset{\mathrm{def}}{=} \min\{n_0 \in \mathbb{N}|\ \forall n \geq n_0,\ \Phi^n \in \mathrm{LEA}_2(\mathcal{H})\}\ .$$

Entanglement breaking, entanglement annihilation, and 2-local entanglement annihilation times of quantum Markov semigroups are defined identically. In the next two sections, we provide bounds on

t_{EB}, t_{EA} and t_{LEA_2}: the upper bounds found in Section 14.3 use strong decoherence of Markovian evolutions together with estimates on the radius of open balls around any full-rank product state. On the other hand, lower bounds found in Section 14.4 mainly use the inclusion $\mathrm{EB}(\mathcal{H}) \subset \mathrm{PPT}(\mathcal{H})$.

14.3. Upper bounds on entanglement loss via strong decoherence

As mentioned just above, the strategy for deriving upper bounds on the various entanglement loss times defined in Section 14.2.3 is as follows: in Section 14.3.1, we get quantitative bounds on the radius of balls surrounding any full-rank separable state on a bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ for various normed topologies. Moreover, we studied in Part IV techniques leading to the derivation of bounds on the time it takes for any state evolving according to a Markov chain/semigroup to come ε-close to equilibrium. Upper bounds on entanglement loss times follow by simply choosing ε as the radius of the separable ball around the adequate state found in Section 14.3.1. This is done in Section 14.3.2.

14.3.1. Separable balls around separable states

In this section, we are interested in finding separable balls around separable states, as we could then also treat bipartite semigroups that converge to a separable state in a similar fashion and determine when they become entanglement breaking. Note, however, that there are faithful separable states that lie on the boundary of the set of separable states and the characterization of that boundary is still a subject of active research [Chen and Đoković, 2015, Kye, 2018]. One simple example that illustrates these difficulties would be the state

$$\tau = \frac{1}{d_{\mathcal{H}} + 1}|\Upsilon\rangle\langle\Upsilon| + \left(1 - \frac{1}{d_{\mathcal{H}} + 1}\right)\frac{\mathbb{1}}{d_{\mathcal{H}}} \tag{14.10}$$

on a bipartite Hilbert space $\mathcal{H}$, which lies on the boundary of separable states. Considering a bipartite generalized depolarizing semigroup converging to this state, we see that it is possible that the channel is asymptotically entanglement breaking (in the sense of having only separable outputs), but might have entangled outputs for all finite t. To see this, take the maximally entangled state as an input for the depolarizing semigroup $(\mathcal{P}_t^{\mathrm{depol},\tau})_{t\geq 0}$ described above. Then

$$\mathcal{P}_{t*}^{\mathrm{depol},\tau}(|\Upsilon\rangle\langle\Upsilon|) = e^{-t}|\Upsilon\rangle\langle\Upsilon| + (1 - e^{-t})\tau,$$

which is entangled for all $t \geq 0$, but not in the limit. This suggests that the situation is much more subtle in this setting. One way to bypass this difficulty is to introduce the following measure of "robustness of separability", inspired by the robustness of entanglement introduced in [Vidal and Tarrach, 1999].

Definition 14.3.1 (Robustness of separability)**.** Let $\rho_{AB} > 0$ be separable on the bipartite Hilbert space $\mathcal{H} \equiv \mathcal{H}_{AB}$. We define its robustness of separability with respect to the maximally mixed state, $R(\rho_{AB})$, as

$$R(\rho_{AB}) = \sup\left\{\lambda \in [0,1] : \exists\, \rho'_{AB} \text{ separable such that } \rho_{AB} = \lambda\,\frac{\mathbb{1}_A \otimes \mathbb{1}_B}{d_{\mathcal{H}}} + (1 - \lambda)\,\rho'_{AB}\right\}.$$

Proposition 14.3.2 (Properties of the robustness of separability)**.** *Let $\rho_{AB} \in \mathrm{SEP}(A : B)$. Then we have the following properties.*

1. We have the reformulation

$$R(\rho_{AB}) = \sup\left\{\lambda \in [0,1] : \frac{1}{1-\lambda}\left(\rho_{AB} - \lambda\frac{\mathbb{1}_A \otimes \mathbb{1}_B}{d_A d_B}\right) \in \mathrm{SEP}(A:B)\right\}.$$

2. We have the bounds

$$0 \le R(\rho_{AB}) \le d_A d_B(\rho_{AB})_{min}, \tag{14.11}$$

where $(\rho_{AB})_{min}$ denotes the minimal eigenvalue of ρ_{AB}, and equality holds in the second inequality for product states:

$$R(\rho_A \otimes \rho_B) = d_A\, d_B\, (\rho_A)_{min}\, (\rho_B)_{min}. \tag{14.12}$$

3. $\rho_{AB} \in \mathrm{int}\,\mathrm{SEP}(A:B)$ if and only if $R(\rho_{AB}) > 0$. Moreover, any state σ_{AB} such that

$$\|\rho_{AB} - \sigma_{AB}\|_2 \le \frac{R(\rho_{AB})}{d_{\mathcal{H}}} \tag{14.13}$$

is separable.

4. R is quasi-concave on $\mathrm{SEP}(A:B)$: if $\rho_{AB} = \sum_j p_j \sigma_{AB}^{(j)}$ where $p_j > 0$, $\sum_j p_j = 1$, then

$$R(\rho_{AB}) \ge \min_j R(\sigma_{AB}^{(j)}).$$

Proof. The first point follows by rearranging the definition of R. Next, $R(\rho_{AB}) \ge 0$ is immediate from the definition; the second inequality in the second point follows from the fact that $\rho_{AB} - \lambda\frac{\mathbb{1}_A \otimes \mathbb{1}_B}{d_A d_B} \not\ge 0$ for $\lambda > d_A d_B(\rho_{AB})_{\min}$. For product states, we may explicilty evaluate $R(\rho_A \otimes \rho_B)$ using the formulation in the first point and by expanding ρ_A in its eigenbasis, $\mathbb{1}_A$ in the same basis, ρ_B in its eigenbasis, and $\mathbb{1}_B$ in the same basis.

For the third point, we first note that if $R(\rho_{AB}) = 0$, then ρ_{AB} is not in the (relative) interior of $\mathrm{SEP}(A:B)$: If $R(\rho_{AB}) = 0$, then for any $\lambda \in (0,1)$,

$$\frac{1}{1-\lambda}\left(\rho_{AB} - \lambda\frac{\mathbb{1}_A \otimes \mathbb{1}_B}{d_A d_B}\right) \notin \mathrm{SEP}(A:B).$$

This quantity is in the affine hull of SEP, and can be made arbitrarily close to ρ_{AB} by taking λ small, which proves ρ_{AB} is not in the relative interior of SEP. The other implication follows from the bound (14.13), which is proven as follows. By definition, we may write

$$\rho_{AB} = R(\rho_{AB})\frac{\mathbb{1}_A \otimes \mathbb{1}_B}{d_A d_B} + (1 - R(\rho_{AB}))\rho'_{AB} \tag{14.14}$$

with ρ'_{AB} separable. Now consider another state σ_{AB} st. $\|\rho_{AB} - \sigma_{AB}\|_2 \le R(\rho_{AB})d_{\mathcal{H}}^{-1}$. Then

$$\sigma_{AB} = \rho_{AB} + (\sigma_{AB} - \rho_{AB}) = R(\rho_{AB})\left(\frac{\mathbb{1}_A \otimes \mathbb{1}_B}{d_A d_B} + \frac{1}{R(\rho_{AB})}\left(\rho_{AB} - \sigma_{AB}\right)\right) + (1 - R(\rho_{AB}))\rho'_{AB}.$$

Since $\|\rho_{AB} - \sigma_{AB}\|_2 \le \frac{R(\rho_{AB})}{d_{\mathcal{H}}}$,

$$\left(\frac{\mathbb{1}_A \otimes \mathbb{1}_B}{d_A d_B} + \frac{1}{R(\rho_{AB})}\left(\rho_{AB} - \sigma_{AB}\right)\right)$$

is a separable state (cf. Theorem 1 of [Gurvits and Barnum, 2002]), from which it follows that σ_{AB} is separable as well, as a convex combination of separable states.

For the last point, it suffices to prove the statement in the case of two states. If $\lambda = \min(R(\rho_{AB}), R(\sigma_{AB}))$, then

$$\frac{1}{1-\lambda}\left(\rho_{AB} - \lambda\frac{\mathbb{1}_A \otimes \mathbb{1}_B}{d_A d_B}\right) \in \mathrm{SEP}(A:B), \quad \text{and} \quad \frac{1}{1-\lambda}\left(\sigma_{AB} - \lambda\frac{\mathbb{1}_A \otimes \mathbb{1}_B}{d_A d_B}\right) \in \mathrm{SEP}(A:B)$$

and so for any $t \in [0,1]$,

$$\frac{1}{1-\lambda}\left(t\rho_{AB} + (1-t)\sigma_{AB} - \lambda\frac{\mathbb{1}_A \otimes \mathbb{1}_B}{d_A d_B}\right) \in \mathrm{SEP}(A:B)$$

too, so $R(t\rho_{AB} + (1-t)\sigma_{AB}) \geq \lambda = \min(R(\rho_{AB}), R(\sigma_{AB}))$ as desired. $\qquad\square$

Remark 14.3.3. Proposition 3 of [Lami and Giovannetti, 2016] shows that $\rho_A \otimes \rho_B \in \mathrm{int}\,\mathrm{SEP}(A:B)$ when ρ_A and ρ_B are full rank. In light of (14.13), the relation (14.12) strengthens this result by giving a quantitative bound:

$$B_2((\rho_A)_{\min}(\rho_B)_{\min}, \rho_A \otimes \rho_B) \subset \mathrm{SEP}(A:B),$$

where $B_p(r, \rho_{AB})$ is the closed ball in $\mathcal{T}_p(\mathcal{H})$-norm of radius r around ρ_{AB}.

Admittedly, it is not a priori clear how to obtain good lower bounds on this quantity for general separable states ρ_{AB} and we leave this for future work. Also note that for the state τ given in (14.10), we have $R(\tau) = 0$.

14.3.2. Upper bounds

We know from Chapter 8 that any finite dimensional faithful quantum Markov semigroup $(\mathcal{P}_t)_{t\geq 0}$ on $\mathcal{B}(\mathcal{H})$ satisfies the so-called *strong decoherence property* (SD): there exist constants $K, \gamma > 0$, possibly depending on $d_{\mathcal{H}}$, such that for any initial state ρ:

$$\|\mathcal{P}_{t*}(\rho - E_{\mathcal{N}*}(\rho))\|_1 \leq K\,\mathrm{e}^{-\gamma t}\,. \tag{SD}$$

Good control over the constants K and γ can be achieved from the functional inequalities studied in Part IV.

Here, we combine the tools gathered in the last subsection, namely estimates on the radius of balls surrounding tensor product states together with the strong decoherence property, in order to estimate from above the entanglement loss in the different situations defined in Section 14.1. This strategy will work only if either the conditional expectation $E_{\mathcal{N}*}$ (in continuous time case), or the projection $\mathcal{P}_{\tilde{\mathcal{N}}(\Phi)}$ (in the discrete time case) output full-rank, separable states whose robustness of separability is uniformly lower bounded by a strictly positive constant. In the entanglement-breaking setting, this is true in particular when the quantum Markov chain is primitive, and the condition becomes necessary in the continuous time setting (cf. Lemma 14.2.7). For the other entanglement loss times, the situation is less clear and we leave the study of these cases to future work.

Proposition 14.3.4. *Let $(\mathcal{P}_t)_{t\geq 0}$ be primitive on $\mathcal{B}(\mathcal{H})$ with full-rank invariant state σ and generator $\mathcal{L}$. Then, assuming that* SD *holds for $(\mathcal{P}_t \otimes \mathrm{id})_{t\geq 0}$:*

$$t_{\mathrm{EB}}((\mathcal{P}_t)_{t\geq 0}) \leq \frac{\ln\left(K\,d_{\mathcal{H}}\|\sigma^{-1}\|_\infty\right)}{\gamma}\,.$$

Proof. By Lemma 14.2.6, we know that the $\|\cdot\|_1$-norm around $\mathbb{1}/d_{\mathcal{H}} \otimes \sigma$ of radius $\sigma_{\min}/d_{\mathcal{H}}$ is included in $\mathrm{SEP}(\mathcal{H}, \mathcal{H})$. In the primitive case, $(\mathcal{P}_{t*} \otimes \mathrm{id}) \circ E_{\mathcal{N}*}(\rho) = \mathrm{Tr}[\rho]\sigma \otimes \frac{\mathbb{1}}{d_{\mathcal{H}}}$. It is therefore clear that for

any t such that $K \, \mathrm{e}^{-\gamma t} \leq \sigma_{\min}/d_{\mathcal{H}}$, $\mathrm{id} \otimes \mathcal{P}_{t*}(|\Upsilon\rangle\langle\Upsilon|)$ is separable, which implies that the channel $\mathcal{P}_{t*}$ itself is entanglement breaking.

$\square$

This proof can also be adapted to get upper bounds on the entanglement annihilating time of a tensor product of semigroups.

14.4. Lower bounds on entanglement loss

Here, we derive lower bounds on the time it takes a Markov semigroup $(\mathcal{P}_t)_{t \geq 0}$ to become entanglement breaking. The idea is simply to use the useful fact that the set of PPT states includes the one of separable states. We recall that a state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is said to have a *positive partial transpose* (PPT) if the operator $(\mathrm{id} \otimes \mathcal{T})(\rho)$ is positive, where the superoperator $\mathcal{T}$ denotes the transposition with respect to any basis (see Proposition 2.11 of [Aubrun and Szarek, 2017]).

14.4.1. Sufficient conditions for entanglement loss

In the next lemma, given a channel Φ we find necessary conditions on k for Φ^k to be 2-locally entanglement annihilating.

Lemma 14.4.1. *Let* $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ *be a quantum channel with* $\det(\Phi) \neq 0$[3]. *If* $\|\Phi^{-k} : \mathcal{T}_2(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})\| \leq d_{\mathcal{H}}$, *then* $\Phi^k \notin \mathrm{LEA}_2(\mathcal{H})$.

Proof. We wish to look for a sufficient condition to imply that Φ^k is not 2-locally entanglement annihilating. If $\Phi^k \in \mathrm{LEA}_2(\mathcal{H})$, then $(\mathrm{id} \otimes \Psi) \circ \left(\left(\Phi^k\right)^* \otimes \left(\Phi^k\right)^* \right)$ would be a positive map for any positive map Ψ. As a map is positive if and only if its adjoint is positive, we have that $(\Phi^k \otimes \Phi^k) \circ (\mathrm{id} \otimes \Psi^*)$ is a positive map for all positive Ψ^*. In particular, this holds for $\mathcal{T}$, the transpose w.r.t. some fixed basis. Therefore, if $(\Phi^k \otimes \Phi^k) \circ (\mathrm{id} \otimes \mathcal{T})(|\Upsilon\rangle\langle\Upsilon|) \not\geq 0$, then $\Phi^k \notin \mathrm{LEA}_2(\mathcal{H})$. We have

$$d_{\mathcal{H}}(\Phi^k \otimes \Phi^k) \circ (\mathrm{id} \otimes \mathcal{T})(|\Upsilon\rangle\langle\Upsilon|) \equiv \Phi^k \otimes \Phi^k(F). \tag{14.15}$$

Therefore, any witness $X_{AB} \geq 0$ with

$$\mathrm{Tr}\left[X_{AB}(\Phi^k \otimes \Phi^k)(F)\right] < 0 \tag{14.16}$$

certifies that $\Phi^k \notin \mathrm{LEA}_2(\mathcal{H})$. We rewrite (14.16) as

$$\mathrm{Tr}\left[((\Phi^k)^* \otimes (\Phi^k)^*)(X_{AB})\,F\right] < 0. \tag{14.17}$$

Take $X_{AB} = P_{\mathrm{asym}} = \frac{\mathbb{1} \otimes \mathbb{1} - F}{2}$. Then the condition becomes

$$d_{\mathcal{H}} = \mathrm{Tr}\left[(\Phi^k)^*(\mathbb{1}) \otimes (\Phi^k)^*(\mathbb{1})F\right] < \mathrm{Tr}\left[((\Phi^k)^* \otimes (\Phi^k)^*)(F)F\right]. \tag{14.18}$$

---

[3] Here, $\det(\Phi)$ simply denotes the product of the eigenvalues of Φ.

where $\mathrm{Tr}[(\Phi^k)^*(\mathbb{1}) \otimes (\Phi^k)^*(\mathbb{1}) F] = \mathrm{Tr}[F] = d_{\mathcal{H}}$. The right-hand side is given by

$$\mathrm{Tr}[((\Phi^k)^* \otimes (\Phi^k)^*)(F) F] = \sum_{ij} \mathrm{Tr}[((\Phi^k)^*(|e_i\rangle\langle e_j|) \otimes (\Phi^k)^*(|e_j\rangle\langle e_i|)) F] \qquad (14.19)$$

$$= \sum_{ij} \mathrm{Tr}[(\Phi^k)^*(|e_i\rangle\langle e_j|)(\Phi^k)^*(|e_j\rangle\langle e_i|)] \qquad (14.20)$$

$$= \sum_{ij} \langle (\Phi^k)^*(|e_j\rangle\langle e_i|), (\Phi^k)^*(|e_j\rangle\langle e_i|)\rangle_{\mathrm{HS}}. \qquad (14.21)$$

Now note that

$$\inf_{A\neq 0} \frac{\|(\Phi^k)^*(A)\|_2}{\|A\|_2} = \inf_{B\neq 0} \frac{\|B\|_2}{\|(\Phi^{-k})^*(B)\|_2} = \frac{1}{\|\Phi^{-k} : \mathcal{T}_2(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})\|},$$

which can be seen by taking $A = \Phi^{-k}(B)$. Thus, it follows that

$$\sum_{ij} \langle (\Phi^k)^*(|e_j\rangle\langle e_i|), (\Phi^k)^*(|e_j\rangle\langle e_i|)\rangle_{\mathrm{HS}} \geq \frac{1}{\|\Phi^{-k} : \mathcal{T}_2(\mathcal{H}) \to \mathcal{T}_2(\mathcal{H})\|} d_{\mathcal{H}}^2$$

and we obtain the claim. $\qquad\square$

Remark 14.4.2. Since $\mathrm{EB}(\mathcal{H}) \subset \mathrm{LEA}_2(\mathcal{H})$, these also constitute conditions for Φ^k to be entanglement breaking. This is because if a channel Ψ is entanglement breaking, then $\Psi \otimes \Psi$ is 2-locally entanglement annihilating.

Corollary 14.4.3. *Let* $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ *be a quantum channel. If*

$$\|\Phi\|_2 > \sqrt{d_{\mathcal{H}}}, \qquad (14.22)$$

then $\Phi \notin \mathrm{LEA}_2(\mathcal{H})$, *and in particular,* $\Phi \notin \mathrm{EB}(\mathcal{H})$.

Proof. Note that, for $k = 1$, the right-hand side of (14.18) can be rewritten as $\|J(\Phi)\|_2^2$, where J is the unnormalized Choi-Jamiolkowski isomorphism defined in Equation (14.1), since

$$\mathrm{Tr}[\Phi^* \otimes \Phi^*(F)F] = \mathrm{Tr}[\Phi \otimes \Phi(F)F]$$

$$= \sum_{ij} \mathrm{Tr}[\Phi(|e_i\rangle\langle e_j|) \otimes \Phi(|e_j\rangle\langle e_i|)F]$$

$$= \sum_{ij} \mathrm{Tr}[\Phi(|e_i\rangle\langle e_j|)\Phi(|e_j\rangle\langle e_i|)]$$

$$= \sum_{ij} \langle \Phi(|e_j\rangle\langle e_i|), \Phi(|e_j\rangle\langle e_i|)\rangle_{\mathrm{HS}}$$

$$= \sum_{ij} \|\Phi(|e_j\rangle\langle e_i|)\|_2^2$$

$$= \|\sum_{ij} |e_i\rangle\langle e_j| \otimes \Phi(|e_i\rangle\langle e_j|)\|_2^2$$

$$= \|J(\Phi)\|_2^2$$

using $\mathrm{Tr}[(A \otimes B)F] = \mathrm{Tr}[AB]$, that the squared 2-norm of a block matrix is the sum of the squared 2-norms of each submatrix, and a matrix representation for the unnormalized Choi matrix of the channel, $J(\Phi)$. Thus, if $\|J(\Phi)\|_2^2 > d_{\mathcal{H}}$, then $\Phi \notin \mathrm{LEA}_2(\mathcal{H})$. We have

$$\|J(\Phi)\|_2^2 = \mathrm{Tr}(\Phi \otimes \mathrm{id}(|\Upsilon\rangle\langle\Upsilon|))((\Phi \otimes \mathrm{id}(|\Upsilon\rangle\langle\Upsilon|)) = \mathrm{Tr}((\Phi^*\Phi \otimes \mathrm{id}(|\Upsilon\rangle\langle\Upsilon|))|\Upsilon\rangle\langle\Upsilon|).$$

Then

$$\operatorname{Tr} \Phi^* \Phi \otimes \operatorname{id} (|\Upsilon\rangle\langle\Upsilon|) |\Upsilon\rangle\langle\Upsilon| = \sum_{i,j,k,l=1}^{d} \operatorname{Tr} \left[\Phi^* \Phi (|e_i\rangle\langle e_j|) \otimes |e_i\rangle\langle e_j| \right] |e_l\rangle\langle e_k| \otimes |e_l\rangle\langle e_k|$$

$$= \sum_{i,j=1}^{d} \operatorname{Tr} \Phi^* \Phi (|e_i\rangle\langle e_j|) (|e_i\rangle\langle e_j|)^\dagger.$$

Note that $\{|e_i\rangle\langle e_j|\}_{i,j=1}^{d_\mathcal{H}}$ is an orthonormal basis of $\mathcal{B}(\mathcal{H})$ and $\operatorname{Tr} \Phi^* \Phi (|e_i\rangle\langle e_j|) (|e_i\rangle\langle e_j|)^\dagger$ corresponds to the Hilbert-Schmidt scalar product between $\Phi^* \Phi (|e_i\rangle\langle e_j|)$ and $|e_i\rangle\langle e_j|$. Therefore, we have that

$$\|J(\Phi)\|_2^2 = \operatorname{Tr} \Phi^* \Phi.$$

$\square$

14.4.2. Lower bounds

In the next theorem, we derive a lower bound on t_{EB} for a quantum Markov semigroup using Corollary 14.4.3.

Theorem 14.4.4 (Lower bound for t_{EB})**.** *For any QMS $(\mathcal{P}_t)_{t\geq 0}$ on $\mathcal{B}(\mathcal{H})$,*

$$t_{\mathrm{EB}}((\mathcal{P}_t)_{t\geq 0}) > \frac{\ln(d_\mathcal{H} + 1)}{2 \max_j |\operatorname{Re}(\lambda_j(\mathcal{L}))|},$$

where $\{\lambda_j\}_{j=1}^{d_\mathcal{H}^2}$ are the eigenvalues of $\mathcal{L}$, the generator of the QMS. In the case that $(\mathcal{P}_t)_{t\geq 0}$ is reversible with respect to a faithful state σ, $\mathcal{L}$ is self-adjoint with respect to $\langle \cdot, \cdot \rangle_\sigma$, and $\max_j |\operatorname{Re}(\lambda_j(\mathcal{L}))| = \|\mathcal{L} : \mathbb{L}_2(\sigma) \to \mathbb{L}_2(\sigma)\|$ is the largest eigenvalue (in modulus) of $\mathcal{L}$.

Proof. We have

$$\|\mathcal{P}_{t*}\|_2^2 \geq \sum_{i=1}^{d_\mathcal{H}^2} |e^{t\lambda_j(\mathcal{L})}|^2 = \sum_{i=1}^{d_\mathcal{H}^2} e^{2t\operatorname{Re}(\lambda_j(\mathcal{L}))} = \sum_{i=1}^{d_\mathcal{H}^2} e^{-2t|\operatorname{Re}(\lambda_j(\mathcal{L}))|} \geq 1 + (d_\mathcal{H}^2 - 1) e^{-2t \max_j |\operatorname{Re}(\lambda_j(\mathcal{L}))|}$$

using that $(\mathcal{P}_{t*})_{t\geq 0}$ is trace-preserving, so $\mathcal{L}$ must have a zero eigenvalue. By Corollary 14.4.3 it follows that

$$1 + (d_\mathcal{H}^2 - 1) e^{-2t \max_j |\operatorname{Re}(\lambda_j(\mathcal{L}))|} > d_\mathcal{H}$$

is a sufficient condition for the semigroup not to be entanglement breaking at time t and the claim follows after rearranging the terms. $\square$

Chapter 15.

Channel coding

In this chapter, we are interested in the estimation of the optimal amount of information that can be sent for different information processing tasks involving quantum inputs and noise. We consider the following tasks: in *classical channel coding*, we are interested in quantifying the amount of information that a sender (Alice) can transmit reliably to a receiver (Bob) using a noisy quantum channel. In particular, since the information is allowed to be encoded in an entangled state, we expect the maximal amount of information transmitted using a memoryless quantum channel to be higher than in the setting of classical communication through a classical channel. Alternatively, in the task of *quantum communication*, Alice wants to transmit a quantum state to Bob. This task is closely related to the one of *private classical communication*, where Alice wants to send private classical information to Bob reliably. These tasks were originally considered in the *asymptotic regime*, i.e. in the limit of arbitrarily large number of uses of the quantum channel. The question then reduces to the one of finding the optimal asymptotic rate of information sent per channel use. This rate is called the *capacity* of the channel for the task under consideration.

For all the tasks that are mentioned above, there exists a *quantum channel coding theorem* which shows that each of their corresponding *capacities*, i.e. of their optimal achievable asymptotic rates, can be expressed in terms of an appropriate entropic quantity. Any proof of a coding theorem consists of two parts: the *direct (or achievability) part* establishes a lower bound on the capacity by providing a protocol that achieves the task under consideration, whereas the *converse part* establishes an upper bound by proving that any protocol will fail in reliably achieving a better rate than the channel capacity. If the probability of error made by trying to achieve a rate that lies above capacity converges to 1 exponentially fast in the limit of infinitely many uses of the channel, the task is said to satisfy a *strong converse property*.

The main difficulty of quantum channel coding in comparison to its classical analogue lies in the fact that the entropic quantities characterizing most of the capacities mentioned above are in general intractable. For instance, the classical capacity $\mathcal{C}(\Phi)$ of a quantum channel Φ is characterized by the *regularized Holevo information*:

$$\mathcal{C}(\Phi) = \chi_{\mathrm{reg}}(\Phi) \equiv \lim_{n\to\infty} \frac{1}{n}\chi(\Phi^{\otimes n}),$$

where the *Holevo information* $\chi(\Phi)$ of a quantum channel Φ is defined in Equation (15.2). In general, the regularized Holevo information does not reduce to its *single-letter* expression: $\chi_{\mathrm{reg}}(\Phi) \neq \chi(\Phi)$, in sharp contrast with the classical setting. This is due to the so-called *superadditivity of the Holevo*

information: there exist channels Φ_1 and Φ_2 such that

$$\chi(\Phi_1 \otimes \Phi_2) > \chi(\Phi_1) + \chi(\Phi_2).$$

Fortunately, recent progress has been made in finding good *strong converse bounds* on various capacities. In the case when the channel is assumed to be time dependent, and in fact arises from a quantum Markov semigroup, we can therefore use the transference methods of Chapter 9 in order to estimate the behavior of the capacity as a function of time. In particular, since entanglement decreases with time, we expect the capacity to follow the same pattern (cf. Chapter 14).

In Section 13.2, we showed how reverse hypercontractivity can be used as a technical tool in order to find finite blocklength strong converse bounds for the task of asymmetric binary quantum hypothesis testing. Since this task is at the core of many protocols in quantum Shannon theory, it is not surprising that the technique found in Section 13.2 is generalizable to other quantum information processing tasks. In this chapter, we prove a finite blocklength strong converse bound for the classical capacity of *classical-quantum* channels.

Layout of the chapter: In Section 15.1, we recall some quantum information processing tasks, among which classical, private and quantum communication. Strong converse bounds available in the literature for each of these tasks are also reviewed. In, Section 15.2, we use functional inequalities together with the transference method of Chapter 9 in order to estimate the behavior of the capacities corresponding to the tasks introduced in Section 15.1. We end this chapter in Section 15.3 by providing a finite blocklength strong converse bound for the classical capacity of classical-quantum channels via quantum reverse hypercontractivity.

15.1. Channel coding with quantum resources

15.1.1. Classical communication

We start by recalling the protocol for classical communication through a quantum channel: here, Alice wants to reliably send a message m belonging to some message set $\mathcal{M}$ to Bob through a channel $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$.

More precisely, let the message sent by Alice be modeled by a random variable M taking values in $\mathcal{M}$ and that we assume to be uniformly distributed. Assume that Alice wants to use the channel n times. The encoding of the input state sent by Alice is modeled by the map $\mathcal{E}^n : \mathcal{M} \to \mathcal{D}(\mathcal{H}_A^{\otimes n})$. The decoding of Alice's message by Bob is then modeled by an $\mathcal{M}$-outcome POVM $\mathbf{D}^n \coloneqq \{D^n(\{m\}) : m \in \mathcal{M}\}$ on $\mathcal{H}_B^{\otimes n}$.

Then, a number $\alpha > 0$ is said to be an *achievable rate for the transmission of classical information through* Φ if for any $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, there exists an encoding $\mathcal{E}^n : \mathcal{M} \to \mathcal{D}(\mathcal{H}_A^{\otimes n})$ and a decoding $\mathbf{D}^n = \{D^n(\{m\}) : m \in \mathcal{M}\}$, where $|\mathcal{M}| = 2^{\lfloor n\alpha \rfloor}$, such that the maximum probability of error in decoding the message sent by Alice

$$p_e^{\max, n} = \max_{m \in \mathcal{M}} \mathrm{Tr}\left[\left(\mathbb{1}_{\mathcal{H}_B^{\otimes n}} - D^n(\{m\}) \right) \Phi^{\otimes n} \circ \mathcal{E}^n(m) \right] \leq \varepsilon.$$

We recall that, given a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the *quantum mutual information* $I(A:B)_\rho$

is defined as[1]

$$I(A:B)_\rho = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) = \inf_{\sigma_B} D(\rho_{AB} \| \rho_A \otimes \sigma_B). \tag{15.1}$$

The *classical capacity* of the channel Φ, denoted by $\mathcal{C}(\Phi)$, is then defined as the supremum over all achievable rates for the transmission of classical information through Φ. The following capacity theorem is due to [Holevo, 1973, Schumacher and Westmoreland, 1997, Holevo, 2006]:

Theorem 15.1.1. *The classical capacity of a quantum channel is characterized by its* regularized Holevo information*:*

$$\mathcal{C}(\Phi) = \lim_{n \to \infty} \frac{\chi(\Phi^{\otimes n})}{n},$$

where $\chi(\Phi)$ *denotes the* Holevo information *of the channel* Φ *and is defined as*

$$\chi(\Phi) \coloneqq \max_\rho I(X:B)_\rho,$$

with the maximum being taken over all classical-quantum states of the form:

$$\rho^{XB} = \sum_x p_X(x) |e_x\rangle\langle e_x| \otimes \Phi(\rho_x).$$

Classical-quantum channels We recall that given an alphabet $\mathcal{X}$ and a Hilbert space $\mathcal{H}$, a *classical-quantum channel* (or c-q channel) is a map $\mathcal{W} : \mathcal{X} \to \mathcal{D}(\mathcal{H})$. Now, assume Alice wants to communicate a message m belonging to a set $\mathcal{M}$ of possible messages making n uses of the c-q channel $\mathcal{W}$. She first encodes her messages into codewords $x^n \coloneqq x_1 \dots x_n \equiv \mathcal{E}^n(m) \in \mathcal{X}^n$, making use of an encoder $\mathcal{E}^n$, and sends these codewords through n uses of the c-q channel $\mathcal{W}$. Bob then measures the output states of the channel via a POVM $\mathbf{D}^n = \{D^n(\{m\}) : m \in \mathcal{M}\}$ and decodes a messages $\hat{m} \in \mathcal{M}$. For c-q channels, the Holevo capacity is known to be additive, and hence the classical capacity is given by

$$\mathcal{C}(\mathcal{W}) \equiv \min_{\sigma \in \mathrm{im}(\mathcal{W})} \max_{\rho \in \mathrm{im}(\mathcal{W})} D(\rho \| \sigma).$$

15.1.2. Entanglement-assisted classical communication

The entanglement-assisted classical capacity of a channel Φ is defined in a similar way to the classical capacity, except that one assumes that the sender and receiver may share unlimited amount of entanglement before starting to send information over the channel. The prior shared of entanglement which they can employ in the protocol can lead to a significant increase in the classical capacity of a quantum channel.

More precisely, let $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ be a quantum channel. A number $\alpha > 0$ is said to be an *achievable rate for the transmission of classical information with entanglement assistance* through Φ if for any $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, there exist Hilbert spaces $\mathcal{H}'_A$ and $\mathcal{H}'_B$ of arbitrary dimension, a bipartite state $\gamma_{A'B'} \in \mathcal{D}(\mathcal{H}'_A \otimes \mathcal{H}'_B)$, *encoding* channels $\mathcal{E}^n_m : \mathcal{B}(\mathcal{H}'_A) \to \mathcal{B}(\mathcal{H}^{\otimes n}_A)$, $m \in \mathcal{M}$, and a *decoding* $\mathbf{D}^n = \{D^n(\{m\}) : m \in \mathcal{M}\}$ on $\mathcal{H}^{\otimes n}_B \otimes \mathcal{H}'_B$, where $|\mathcal{M}| = 2^{\lfloor n\alpha \rfloor}$, such that

$$p^{\max, n}_e = \max_{m \in \mathcal{M}} \mathrm{Tr}\left[(\mathbb{1}_{\mathcal{H}^{\otimes n}_B \otimes \mathcal{H}'_B} - D^n(\{m\})) (\Phi^{\otimes n} \circ \mathcal{E}^n_m \otimes \mathrm{id}_{\mathcal{H}'_B \to \mathcal{H}'_B})(\gamma) \right] \leq \varepsilon.$$

---

[1]Since it is more natural in quantum information theory, the logarithms in this section are taken in base 2 and denoted by log, as opposed to the other chapters.

The *entanglement-assisted classical capacity* of Φ, denoted by $\mathcal{C}_{\mathrm{EA}}(\Phi)$, is then the supremum over all achievable rates for the entanglement assisted classical information transmission through Φ.

The following result, due to [Bennett et al., 2002, Bennett et al., 1999, Horodecki et al., 2009] constitutes an extension of Theorem 15.1.1 to the entanglement-assisted scenario:

Theorem 15.1.2. *The entanglement-assisted classical capacity of a channel Φ is additive: for any k uses of the channel: $\mathcal{C}_{\mathrm{EA}}(\Phi^{\otimes k}) = k\,\mathcal{C}_{\mathrm{EA}}(\Phi)$. Moreover, given a quantum channel $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$;*

$$\mathcal{C}_{\mathrm{EA}}(\Phi) = \chi_{\mathrm{EA}}(\Phi),\tag{15.2}$$

where $\chi_{\mathrm{EA}}(\Phi)$ denotes the entanglement-assisted Holevo capacity *of the channel Φ and is defined as*

$$\chi_{\mathrm{EA}}(\Phi) := \sup_{\eta} I(A:B)_{(\mathrm{id}_{\mathcal{H}_A}\otimes\Phi)(\eta)},\tag{15.3}$$

where the supremum is taken over any pure state η on $\mathcal{H}_A \otimes \mathcal{H}_A$.

It was shown in [Berta et al., 2011, Bennett et al., 2014, Gupta and Wilde, 2015] that the entanglement assisted capacity has the strong converse property: if the rate of communication in any given coding scheme exceeds the capacity, the error probability tends to one in the limit $n \to \infty$ of channel uses.

15.1.3. Private classical communication

The private capacity quantifies the rate at which classical information that Alice and Bob would like to keep secret from an adversary Eve modeled by the environment can be reliably transmitted by a given quantum channel in the asymptotic limit of many channel uses. In loose terms, Alice wants to reliably send a message m belonging to a finite message set $\mathcal{M}$ to Bob through a channel $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ in such a way that Eve's state ends up close to a constant state, regardless of which message Alice transmits through the channel Φ, in the limit of a large number of uses of the channel Φ.

More precisely, let the message sent by Alice be modeled by a random variable M taking values in $\mathcal{M}$ and that we assume to be uniformly distributed. Assume that Alice is allowed multiple n uses of the channel Φ. Then, in order to make sure that the state retrieved by Eve is close to a fixed state $\omega^{E_n} \in \mathcal{B}(\mathcal{H}_E)$, $\mathcal{H}_E \equiv \mathcal{H}_B$, Alice randomizes her input state before sending it to the channel according to a uniformly distributed random variable K taking values in the so-called *privacy amplification set* $\mathcal{K}$. Therefore, the encoding of the input state sent by Alice is modeled by the map $\mathcal{E}^n : \mathcal{M} \times \mathcal{K} \to \mathcal{D}(\mathcal{H}_A^{\otimes n})$ so that, on average:

$$\mathcal{E}^n(M) := \mathbb{E}_K[\mathcal{E}^n(M,K)] = \frac{1}{|\mathcal{K}|}\sum_{k\in\mathcal{K}}\mathcal{E}^n(M,k).\tag{15.4}$$

The *decoding* of Alice's message by Bob is modeled by an $\mathcal{M} \times \mathcal{K}$-outcome POVM $\mathbf{D}^n := \{D^n(\{m,k\}) : (m,k) \in \mathcal{M} \times \mathcal{K}\}$ on $\mathcal{H}_B^{\otimes n}$. The difference between private classical communication and the classical channel coding presented in Section 15.1.1 lies in the so-called *privacy condition*: recall that, given a channel $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$, there exists a channel $\hat{\Phi} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_E)$, $\mathcal{H}_E \equiv \mathcal{H}_B$, called a *complementary channel* of Φ, such that there exists an isometric extension $U_\Phi : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$ so that for any $\rho \in \mathcal{D}(\mathcal{H}_A)$:

$$\hat{\Phi}(\rho) = \mathrm{Tr}_B(U_\Phi\,\rho\,U_\Phi^*).\tag{15.5}$$

Then, given $\delta > 0$, δ-*privacy* is said to be achieved if there exists a state $\rho_0^n \in \mathcal{D}(\mathcal{H}_E^{\otimes n})$ so that

$$\max_{m \in \mathcal{M}} \|\hat{\Phi}(\mathcal{E}^n(m)) - \rho_0^n\|_1 \le \delta. \tag{15.6}$$

Note that the condition of δ-privacy (15.6) does not depend on the choice of the complementary channel, since it is well-known that any two such channels are equal up to a local unitary conjugation, and by invariance of $\|.\|_1$ under such unitaries.

Then, a number $\alpha > 0$ is said to be an *achievable rate for the transmission of private classical information* through Φ if for any $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$, there exists an encoding $\mathcal{E}^n : \mathcal{M} \times \mathcal{K} \to \mathcal{D}(\mathcal{H}_A^{\otimes n})$ and decoding $\mathbf{D}^n = \{D^n(\{m, k\}) : (m, k) \in \mathcal{M} \times \mathcal{K}\}$, where $|\mathcal{M}| = 2^{\lfloor n\alpha \rfloor}$, and a state $\rho_0^n \in \mathcal{D}(\mathcal{H}_E^{\otimes n})$, such that

$$p_e^{\max, n} = \max_{m \in \mathcal{M}, k \in \mathcal{K}} \mathrm{Tr}\left[\left(\mathbb{1}_{\mathcal{H}_B^{\otimes n}} - D^n(\{m, k\}) \right) \Phi^{\otimes n} \circ \mathcal{E}^n(m, k) \right] \le \varepsilon, \quad \text{and} \quad \max_{m \in \mathcal{M}} \|\hat{\Phi}(\mathcal{E}^n(m)) - \rho_0^n\|_1 \le \varepsilon.$$

The *private capacity* of the channel Φ, denoted by $\mathcal{P}(\Phi)$, is then defined as the supremum over all achievable rates for the transmission of private classical information through Φ. The following private capacity theorem is due to [Devetak, 2005, Cai et al., 2004]:

Theorem 15.1.3. *Given a quantum channel $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$:*

$$\mathcal{P}(\Phi) = \lim_{n \to \infty} \frac{P(\Phi^{\otimes n})}{n}, \tag{15.7}$$

where the private information *of the channel Φ is defined as follows:*

$$P(\Phi) = \max_{\rho^{XA}} \left(I(X : B)_\sigma - I(X : E)_\sigma \right),$$

with the maximization of the mutual informations being taken over all c-q states ρ^{XA}, and for $\sigma := U_\Phi \rho^{XA} U_\Phi^*$, where U_Φ any isometric extension of Φ.

15.1.4. Quantum communication

Roughly speaking, the quantum capacity of a channel is the average amount of qubits that can be accurately transmitted per use of the channel, in the limit of a large number of channel uses, acting on a collection of possibly entangled registers. More precisely:

Definition 15.1.4. Let $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ be a quantum channel. A number $\alpha > 0$ is said to be an *achievable rate for the transmission of quantum information* through Φ if for any $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$, there exist an *encoding* channel $\mathcal{E}^n : \mathcal{B}((\mathbb{C}^2)^{\otimes \lfloor \alpha n \rfloor}) \to \mathcal{B}(\mathcal{H}_A^{\otimes n})$ and a *decoding* map $\Phi_D^n : \mathcal{B}(\mathcal{H}_B^{\otimes n}) \to \mathcal{B}((\mathbb{C}^2)^{\otimes \lfloor \alpha n \rfloor})$ such that

$$\| \mathrm{id}_{\mathcal{B}(\mathbb{C}^2)}^{\otimes \lfloor \alpha n \rfloor} - \Phi_D^n \circ \Phi^{\otimes n} \circ \mathcal{E}^n \|_\diamond \le \varepsilon. \tag{15.8}$$

The *quantum capacity* of Φ, denoted by $\mathcal{Q}(\Phi)$, is then defined as the supremum over all achievable rates for the quantum information transmission through Φ.

We recall that, given a state $\sigma \in \mathcal{D}(\mathcal{H}_A)$ and a channel $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$, the *coherent information* of the channel Φ corresponding to the state σ is defined as

$$I_c(\sigma; \Phi) := S(\Phi(\sigma)) - S((\Phi \otimes \mathrm{id}_{\mathcal{B}(\mathcal{H}_A)})(|\psi\rangle\langle\psi|)), \tag{15.9}$$

where $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_A$ denotes any purification of the state σ. Namely, the coherent information quantifies the correlations that persist after Φ is applied to a purification of σ. Next, the *maximum coherent information* of the channel Φ is defined as

$$I_c(\Phi) := \max_{\sigma \in \mathcal{D}(\mathcal{H}_A)} I_c(\sigma; \Phi). \tag{15.10}$$

Next, we recall the quantum capacity theorem, due to [Lloyd, 1997, Devetak, 2005, Hayden et al., 2008]:

Theorem 15.1.5. *Given a quantum channel* $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$:

$$\mathcal{Q}(\Phi) = \lim_{n \to \infty} \frac{I_c(\Phi^{\otimes n})}{n}.$$

15.2. Application to the estimation of capacities of QMS

In this section, we obtain strong converse bounds on the private and quantum capacities of a reversible quantum dynamical semigroups $(\mathcal{P}_t)_{t \geq 0}$ converging to its associated conditional expectation $E_{\mathcal{F}}$, where $\mathcal{F} \subset \mathcal{B}(\mathcal{H})$ is the algebra of fixed points of $(\mathcal{P}_t)_{t \geq 0}$. These will mainly be based on relating norm estimates to bounds on entropic quantities derived from the sandwiched Rényi entropies (cf. Section 1.2). As seen in Section 15.1, these relative entropies constitute the basic quantities used in numerous strong converse results for various capacities of quantum channels. Intuitively speaking, as the QMS $(\mathcal{P}_t)_{t \geq 0}$ converges to $E_{\mathcal{F}}$, we expect that its capacity also converges to that of the conditional expectation, and we wish to quantify this convergence using non-commutative functional analytical tools. We recall for $\mathcal{H} \equiv \mathbb{C}^d$ the following standard decomposition of the fixed point algebra $\mathcal{F} \subset \mathcal{B}(\mathcal{H}) \equiv \mathbb{M}_d$:

$$\mathcal{F} = \bigoplus_k \mathbb{M}_{d_k} \otimes \mathbb{1}_{n_k}. \tag{15.11}$$

15.2.1. Entropy Comparison Theorem

In [Gao et al., 2017], the authors proved the following factorization property: given the representation $\alpha : g \mapsto \alpha_g(.) = U_g(.)U_g^*$ of a finite or compact Lie group G on the algebra $\mathcal{B}(\mathcal{H})$ of linear operators on a finite dimensional Hilbert space $\mathcal{H}$, and for any $t \geq 0$, define the co-representation $\pi : g \mapsto \alpha_{g^{-1}}(.)$. Then we may transfer properties of completely positive maps on $\mathbb{L}_\infty(G)$ to completely positive maps on $\mathcal{B}(\mathcal{H})$. Indeed, for every positive function k on G, we define

$$\Phi_k(X) := \int_G k(g) U_g^* X U_g \, d\mu_G(g). \tag{15.12}$$

Here μ_G is the Haar measure. Similarly to what has been already discussed in Section 9.1, since $d_{\mathcal{H}}^{-1}\mathbb{1}$ is a full-rank invariant state of Φ_k, the set $\mathcal{F}$ of fixed points of Φ_k is an algebra (see Theorem 6.12 of [Wolf, 2012]) and is given by the commutant of $\{U_g\}_{g \in G}$ (Theorem 6.13 of [Wolf, 2012]):

$$\mathcal{F} = \{X \in \mathcal{B}(\mathcal{H}) \,|\, X U_g = U_g X\} = \{U_g\}'_{g \in G}.$$

Note that the following natural bimodule property holds: for any $X_1, X_2 \in \mathcal{F}$:

$$\Phi_k(X_1 X X_2) = X_1 \Phi_k(X) X_2.$$

We then have

$$\pi \circ \Phi_k = (\varphi_k \otimes \mathrm{id}_{\mathcal{B}(\mathcal{H})}) \circ \pi \,,$$

where $\varphi_k : \mathbb{L}_\infty(G) \to \mathbb{L}_\infty(G)$ is defined by

$$\varphi_k(f)(g) = \int k(gh^{-1}) f(h) \, d\mu_G(h) \,.$$

We will denote by

$$E_{\mathcal{F}}(\rho) = \int U_g^* \rho U_g \, d\mu_G(g)$$

the conditional expectation onto the fixed-point algebra. A commuting square of the form of Equation (9.3) then holds, which in particular implies that the natural inclusion

$$\mathbb{L}_p^q(\mathcal{F} \subset \mathcal{B}(\mathcal{H})) \subset \mathbb{L}_p^q(\mathcal{B}(\mathcal{H}) \subset \mathbb{L}_\infty(G, \mathcal{B}(\mathcal{H})))$$

is completely isometric (see [Junge and Parcet, 2010] for more details).

The next theorem constitutes the basis of all the capacity estimates that we provide in the sequel:

Theorem 15.2.1. *Let $X \in \mathcal{F}^+$ and $k : G \to \mathbb{R}^+$ a bounded measurable function such that $\int k \, d\mu_G = 1$. Then, for any $Y \in \mathcal{B}(\mathcal{H})$, and any $p \geq 1$ of Hölder conjugate $\hat{p}$:*

$$\|X^{-\frac{1}{2\hat{p}}} \Phi_k(Y) X^{-\frac{1}{2\hat{p}}}\|_{\mathbb{L}_p(d_{\mathcal{H}}^{-1} \mathbb{1}_{\mathcal{H}})} \leq \|k\|_{\mathbb{L}_p(\mu_G)} \|X^{-\frac{1}{2\hat{p}}} E_{\mathcal{F}}(Y) X^{-\frac{1}{2\hat{p}}}\|_{\mathbb{L}_p(d_{\mathcal{H}}^{-1} \mathbb{1}_{\mathcal{H}})} \,.$$

Moreover, for any states $\rho, \sigma \in \mathcal{D}_+(\mathcal{H})$ such that $\sigma \in \mathcal{F}^+$:

$$D(\Phi_k(\rho)\|\sigma) \leq D(E_{\mathcal{F}}(\rho)\|\sigma) + \int_G k \log k \, d\mu_G \,. \tag{15.13}$$

In particular, choosing $\sigma = E_{\mathcal{F}}(\rho)$:

$$D(\Phi_k(\rho)\|E_{\mathcal{F}}(\rho)) \leq \int_G k \log k \, d\mu_G \,. \tag{15.14}$$

The proof of Theorem 15.2.1 should be interpreted as an extension of the one of Theorem 1 of [Gao et al., 2018a] to the case of an infinite dimensional classical environment. First let us introduce some notations. Given two (possibly infinite dimensional) Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, we denote by $\mathcal{T}_p(\mathcal{H}, \mathcal{K})$ the space of operators $X : \mathcal{H} \to \mathcal{K}$ with norm

$$\|X\|_{\mathcal{T}_p(\mathcal{H},\mathcal{K})} := \|XX^*\|_{\mathcal{T}_{p/2}(\mathcal{K})}^{1/2} = \|X^*X\|_{\mathcal{T}_{p/2}(\mathcal{H})}^{1/2} \,. \tag{15.15}$$

Next, given two Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ and a positive invertible element $\sigma \in \mathcal{B}(\mathcal{H})$ we recall that the Kosaki norms

$$\|X\|_{\widehat{\mathbb{L}}_p(\sigma)} := \|\sigma^{\frac{1}{p}} X\|_{\mathcal{T}_p(\mathcal{H},\mathcal{K})}$$

form an interpolation family (see. Theorem 4 of [Gao et al., 2018a] for more details):

$$\left[\widehat{\mathbb{L}}_\infty(\sigma), \widehat{\mathbb{L}}_2(\sigma)\right]_{\frac{1}{p}} = \widehat{\mathbb{L}}_{2p}(\sigma) \,.$$

Importantly, we recall that if a pair of spaces X_0 and X_1 form an interpolation scale $[X_0, X_1]_\theta$, and if

$Y_0 \subseteq X_0$ and $Y_1 \subseteq X_1$ are given by the *same* projector, then $[Y_0, Y_1]_\theta$ is also an interpolation scale. We are now ready to prove Theorem 15.2.1:

Proof of Theorem 15.2.1 : Let $\mathcal{H}'$ be a copy of $\mathcal{H}$, and define the following dual operators

$$
\hat{\eta} : \begin{cases} \mathcal{H} \otimes \mathbb{L}_2(\mu_G) \to \mathcal{H}' \\ \chi \mapsto \int_G U_g^* \sqrt{\rho}\, \chi(g)\, d\mu_G(g) \end{cases}
\qquad , \qquad
\hat{\eta}^* : \begin{cases} \mathcal{H}' \to \mathcal{H} \otimes \mathbb{L}_2(\mu_G) \\ \psi \mapsto (g \mapsto \sqrt{\rho}\, U_g \psi). \end{cases}
$$

The following then holds:

$$
\begin{aligned}
\|\sigma^{-\frac{1}{2p}} \hat{\eta} (\mathbb{1}_{\mathcal{H}} \otimes L_{k^{\frac{1}{2}}})\|^2_{\mathcal{T}_{2p}(\mathcal{H} \otimes \mathbb{L}_2(\mu_G), \mathcal{H}')} &= \|\sigma^{-\frac{1}{2p}} \hat{\eta} (\mathbb{1}_{\mathcal{H}} \otimes L_k) \hat{\eta}^* \sigma^{-\frac{1}{2p}}\|_{\mathcal{T}_p(\mathcal{H}')} \\
&= \left\| \int_G k(g)\, \sigma^{-\frac{1}{2p}} U_g^* \rho U_g\, d\mu_G(g) \sigma^{-\frac{1}{2p}} \right\|_{\mathcal{T}_p(\mathcal{H}')} \\
&= \|\sigma^{-\frac{1}{2p}} \Phi_k(\rho)\, \sigma^{-\frac{1}{2p}}\|_{\mathcal{T}_p(\mathcal{H}')}.
\end{aligned}
$$

Therefore, for the claim to hold, it suffices to show that

$$
\|\sigma^{-\frac{1}{2p}} \hat{\eta} (\mathbb{1}_{\mathcal{H}} \otimes L_{k^{\frac{1}{2}}})\|_{\mathcal{T}_{2p}(\mathcal{H} \otimes \mathbb{L}_2(\mu_G), \mathcal{H}')} \le \|k^{\frac{1}{2}}\|_{\mathbb{L}_{2p}(\mu_G)} \|\sigma^{-\frac{1}{2p}} \hat{\eta}\|^2_{\mathcal{T}_{2p}(\mathcal{H} \otimes \mathbb{L}_2(\mu_G), \mathcal{H}')}.
$$

This will be proved by complex interpolation: without loss of generality, choose the right hand side to be equal to 1. Then,

$$
\|k^{\frac{1}{2}}\|_{\mathbb{L}_{2p}(\mu_G)} = \left(\int_G k(g)^p\, d\mu_G(g) \right)^{\frac{1}{2p}} \ge \left(\int_G k(g)\, d\mu_G(g) \right)^{\frac{1}{2}} = 1,
$$

where we used Jensen's inequality and the normalization condition on k. This implies by assumption that $\|\sigma^{-\frac{1}{2p}} \hat{\eta}\|_{\mathcal{T}_{2p}(\mathcal{H} \otimes \mathbb{L}_2(\mu_G), \mathcal{H}')} \equiv \|\sigma^{\frac{1}{2p}} \xi_0\|_{\mathcal{T}_{2p}(\mathcal{H} \otimes \mathbb{L}_2(\mu_G), \mathcal{H}')} \le 1$, where $\xi_0 := \sigma^{-\frac{1}{2}} \hat{\eta}$. Next, there exists a continuous function on the strip $\xi : S = \{z : 0 \le \mathrm{Re}(z) \le 1\} \to \widehat{\mathbb{L}}_2(\sigma) + \widehat{\mathbb{L}}_\infty(\sigma)$ that is analytic on its interior and takes values in a finite dimensional subspace of $\widehat{\mathbb{L}}_2(\sigma) + \widehat{\mathbb{L}}_\infty(\sigma)$ such that $\xi(1/p) = \xi_0$ and for all $t \in \mathbb{R}$ [Stafney, 1969],

$$
\|\xi(it)\|_{\widehat{\mathbb{L}}_\infty(\sigma)}, \; \|\xi(1+it)\|_{\widehat{\mathbb{L}}_2(\sigma)} \le 1.
$$

Next, define the analytic function $T(z) := \sigma^{\frac{z}{2}} \xi(z) (\mathbb{1}_{\mathcal{H}} \otimes L_a^{pz})$, where $a := \dfrac{k^{\frac{1}{2}}}{\|k^{\frac{1}{2}}\|_{\mathbb{L}_{2p}(\mu_G)}}$. Then

$$
\|T(it)\|_{\mathcal{T}_\infty(\mathcal{H} \otimes \mathbb{L}_2(\mu_G), \mathcal{H}')} = \|\sigma^{\frac{it}{2}} \xi(it) (\mathbb{1}_{\mathcal{H}} \otimes L_a^{tip})\|_{\mathcal{T}_\infty(\mathcal{H} \otimes \mathbb{L}_2(\mu_G), \mathcal{H}')} = \|\xi(it)\|_{\widehat{\mathbb{L}}_\infty(\sigma)} \le 1,
$$

where we used the unitary invariance of Schatten norms. Similary,

$$
\begin{aligned}
\|T(1+it)\|_{\mathcal{T}_2(\mathcal{H} \otimes \mathbb{L}_2(\mu_G), \mathcal{H}')} &= \|\sigma^{\frac{1+it}{2}} \xi(1+it) (\mathbb{1}_{\mathcal{H}} \otimes L_a^{p(1+it)})\|_{\mathcal{T}_2(\mathcal{H} \otimes \mathbb{L}_2(\mu_G), \mathcal{H}')} \\
&= \|\sigma^{\frac{1}{2}} \xi(1+it) (\mathbb{1}_{\mathcal{H}} \otimes L_a^p)\|_{\mathcal{T}_2(\mathcal{H} \otimes \mathbb{L}_2(\mu_G), \mathcal{H}')} \\
&= \mathrm{Tr}\left(\sigma^{\frac{1}{2}} \xi(1+it) (\mathbb{1}_{\mathcal{H}} \otimes L_a^{2p}) \xi(1+it)^* \sigma^{\frac{1}{2}} \right) \\
&\le \|\xi(1+it)\|_{\widehat{\mathbb{L}}_2(\sigma)} \|a^p\|_{\mathbb{L}_2(\mu_G)} \le 1.
\end{aligned}
$$

We conclude by Stein's interpolation theorem (Theorem 1.1.10) that

$$
\|T(1/p)\|_{\mathcal{T}_{2p}(\mathcal{H} \otimes \mathbb{L}_2(\mu_G), \mathcal{H}')} = \frac{1}{\|k^{\frac{1}{2}}\|_{\mathbb{L}_2(\mu_G)}} \|\xi_0 (\mathbb{1}_{\mathcal{H}} \otimes L_{k^{\frac{1}{2}}})\|_{\widehat{\mathbb{L}}_{2p}(\sigma)} \le 1,
$$

which is what remained to be proved. Differentiation at $p = 1$ gives the entropic inequality. $\qquad\square$

15.2.2. Strong converse bounds on capacities of QMS

Given a primitive classical Markov semigroup $(P_t)_{t\geq 0}$ defined on a group compact Lie or finite group G with invariant measure μ_G, we recall the definition of its mixing time:

$$\tau_{\mathrm{mix}}(\varepsilon) \coloneqq \inf\left\{ t \geq 0 : \ \|\nu_t - \mu_G\|_{\mathrm{TV}} \leq \varepsilon \ \ \forall \nu << \mu_G \right\} \geq \inf\left\{ t \geq 0 : \ \sup_{g\in G}|k_t(g) - 1| \leq 2\varepsilon \ \ \forall g \in G \right\}. \quad (15.16)$$

Private and quantum capacity   Note that we always have $\mathcal{Q}(\Phi) \leq \mathcal{P}(\Phi)$ and, thus, any bound on the private capacity extends to a bound on the quantum capacity. Moreover, we may also consider variations of these capacities in which we also allow for unlimited classical communication between the sender and the receiver of the output of the quantum channel. These are usually called the two-way private and quantum capacities and we will denote them by $\mathcal{P}_{\leftrightarrow}(\Phi)$ and $\mathcal{Q}_{\leftrightarrow}(\Phi)$, respectively. Clearly, we have $\mathcal{P}(\Phi) \leq \mathcal{P}_{\leftrightarrow}(\Phi)$. We refer to e.g. [Christandl and Müller-Hermes, 2017] for a precise definition of these quantities.

In Theorem 13 of [Wilde et al., 2017] the authors show that the *relative entropy of entanglement* of a quantum channel Φ [Vedral and Plenio, 1998, Pirandola et al., 2009], defined as

$$E_R(\Phi) = \sup_{\rho\in\mathcal{D}(\mathcal{H}\otimes\mathcal{H})} \inf_{\sigma\in\mathrm{SEP}(\mathcal{H}:\mathcal{H})} D\left(\Phi\otimes\mathrm{id}\left(\rho\right)\|\sigma\right)$$

is a strong converse bound on the private capacity $\mathcal{P}(\Phi)$ of a quantum channel. Here $\sigma \in \mathrm{SEP}(\mathcal{H} : \mathcal{H})$ means that we are taking the infimum over all separable states.

Theorem 15.2.2. *Let $(\mathcal{P}_t)_{t\geq 0}$ be a (not necessarily ergodic) QMS transferred from a classical semigroup $(P_t)_{t\geq 0}$ of corresponding density kernel $(k_t)_{t\geq 0}$ over a finite or compact Lie group G via the projective representation $U : G \to \mathcal{U}(\mathcal{H})$. Then, for each $t \geq 0$, the quantum and private capacities of $\mathcal{P}_{t*}$ satisfy the following strong converse bound*

$$\mathcal{Q}\left(\mathcal{P}_{t*}\right), \mathcal{P}\left(\mathcal{P}_{t*}\right) \leq \max_k \log d_k + \int_G k_t \log k_t \, d\mu_G \,,$$

where d_k is the dimension of the kth block defined through Equation (15.11). Moreover, for any $t \geq \tau_{\mathrm{mix}}(\varepsilon/2)$:

$$\mathcal{Q}\left(\mathcal{P}_{t*}\right), \mathcal{P}\left(\mathcal{P}_{t*}\right) \leq \max_k \log d_k + \frac{\varepsilon}{\ln 2} \,.$$

Proof. For any state $\rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H})$ it follows from (15.13) and the fact that $\int_G k_t \log k_t \, d\mu_G \leq \frac{\varepsilon}{\ln 2}$ that for any separable state $\sigma = \sum_i p_i \, \sigma_i \otimes \rho_i$, $\sigma_i \in \mathcal{F}$,

$$D((\mathcal{P}_{t*} \otimes \mathrm{id})(\rho)\|\sigma) \leq D((E_{\mathcal{F}} \otimes \mathrm{id})(\rho)\|\sigma) + \frac{\epsilon}{\ln 2} \,.$$

This is because the semigroup $(\mathcal{P}_t \otimes \mathrm{id})_{t\geq 0}$ can be interpreted as being transferred via the representation $G \ni g \mapsto U_g \otimes \mathbb{1}$, so that $\sigma \in \mathcal{F}(\mathcal{L} \otimes \mathrm{id})$. To obtain the statement for the relative entropy of entanglement, we explore the key insight of Lemma 1 in [Gao et al., 2018a], which is that the infimum is approached from separable states of the above form. Therefore, even when $(\mathcal{P}_t)_{t\geq 0}$ is not primitive, we can still

say that for any $t \geq 0$:

$$E_R(\mathcal{P}_{t*}) = \sup_{\rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H})} \inf_{\sigma \in \text{SEP}(\mathcal{H}:\mathcal{H})} D((\mathcal{P}_{t*} \otimes \text{id})(\rho) \| \sigma)$$

$$\leq \sup_{\rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H})} \inf_{\sigma = \sum_i p_i \sigma_i \otimes \rho_i, \, \sigma_i \in \mathcal{F}} D((\mathcal{P}_{t*} \otimes \text{id})(\rho) \| \sigma)$$

$$\leq \sup_{\rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H})} \inf_{\sigma = \sum_i p_i \sigma_i \otimes \rho_i, \, \sigma_i \in \mathcal{F}} D((E_\mathcal{F} \otimes \text{id})(\rho) \| \sigma) + \frac{\epsilon}{\ln 2}$$

$$= \sup_{\rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H})} \inf_{\sigma \in \text{SEP}(\mathcal{H}:\mathcal{H})} D((E_\mathcal{F} \otimes \text{id})(\rho) \| \sigma) + \frac{\epsilon}{\ln 2}.$$

The last line follows from the fact that, for any fixed $\rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H})$,

$$\inf_{\sigma \in \text{SEP}(\mathcal{H}:\mathcal{H})} D((E_\mathcal{F} \otimes \text{id})(\rho) \| \sigma) \geq \inf_{\sigma \in \text{SEP}(\mathcal{H}:\mathcal{H})} D((E_\mathcal{F} \otimes \text{id})(\rho) \| (E_\mathcal{F} \otimes \text{id})(\sigma))$$

$$\geq \inf_{\sigma = \sum_i p_i \sigma_i \otimes \rho_i, \, \sigma_i \in \mathcal{F}} D((E_\mathcal{F} \otimes \text{id})(\rho) \| \sigma)$$

$$\geq \inf_{\sigma \in \text{SEP}(\mathcal{H}:\mathcal{H})} D((E_\mathcal{F} \otimes \text{id})(\rho) \| \sigma).$$

Therefore, we directly find $E_R(\mathcal{P}_{t*}) \leq E_R(E_\mathcal{F}) + \frac{\epsilon}{\ln 2}$. We also know by Example 2 in [Gao et al., 2018a] that $E_R(E_\mathcal{F}) = \max_k \log d_k$ is the logarithm of the maximal matrix block size. We conclude by using Theorem 13 of [Wilde et al., 2017]. $\qquad \square$

Two-way private and quantum capacity We can also derive strong converses on the two-way quantum and private capacities based on the results of [Christandl and Müller-Hermes, 2017]. These will be based on estimates like those of Theorem 9.2.1, as we will show that they can be related to the max-relative entropy of entanglement of the channel. More specifically, in [Christandl and Müller-Hermes, 2017] the authors show that for a quantum channel $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ the quantity

$$E_{\max}(\Phi) = \sup_{\rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H})} \inf_{\sigma \in \text{SEP}(\mathcal{H}:\mathcal{H})} D_{\max}((\Phi \otimes \text{id})(\rho) \| \sigma) \qquad (15.17)$$

is a strong converse upper bound on the two-way private and quantum capacities of Φ.

By a reasoning similar to Theorem 15.2.2 that is based on Lemma 1(iii) of [Gao et al., 2018a], we find the following:

Theorem 15.2.3. *Using the same assumptions as above, for any $t \geq 0$:*

$$\mathcal{Q}_{\leftrightarrow}(\mathcal{P}_{t*}), \mathcal{P}_{\leftrightarrow}(\mathcal{P}_{t*}) \leq \max_k \log d_k + \log \|k_t\|_\infty.$$

These are strong converse bounds.

Proof. In [Christandl and Müller-Hermes, 2017] the authors show that $E_{\max}$ is a strong converse bound on the two-way private and quantum capacities. The claim then easily follows from Theorem 15.2.1 and Example 2 of [Gao et al., 2018a]. $\qquad \square$

Entanglement-assisted classical capacity In the next theorem, we obtain upper bounds on the entanglement-assisted classical capacity of a transferred QMS:.

Theorem 15.2.4. *Let $(\mathcal{P}_t)_{t \geq 0}$ be a (not necessarily ergodic) QMS transferred from a classical semigroup $(P_t)_{t \geq 0}$ over a finite or compact Lie group G via the projective representation $U : G \to \mathcal{U}(\mathcal{H})$. Then,*

for each $t \geq 0$, the entanglement-assisted classical capacity of $\mathcal{P}_{t*}$ satisfies the following strong converse bound

$$\mathcal{C}_{\mathrm{EA}}(\mathcal{P}_{t*}) \leq \log\left(\sum_k d_k^2\right) + \int_G k_t \log k_t \, d\mu_G,$$

where d_k is the dimension of the kth block defined through Equation (15.11). Moreover we have that, for any $t \geq \tau_{\mathrm{mix}}(\varepsilon/2)$,

$$\mathcal{C}_{\mathrm{EA}}(\mathcal{P}_{t*}) \leq \log\left(\sum_k d_k^2\right) + \frac{\varepsilon}{\ln 2}.$$

Proof. The proof follows the lines of Corollary 3 of [Gao et al., 2018a]. Since the entanglement-assisted classical capacity of a channel, whose expression is given in terms of the mutual information by Theorem 15.1.2, satisfies the strong converse property, it suffices to bound the latter in order to conclude. This is simply done as follows: using the alternative expression for the mutual information as given in Equation (15.1), for any $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, $\mathcal{H}_A \simeq \mathcal{H}_B \simeq \mathcal{H}$,

$$
\begin{aligned}
I(A:B)_{(\mathrm{id}_{\mathcal{H}_A} \otimes \mathcal{P}_{t*})(\rho)} &= \inf_{\sigma_B \in \mathcal{D}(\mathcal{H}_B)} D\big((\mathrm{id}_{\mathcal{H}_A} \otimes \mathcal{P}_{t*})(\rho) \| \rho_A \otimes \sigma_B\big) \\
&\leq \inf_{\sigma_B \in \mathcal{D}(\mathcal{H}_B) \cap \mathcal{F}} D\big((\mathrm{id}_{\mathcal{H}_A} \otimes \mathcal{P}_{t*})(\rho) \| \rho_A \otimes \sigma_B\big) \\
&\leq \inf_{\sigma_B \in \mathcal{D}(\mathcal{H}_B) \cap \mathcal{F}} D\big((\mathrm{id}_{\mathcal{H}_A} \otimes E_{\mathcal{F}})(\rho) \| \rho_A \otimes \sigma_B\big) + \int_G k_t \log k_t \, d\mu_G \\
&= \log\left(\sum_k d_k^2\right) + \int_G k_t \log k_t \, d\mu_G,
\end{aligned}
$$

where the second inequality arises as in (15.13) applied to the QMS $(\mathrm{id} \otimes \mathcal{P}_t)_{t \geq 0}$ transferred from $(P_t)_{t \geq 0}$ via the representation $G \ni g \mapsto U_g \otimes \mathbb{1}$, so that $\rho_A \otimes \sigma_B \in \mathcal{F}(\mathrm{id} \otimes \mathcal{L})$. The last line uses Proposition 5 of [Gao et al., 2018a]. The result then follows after taking the supremum over all input states ρ. $\qquad\square$

Classical capacity By the same reasoning, and a use of the identity $\mathcal{C}(E_{\mathcal{F}}) = \log \sum_k d_k$ from Proposition 5 of [Gao et al., 2018a], we find the following similar result for the classical capacity:

Theorem 15.2.5. *Let $(\mathcal{P}_t)_{t \geq 0}$ be a (not necessarily ergodic) QMS transferred from a classical semigroup $(P_t)_{t \geq 0}$ over a finite or compact Lie group G via the projective representation $U : G \to \mathcal{U}(\mathcal{H})$. Then, for each $t \geq 0$, the classical capacity of $\mathcal{P}_{t*}$ satisfies the following strong converse bound*

$$\mathcal{C}(\mathcal{P}_{t*}) \leq \log\left(\sum_k d_k\right) + \int_G k_t \log k_t \, d\mu_G,$$

where d_k is the dimension of the kth block defined through Equation (15.11). Moreover, we have that, for any $t \geq \tau_{\mathrm{mix}}(\varepsilon/2)$,

$$\mathcal{C}(\mathcal{P}_{t*}) \leq \log\left(\sum_k d_k\right) + \frac{\varepsilon}{\ln 2}.$$

15.2.3. Examples

Here, we briefly illustrate the bounds found in the previous section on simple examples:

The depolarizing QMS In Section 9.3.1, we found decoherence times for the depolarizing QMS on $\mathbb{C}^n$ in terms of the mixing times of the uniform random walk on the complete graph of n^2 vertices. Recalling the bound (3.13) and the log-Sobolev and Poincaré constants for the discrete torus given in (3.30), the mixing time for this classical Markov chain is upper bounded as follows:

$$\tau_{\mathrm{mix}}(\varepsilon) \leq n^2 \frac{1 - \ln 2\varepsilon}{n^2 - 1} + \frac{n^2 \ln(n^2 - 1)}{2(n^2 - 2)} \ln\ln n^2 \,.$$

The following estimates are direct consequences of Theorems 15.2.2 to 15.2.5:

Proposition 15.2.6. *The depolarizing QMS* $(\mathcal{P}_t^{\mathrm{depol}})_{t \geq 0}$ *satisfies the following strong converse bounds: for any* $t \geq n^2 \frac{1 - \ln\varepsilon}{n^2 - 1} + \frac{n^2 \ln(n^2 - 1)}{2(n^2 - 2)} \ln\ln n^2$,

$$\mathcal{Q}(\mathcal{P}_{t*}^{\mathrm{depol}}), \mathcal{P}(\mathcal{P}_{t*}^{\mathrm{depol}}), \mathcal{Q}_{\leftrightarrow}(\mathcal{P}_{t*}^{\mathrm{depol}}), \mathcal{P}_{\leftrightarrow}(\mathcal{P}_{t*}^{\mathrm{depol}}), \mathcal{C}_{\mathrm{EA}}(\mathcal{P}_{t*}^{\mathrm{depol}}), \mathcal{C}(\mathcal{P}_{t*}^{\mathrm{depol}}) \leq \frac{\varepsilon}{\ln 2} \,.$$

The dephasing QMS In Section 9.3.2, we found decoherence times for the dephasing QMS on $\mathbb{C}^n$ in terms of the mixing times of classical Markov chains on the discrete and continuous torus. In particular, we obtained stronger bounds from transferring the latter. Recalling the bound (3.13) and the constant for the discrete torus given in (3.30), the mixing time for this classical Markov chain is upper bounded as follows:

$$\tau_{\mathrm{mix}}(\varepsilon) \leq n \frac{1 - \ln 2\varepsilon}{n - 1} + \frac{n \ln(n - 1)}{2(n - 2)} \ln\ln n \,.$$

In order to obtain small time estimates on the various capacities for the dephasing QMS, we also need an estimate on the kernel of a corresponding classical Markov process: The following bound on the heat kernel on the n-dimensional torus is a simple consequence of Equation (9.12) and a bound that one can find in the proof of Theorem 5.3 of [Saloff-Coste, 1994]: for any $0 \leq t \leq 1$,

$$\sup_{g \in G} |k_t^{\mathrm{heat}}(g)| = \|P_t^{\mathrm{heat}} : \mathbb{L}_1(\mu_G) \to \mathbb{L}_\infty(G)\| \leq \|P_{t/2}^{\mathrm{heat}} : \mathbb{L}_1(\mu_G) \to \mathbb{L}_2(\mu_G)\| \|P_{t/2}^{\mathrm{heat}} : \mathbb{L}_2(\mu_G) \to \mathbb{L}_\infty(G)\|$$

$$= \|k_{t/2}^{\mathrm{heat}}\|_2^2 = k_t^{\mathrm{heat}}(e) \leq \left(\frac{5}{t}\right)^{\frac{n}{2}} \,.$$

The following estimates are direct consequences of Theorems 15.2.2 to 15.2.5:

Proposition 15.2.7. *The dephasing QMS* $(\mathcal{P}_t^{\mathrm{deph}})_{t \geq 0}$ *satisfies the following strong converse bounds: for any* $t \geq n \frac{1 - \ln\varepsilon}{n - 1} + \frac{n \ln(n - 1)}{2(n - 2)} \ln\ln n$

$$\mathcal{C}_{\mathrm{EA}}(\mathcal{P}_{t*}^{\mathrm{deph}}), \mathcal{C}(\mathcal{P}_{t*}^{\mathrm{deph}}) \leq \log(n) + \frac{\varepsilon}{\ln 2}$$

$$\mathcal{Q}(\mathcal{P}_{t*}^{\mathrm{deph}}), \mathcal{P}(\mathcal{P}_{t*}^{\mathrm{deph}}), \mathcal{Q}_{\leftrightarrow}(\mathcal{P}_{t*}^{\mathrm{deph}}), \mathcal{P}_{\leftrightarrow}(\mathcal{P}_{t*}^{\mathrm{deph}}) \leq \frac{\varepsilon}{\ln 2} \,.$$

On the other hand, for any $0 \leq t \leq 1$,

$$\mathcal{C}_{\mathrm{EA}}(\mathcal{P}_{t*}^{\mathrm{deph}}), \mathcal{C}(\mathcal{P}_{t*}^{\mathrm{deph}}) \leq \log(n) + \frac{n}{2} \log\left(\frac{5}{t}\right)$$

$$\mathcal{Q}(\mathcal{P}_{t*}^{\mathrm{deph}}), \mathcal{P}(\mathcal{P}_{t*}^{\mathrm{deph}}), \mathcal{Q}_{\leftrightarrow}(\mathcal{P}_{t*}^{\mathrm{deph}}), \mathcal{P}_{\leftrightarrow}(\mathcal{P}_{t*}^{\mathrm{deph}}) \leq \frac{n}{2} \log\left(\frac{5}{t}\right) \,.$$

15.3. Finite n strong converse via reverse hypercontractivity

The strong converse property of the capacity of a classical-quantum (c-q) channel was proved independently in [Ogawa and Nagaoka, 1999, Winter, 1999]. In this section, we use quantum reverse hypercontractivity to obtain a finite blocklength strong converse bound for transmission of information through c-q channels. Given n uses of the c-q channel $\mathcal{W}$, the triple $(|\mathcal{M}|, \mathcal{E}^n, \mathbf{D}^n)$ defines a code which we denote as $\mathcal{C}_n$. We recall that its *maximum probability of error* is given by

$$p_{\max}(\mathcal{C}_n; \mathcal{W}) := \max_{m \in \mathcal{M}} \left[1 - \mathrm{Tr} \left(D^n(\{m\}) \, \mathcal{W}^{\otimes n} \circ \mathcal{E}^n(m) \right) \right].$$

Theorem 15.3.1. *Let $\mathcal{W} : \mathcal{X} \to \mathcal{D}(\mathcal{H}_B)$, $\dim(\mathcal{H}_B) = d$, be a c-q channel with $\mathcal{W}(x) = \rho_x$ being faithful for all $x \in \mathcal{X}$. Then, for any code $\mathcal{C}_n := (|\mathcal{M}|, \mathcal{E}^n, \mathbf{D}^n)$ with $p_{\max}(\mathcal{C}_n; \mathcal{W}) \equiv \varepsilon$ we have*

$$I(X^n : B^n) \geq \log |\mathcal{M}| - 2 \sqrt{\frac{d \cdot n}{\ln 2} \, \log \frac{1}{1 - \varepsilon}} - \log \frac{1}{1 - \varepsilon},$$

where $d = \dim \mathcal{H}_B$ and the mutual information is computed for the states

$$\rho_{X^n B^n} = \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} |x^n(m)\rangle \langle x^n(m)| \otimes \rho_{x^n(m)}.$$

Proof. For every $x^n = (x_1, \ldots, x_n) \in \mathcal{X}^n$ let $\mathcal{P}_{t,x^n} = \mathcal{P}_t^{\mathrm{depol}, \rho_{x_1}} \otimes \cdots \otimes \mathcal{P}_t^{\mathrm{depol}, \rho_{x_n}}$ with

$$\mathcal{P}_t^{\mathrm{depol}, \rho_x} = \mathrm{e}^{-t} X + (1 - \mathrm{e}^{-t}) \, \mathrm{Tr}(\rho_x X) \, \mathbb{1}.$$

Then following similar steps as in the proof of Theorem 13.1.2, using Theorem 10.1.1, Lemma 13.1.1 and the Araki-Lieb-Thirring inequality, for every $D^n(\{m\})$ we have

$$\mathrm{Tr} \left(\rho_{B^n} \mathcal{P}_{t,x^n}(D^n(\{m\})) \right) \geq \left[\mathrm{Tr} \left(\rho_{x^n} D^n(\{m\}) \right) \right]^{1/(1 - e^{-t})} 2^{-D(\rho_{x^n} \| \rho_{B^n})}.$$

Letting $x^n = x^n(m)$, using $\mathrm{Tr} \left(\rho_{x^n(m)} D^n(\{m\}) \right) \geq 1 - \varepsilon$, taking logarithm of both sides and averaging over the choice of $m \in \mathcal{M}$ we obtain

$$\frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \log \mathrm{Tr} \left(\rho_{B^n} \mathcal{P}_{t,x^n(m)}(D^n(\{m\})) \right) \geq -\frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} D(\rho_{x^n(m)} \| \rho_{B^n}) + \frac{1}{1 - \mathrm{e}^{-t}} \log(1 - \varepsilon)$$

$$= -I(X^n : B^n) + \frac{1}{1 - \mathrm{e}^{-t}} \log(1 - \varepsilon)$$

$$\geq -I(X^n : B^n) + \left(1 + \frac{1}{t} \right) \log(1 - \varepsilon).$$

Now define $\Psi_t(X) = \mathrm{e}^{-t} X + (1 - \mathrm{e}^{-t}) \mathrm{Tr}(X) \mathbb{1}$. Following similar steps as in the proof of Theorem 13.1.2, using $\rho_x \leq \mathbb{1}$ it can be shown that $\Psi_t^{\otimes n} - \mathcal{P}_{t,x^n(m)}$ is completely positive. Therefore, $\mathcal{P}_{t,x^n(m)}(D^n(\{m\})) \leq \Psi_t^{\otimes n}(D^n(\{m\}))$ and we have

$$-I(X^n : B^n) + \left(1 + \frac{1}{t} \right) \log(1 - \varepsilon) \leq \frac{1}{|\mathcal{M}|} \sum_m \log \mathrm{Tr} \left(\rho_{B^n} \Psi_t^{\otimes n}(D^n(\{m\})) \right)$$

$$\leq \log \left(\frac{1}{|\mathcal{M}|} \sum_m \mathrm{Tr} \left(\rho_{B^n} \Psi_t^{\otimes n}(D^n(\{m\})) \right) \right)$$

$$= \log \left(\frac{1}{|\mathcal{M}|} \mathrm{Tr} \left(\rho_{B^n} \Psi_t^{\otimes n}(\mathbb{1}_B^{\otimes n}) \right) \right),$$

where the second line follows from the concavity of the logarithm function and in the third line we use the fact that $\{D^n(\{m\}) : m \in \mathcal{M}\}$ is a POVM. On the other hand,

$$\Psi_t^{\otimes n}(\mathbb{1}_B^{\otimes n}) = \left(e^{-t} + (1 - e^{-t})d\right)^n \mathbb{1}_B^{\otimes n} \leq e^{(d-1)nt}\, \mathbb{1}_B^{\otimes n}$$

Therefore,

$$-I(X^n : B^n) + \left(1 + \frac{1}{t}\right)\log(1 - \varepsilon) \leq -\log|\mathcal{M}| + \frac{d \cdot n \cdot t}{\ln 2}.$$

Optimizing over the choice of $t > 0$, the desired result follows. $\qquad\square$

The above theorem together with the additivity of the capacity of c-q channels directly imply that for any code of rate larger than $\mathcal{C}(\mathcal{W})$, the maximum probability of error goes to one, as $n \to \infty$:

Corollary 15.3.2. *Fir $\alpha > \mathcal{C}(\mathcal{W})$ and assume that $\frac{\log|\mathcal{M}|}{n} \geq \alpha > \mathcal{C}(\mathcal{W})$. Then*

$$p_{\max}(\mathcal{C}_n; \mathcal{W}_G) \geq 1 - e^{-nf},$$

where, for $\gamma = \frac{d}{\ln(2)}$,

$$f = \left(\sqrt{\gamma + (\alpha - \mathcal{C}(\mathcal{W}))} - \sqrt{\gamma}\right)^2,$$

which converges to 0 as $\alpha \to \mathcal{C}(\mathcal{W})$.

Proof. Follows directly from solving Theorem 15.3.1 for $\frac{1}{1 - p_{\max}(\mathcal{C}_n; \mathcal{W})}$. $\qquad\square$

Conclusion and open questions

This thesis investigated the convergence towards equilibrium of quantum Markovian evolutions. Functional inequalities represent a powerful method to estimate the mixing time of such processes in the commutative setting. Such methods were extended to the non-commutative framework inherent to quantum theory over the past few decades. However, these techniques only made sense for primitive evolutions, that is those converging to the same unique full-rank invariant state irrespective of the initial state. One of the main contributions of this thesis was to relax the condition of primitivity of the semigroup.

In finite dimensions, the asymptotic behavior of non-primitive processes is well-known as long as one assumes the existence of a full-rank invariant state (condition of faithfulness). In particular, faithful QMS are known to converge to an algebra of effective observables, known to be equal to its so-called decoherence-free subalgebra. Hence, determining the speed of convergence to this algebra becomes a well-posed problem. To answer this question, we introduced the notion of a *decoherence-free logarithmic Sobolev inequality* and the corresponding *decoherence-free hypercontractivity*. The latter corresponds to looking at the contractivity properties of a QMS with respect to weighted amalgamated $\mathbb{L}_p$ norms.

Another main part of the thesis was to explore [Carlen and Maas, 2017]'s recent proposal for extending the notion of displacement convexity to the quantum realm. In particular, we showed that their definition of a quantum Wasserstein distance allows one to recover most of the implications between the various functional inequalities in the non-commutative setting. We also introduced the related notion of a non-commutative transportation cost inequality from which one can derive concentration inequalities for quantum states.

However, many crucial questions still remain to be answered at this point, and we would like to draw the reader's attention to a couple of them that we believe to be the conceptually most interesting and challenging ones.

Displacement convexity for continuous variable quantum systems

The modified Wasserstein distance associated to a Markov chain $(P_t)_{t\geq 0}$ introduced by [Maas, 2011] has the advantage of providing the set of probability measures on a finite sample space with a Riemannian metric. Despite being different from the usual Wasserstein distance of order 2 defined e.g. in terms of the graph distance associated to $(P_t)_{t\geq 0}$, its convergence in the sense of Gromov-Hausdorff to the Wasserstein distance on a smooth Riemannian manifold in the limit of small mesh was shown to hold in some particular cases (see e.g. [Gigli and Maas, 2013]). We believe that finding a version of such a convergence in the quantum case constitutes a fundamental question. However, quantum Markov semigroups are known to lead to classical semigroups of very different nature (diffusions vs. birth and death) depending on the invariant subalgebra that one restricts them to (cf. Theorem 5.5.1). Therefore, it seems that the non-commutative setting could lead to a better understanding of the relation between the discrete vs. continuous frameworks.

On the other hand, [Carlen and Maas, 2017] proposed a possible extension of their quantum Wasserstein distance to the case of the quantum Ornstein Uhlenbeck semigroup on the CCR algebra. We believe that much remains to be done in this direction. The extension of the definition of the Wasserstein distance proposed in Section 12.8 constitutes a first small step towards a fully rigorous analysis of this framework.

Modified logarithmic Sobolev inequality for quantum Gibbs samplers

The examples considered in this thesis are amongst the simplest noise models that one can think of to verify the validity of the concepts introduced. In recent years, much effort has been devoted to derive functional inequalities for quantum Markov semigroups that are physically relevant. We refer to [Kastoryano and Brandão, 2016] and the references therein, where the authors proved the positivity of the Poincaré constant for a certain class of quantum Gibbs samplers under a mixing condition. Showing the positivity of the modified logarithmic Sobolev constant for these semigroups remains a challenging problem to tackle. This is because the traditional classical techniques relying on conditioning of the inequalities to finite regions break down because of the entanglement that persists at the boundary. Recently, new proofs in the classical setting were found that rely on a discrete Bakry-Émery approach [Dai Pra and Posta, 2013, Fathi and Maas, 2016]. We hope that the theory of quantum optimal transport developed in Chapter 12 will allow us to derive similar results in the quantum setting.

Index

Bibliography

[Agredo et al., 2014] Agredo, J., Fagnola, F., and Rebolledo, R. (2014). Decoherence free subspaces of a quantum Markov semigroup. *Journal of Mathematical Physics*, 55(11):112201.

[a.J. Stam, 1959] a.J. Stam (1959). Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Information and Control*, 2:101–112.

[Albeverio and Høegh-Krohn, 1977] Albeverio, S. and Høegh-Krohn, R. (1977). Dirichlet forms and Markov semigroups on C*-algebras. *Communications in Mathematical Physics*, 56(2):173–187.

[Ambrosio and Gigli, 2013] Ambrosio, L. and Gigli, N. (2013). *A User's Guide to Optimal Transport*, pages 1–155. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Ambrosio et al., 2008] Ambrosio, L., Gigli, N., and Savaré, G. (2008). *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media.

[Applebaum, 2004] Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus*. Cambridge University Press, Cambridge.

[Araki, 1976] Araki, H. (1976). Relative entropy of state of von Neumann algebras. *Publ. RIMS Kyoto Univ.*, 9:809–833.

[Araki, 1977] Araki, H. (1977). Relative entropy for states of von Neumann algebras II. *Publ. Res. Inst. Math. Sci.*, 13(1):173–192.

[Araki, 1990] Araki, H. (1990). On an inequality of lieb and thirring. *Letters in Mathematical Physics*, 19(2):167–170.

[Aubrun and Szarek, 2017] Aubrun, G. and Szarek, S. J. (2017). *Alice and Bob Meet Banach: The Interface of Asymptotic Geometric Analysis and Quantum Information Theory*. Number volume 223 in Mathematical surveys and monographs. American Mathematical Society, Providence, Rhode Island. 00010.

[Audenaert et al., 2016] Audenaert, K., Datta, N., and Ozols, M. (2016). Entropy power inequalities for qudits. *Journal of Mathematical Physics*, 57(5):052202.

[Audenaert et al., 2007] Audenaert, K. M. R., Calsamiglia, J., Muñoz-Tapia, R., Bagan, E., Masanes, L., Acin, A., and Verstraete, F. (2007). Discriminating States: The Quantum Chernoff Bound. *Physical Review Letters*, 98(16):160501.

[Audenaert et al., 2012] Audenaert, K. M. R., Mosonyi, M., and Verstraete, F. (2012). Quantum state discrimination bounds for finite sample size. *Journal of Mathematical Physics*, 53(12).

[Audenaert et al., 2008] Audenaert, K. M. R., Nussbaum, M., Szkoła, A., and Verstraete, F. (2008). Asymptotic error rates in quantum hypothesis testing. *Communications in Mathematical Physics*, 279(1):251–283.

Bibliography

[Bakry, 1989] Bakry, D. (1989). Sur l'interpolation complexe des semigroupes de diffusion. In Azéma, J., Yor, M., and Meyer, P. A., editors, *Séminaire de Probabilités XXIII*, pages 1–20, Berlin, Heidelberg. Springer Berlin Heidelberg.

[Bakry, 1994] Bakry, D. (1994). *L'hypercontractivité et son utilisation en théorie des semigroupes*, pages 1–114. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Bakry and Émery, 1984] Bakry, D. and Émery, M. (1984). Hypercontractivité de semi-groupes de diffusion. *Comptes rendus des séances de l'Académie des sciences. Série 1, Mathématique*, 299(15):775–778.

[Bakry and Émery, 1985] Bakry, D. and Émery, M. (1985). Diffusions hypercontractives. In *Séminaire de Probabilités XIX 1983/84*, pages 177–206. Springer.

[Bakry et al., 2014] Bakry, D., Gentil, I., and Ledoux, M. (2014). *Analysis and Geometry of Markov Diffusion Operators*. Springer International Publishing.

[Bakry, 1990] Bakry, Dominique, M. D. (1990). Inégalités de Sobolev et minorations du semi-groupe de la chaleur. *Annales de la Faculté des sciences de Toulouse : Mathématiques*, 11(2):23–66.

[Ball et al., 1994] Ball, K., Carlen, E. A., and Lieb, E. H. (1994). Sharp uniform convexity and smoothness inequalities for trace norms. *Inventiones mathematicae*, 115(1):463–482.

[Bardet, 2017] Bardet, I. (2017). Estimating the decoherence time using non-commutative functional inequalities. *arXiv preprint arXiv:1710.01039*.

[Barron, 1984] Barron, A. R. (1984). Monotonic central limit theorem for densities. *Department of Statistics, Stanford University, California, Tech. Rep*, 50.

[Barron, 1986] Barron, A. R. (1986). Entropy and the Central Limit Theorem. *The Annals of Probability*, 14(1):336–342.

[Barthe, 1997] Barthe, F. (1997). Inégalités de Brascamp-Lieb et convexité. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 324(8):885 – 888.

[Barthe and Kolesnikov, 2008] Barthe, F. and Kolesnikov, A. V. (2008). Mass transport and variants of the logarithmic Sobolev inequality. *Journal of Geometric Analysis*, 18(4):921–979.

[Bäuml et al., 2015] Bäuml, S., Christandl, M., Horodecki, K., and Winter, A. (2015). Limitations on quantum key repeaters. *Nature Communications*, 6(1).

[Beckner, 1975] Beckner, W. (1975). Inequalities in Fourier analysis. *Annals of Mathematics*, 102(1):159–182.

[Beigi, 2013] Beigi, S. (2013). Sandwiched Rényi divergence satisfies data processing inequality. *Journal of Mathematical Physics*, 54(12):122202.

[Beigi et al., 2018] Beigi, S., Datta, N., and Rouzé, C. (2018). Quantum reverse hypercontractivity: its tensorization and application to strong converses. *arXiv preprint arXiv:1804.10100*.

[Beigi and King, 2016] Beigi, S. and King, C. (2016). Hypercontractivity and the logarithmic Sobolev inequality for the completely bounded norm. *Journal of Mathematical Physics*, 57(1):015206.

[Belavkin, 1982] Belavkin, V. P., S. P. (1982). C*-algebraic generalization of relative entropy and entropy. *Annales de l'I.H.P. Physique théorique*, 37(1):51–58.

[Belavkin and Ohya, 2002] Belavkin, V. P. and Ohya, M. (2002). Entanglement, quantum entropy and mutual information. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 458, pages 209–231. The Royal Society.

[Benamou and Brenier, 2000] Benamou, J.-D. and Brenier, Y. (2000). A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393.

[Bennett et al., 2014] Bennett, C. H., Devetak, I., Harrow, A. W., Shor, P. W., and Winter, A. (2014). The quantum Reverse shannon theorem and resource tradeoffs for simulating quantum channels. *IEEE Transactions on Information Theory*, 60(5):2926–2959.

[Bennett et al., 1999] Bennett, C. H., Shor, P. W., Smolin, J. A., and Thapliyal, A. V. (1999). Entanglement-assisted classical capacity of noisy quantum channels. *Physical Review Letters*, 83:3081–3084.

[Bennett et al., 2002] Bennett, C. H., Shor, P. W., Smolin, J. A., and Thapliyal, A. V. (2002). Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Transactions on Information Theory*, 48(10):2637–2655.

[Bergh and Löfström, 2012] Bergh, J. and Löfström, J. (2012). *Interpolation spaces: an introduction*, volume 223. Springer Science & Business Media.

[Berta et al., 2011] Berta, M., Christandl, M., and Renner, R. (2011). The quantum reverse Shannon theorem based on one-shot information theory. *Communications in Mathematical Physics*, 306(3):579.

[Berta et al., 2017] Berta, M., Fawzi, O., and Tomamichel, M. (2017). On variational expressions for quantum relative entropies. *Letters in Mathematical Physics*, 107(12):2239–2265.

[Berta et al., 2018] Berta, M., Scholz, V. B., and Tomamichel, M. (2018). Rényi divergences as weighted non-commutative vector-valued L_p-spaces. *Annales Henri Poincaré*, 19(6):1843–1867.

[Bhatia, 1997] Bhatia, R. (1997). *Matrix analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer-Verlag, New York.

[Bhatia, 2015] Bhatia, R. (2015). *Positive Definite Matrices*. Princeton Series in Applied Mathematics. Princeton University Press.

[Biane, 1997] Biane, P. (1997). Free Hypercontractivity. *Communications in mathematical physics*, 184:457–474.

[Birman and Solomyak, 1967] Birman, M. S. and Solomyak, M. Z. (1967). *Stieltjes Double-Integral Operators*, pages 25–54. Springer US, Boston, MA.

[Birman and Solomyak, 1993] Birman, M. S. and Solomyak, M. Z. (1993). Operator integration, perturbations, and commutators. *Journal of Soviet Mathematics*, 63(2):129–148.

[Blachman, 1965] Blachman, N. M. (1965). The Convolution Inequality for Entropy Powers. *IEEE Transactions on Information Theory*, 11(2):267–271.

[Blanchard et al., 2000] Blanchard, P., Giulini, D., Joos, E., Kiefer, C., and Stamatescu, I.-O. (2000). *Decoherence: Theoretical, Experimental, and Conceptual Problems: Proceedings of a Workshop Held at Bielefeld Germany, 10–14 November 1998*, volume 538. Springer Science & Business Media.

Bibliography

[Blanchard et al., 2007] Blanchard, P., Hellmich, M., Ługiewicz, P., and Olkiewicz, R. (2007). Quantum dynamical semigroups for finite and infinite bose systems. *Journal of Mathematical Physics*, 48(1):012106.

[Blanchard and Olkiewicz, 2003] Blanchard, P. and Olkiewicz, R. (2003). Decoherence induced transition from quantum to classical dynamics. *Reviews in Mathematical Physics*, 15(03):217–243.

[Blower, 2003] Blower, G. (2003). The Gaussian isoperimetric inequality and transportation. *Positivity*, 7(3):203–224.

[Bobkov and Goetze, 1999] Bobkov, S. and Goetze, F. (1999). Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *Journal of Functional Analysis*, 163(1):1 – 28.

[Bobkov and Ledoux, 2000] Bobkov, S. and Ledoux, M. (2000). From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities. *Geometric & Functional Analysis GAFA*, 10(5):1028–1052.

[Bobkov and Tetali, 2003] Bobkov, S. and Tetali, P. (2003). Modified log-Sobolev inequalities, mixing and hypercontractivity. *Proceedings of the thirty-fifth ACM symposium on Theory of computing - STOC '03*, page 287.

[Bobkov et al., 2001] Bobkov, S. G., Gentil, I., and Ledoux, M. (2001). Hypercontractivity of Hamilton–Jacobi equations. *Journal de Mathématiques Pures et Appliquées*, 80(7):669 – 696.

[Bobkov and Tetali, 2006] Bobkov, S. G. and Tetali, P. (2006). Modified logarithmic Sobolev inequalities in discrete settings. *Journal of Theoretical Probability*, 19(2):289–336.

[Bolley and Villani, 2005] Bolley, F. and Villani, C. (2005). Weighted Csiszár-Kullback-Pinsker inequalities and applications to transportation inequalities. *Annales de la Faculté des Sciences de Toulouse. Mathématiques.*, 6, 14(3):331–352.

[Borell, 1982] Borell, C. (1982). Positivity improving operators and hypercontractivity. *Mathematische Zeitschrift*, 180:225–234.

[Brascamp and Lieb, 1976a] Brascamp, H. J. and Lieb, E. H. (1976a). Best constants in Young's inequality, its converse, and its generalization to more than three functions. *Advances in Mathematics*, 20(2):151 – 173.

[Brascamp and Lieb, 1976b] Brascamp, H. J. and Lieb, E. H. (1976b). On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *Journal of Functional Analysis*, 22(4):366 – 389.

[Bratteli and Robinson, 1976] Bratteli, O. and Robinson, D. W. (1976). Unbounded derivations of von neumann algebras. *Annales de l'I.H.P. Physique théorique*, 25(2):139–164.

[Bratteli and Robinson, 1979] Bratteli, O. and Robinson, D. W. (1979). *Operator Algebras and Quantum Statistical Mechanics I.* Springer Berlin Heidelberg, Berlin, Heidelberg.

[Bratteli and Robinson, 1997] Bratteli, O. and Robinson, D. W. (1997). *Operator Algebras and Quantum Statistical Mechanics II.* Springer Berlin Heidelberg, Berlin, Heidelberg.

[Burgarth et al., 2013] Burgarth, D., Chiribella, G., Giovannetti, V., Perinotti, P., and Yuasa, K. (2013). Ergodic and mixing quantum channels in finite dimensions. *New Journal of Physics*, 15(7):073045.

[Cai et al., 2004] Cai, N., Winter, A., and Yeung, R. W. (2004). Quantum privacy and quantum wiretap channels. *Problems of Information Transmission*, 40(4):318–336.

[Capel et al., 2018] Capel, A., Lucia, A., and Pérez-García, D. (2018). Quantum conditional relative entropy and quasi-factorization of the relative entropy. *Journal of Physics A: Mathematical and Theoretical*.

[Caputo et al., 2015] Caputo, P., Menz, G., and Tetali, P. (2015). Approximate tensorization of entropy at high temperature. *Annales de la Faculté des sciences de Toulouse : Mathématiques*, Ser. 6, 24(4):691–716.

[Carbone, 2004] Carbone, R. (2004). Optimal log-Sobolev inequality and hypercontractivity for positive semigroups on $M_2(\mathbb{C})$. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 07(03):317–335.

[Carbone and Fagnola, 2000] Carbone, R. and Fagnola, F. (2000). Exponential L_2-convergence of quantum Markov semigroups on $\mathcal{B}(\mathcal{H})$. *Mathematical Notes*, 68(4):452–463.

[Carbone and Fagnola, 2003] Carbone, R. and Fagnola, F. (2003). The Feller property of a class of quantum Markov semigroups II. In *Quantum Probability And Infinite-Dimensional Analysis*, pages 57–76. World Scientific.

[Carbone and Martinelli, 2015] Carbone, R. and Martinelli, A. (2015). Logarithmic Sobolev inequalities in non-commutative algebras. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 18(02):1550011.

[Carbone and Sasso, 2008] Carbone, R. and Sasso, E. (2008). Hypercontractivity for a quantum Ornstein-Uhlenbeck semigroup. *Probability Theory and Related Fields*, 140(3-4):505–522.

[Carbone et al., 2011] Carbone, R., Sasso, E., and Umanita, V. (2011). Decoherence for positive semigroups on $M_2(\mathbb{C})$. *Journal of Mathematical Physics*, 52(3).

[Carbone et al., 2013] Carbone, R., Sasso, E., and Umanità, V. (2013). Decoherence for Quantum Markov Semi-Group on Matrix Algebras. *Annales Henri Poincaré*, 14(4):681–697.

[Carbone et al., 2014] Carbone, R., Sasso, E., and Umanita, V. (2014). Ergodic quantum Markov semigroups and decoherence. *J. Oper. Theory*, 72(2):293–312.

[Carbone et al., 2015] Carbone, R., Sasso, E., and Umanita, V. (2015). Environment induced decoherence for Markovian evolutions. *Journal of Mathematical Physics*, 56(9).

[Carlen, 1991] Carlen, E. A. (1991). Superadditivity of Fisher's information and logarithmic Sobolev inequalities. *Journal of Functional Analysis*, 101(1):194 – 211.

[Carlen, 2008] Carlen, E. A. (2008). Brascamp-lieb inequalities for non-commutative integration. *Documenta Mathematica*, 13:553–584.

[Carlen and Cordero-Erausquin, 2009] Carlen, E. A. and Cordero-Erausquin, D. (2009). Subadditivity of the entropy and its relation to Brascamp-Lieb type inequalities. *Geometric and Functional Analysis*, 19(2):373–405.

Bibliography

[Carlen et al., 1987] Carlen, E. A., Kusuoka, S., and Stroock, D. W. (1987). Upper bounds for symmetric Markov transition functions. *Annales de l'institut Henri Poincaré*, 23(2):245–287.

[Carlen and Lieb, 1993] Carlen, E. A. and Lieb, E. H. (1993). Optimal hypercontractivity for Fermi fields and related non-commutative integration inequalities. *Communications in Mathematical Physics*, 155(1):27–46.

[Carlen et al., 2016] Carlen, E. A., Lieb, E. H., and Loss, M. (2016). On a quantum entropy power inequality of Audenaert, Datta, and Ozols. *Journal of Mathematical Physics*, 57(6):062203.

[Carlen and Loss, 1993] Carlen, E. A. and Loss, M. (1993). Sharp constant in Nash's inequality. *International Mathematics Research Notices*, 1993(7):213–215.

[Carlen and Maas, 2014] Carlen, E. A. and Maas, J. (2014). An analog of the 2-Wasserstein metric in non-commutative probability under which the Fermionic Fokker-Planck equation is gradient flow for the entropy. *Communications in Mathematical Physics*, 331(3):887–926.

[Carlen and Maas, 2017] Carlen, E. A. and Maas, J. (2017). Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance. *Journal of Functional Analysis*, 273(5):1810 – 1869.

[Carlen, 1987] Carlen, E. A., K. S. S. D. W. (1987). Upper bounds for symmetric Markov transition functions. *Annales de l'I.H.P. Probabilités et statistiques*, 23(S2):245–287.

[Cavalletti and Mondino, 2017] Cavalletti, F. and Mondino, A. (2017). Sharp and rigid isoperimetric inequalities in metric-measure spaces with lower Ricci curvature bounds. *Inventiones mathematicae*, 208(3):803–849.

[Cencov, 2000] Cencov, N. N. (2000). *Statistical decision rules and optimal inference*. Number 53. American Mathematical Soc.

[Cesi, 2001] Cesi, F. (2001). Quasi-factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields. *Probability Theory and Related Fields*, 120(4):569–584.

[Chafai, 2005] Chafai, D. (2005). Inégalités de Poincaré et de Gross pour les mesures de Bernoulli, de Poisson, et de Gauss. working paper or preprint.

[Chebotarev and Fagnola, 1998] Chebotarev, A. and Fagnola, F. (1998). Sufficient conditions for conservativity of minimal quantum dynamical semigroups. *Journal of Functional Analysis*, 153(2):382 – 404.

[Cheeger, 1969] Cheeger, J. (1969). A lower bound for the smallest eigenvalue of the laplacian. In *Proceedings of the Princeton conference in honor of Professor S. Bochner*, pages 195–199.

[Chen and Đoković, 2015] Chen, L. and Đoković, D. Ž. (2015). Boundary of the set of separable states. *Proc. R. Soc. A*, 471(2181):20150102. 00010.

[Chen et al., 2017a] Chen, Y., Georgiou, T. T., Ning, L., and Tannenbaum, A. (2017a). Matricial Wasserstein-1 distance. *IEEE control systems letters*, 1(1):14–19.

[Chen et al., 2017b] Chen, Y., Georgiou, T. T., and Tannenbaum, A. (2017b). Matrix optimal mass transport: a quantum mechanical approach. *IEEE Transactions on Automatic Control*.

[Christandl and Ferrara, 2017] Christandl, M. and Ferrara, R. (2017). Private states, quantum data hiding, and the swapping of perfect secrecy. *Physical Review Letters*, 119(22).

[Christandl and Müller-Hermes, 2017] Christandl, M. and Müller-Hermes, A. (2017). Relative entropy bounds on quantum, private and repeater capacities. *Communications in Mathematical Physics*, 353(2):821–852.

[Christandl et al., 2018] Christandl, M., Müller-Hermes, A., and Wolf, M. M. (2018). When do composed maps become entanglement breaking? *arXiv preprint arXiv:1807.01266*.

[Christensen and Evans, 1979] Christensen, E. and Evans, D. E. (1979). Cohomology of operator algebras and quantum dynamical semigroups. *Journal of the London Mathematical Society*, s2-20(2):358–368.

[Cipriani, 1997] Cipriani, F. (1997). Dirichlet forms and Markovian semigroups on standard forms of von Neumann algebras. *Journal of Functional Analysis*, 147(2):259 – 300.

[Cipriani et al., 2000] Cipriani, F., Fagnola, F., and Lindsay, J. M. (2000). Spectral Analysis and Feller Property for Quantum Ornstein-Uhlenbeck Semigroups. *Communications in Mathematical Physics*, 210(1):85–105.

[Cipriani and Sauvageot, 2003] Cipriani, F. and Sauvageot, J.-L. (2003). Derivations as square roots of Dirichlet forms. *Journal of Functional Analysis*, 201(1):78 – 120.

[Cordero-Erausquin, 1999] Cordero-Erausquin, D. (1999). Inégalité de Prékopa-Leindler sur la sphère. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, 329(9):789–792.

[Cordero-Erausquin, 2002] Cordero-Erausquin, D. (2002). Some applications of mass transport to Gaussian-type inequalities. *Archive for Rational Mechanics and Analysis*, 161(3):257–269.

[Cordero-Erausquin et al., 2001] Cordero-Erausquin, D., McCann, R. J., and Schmuckenschläger, M. (2001). A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Inventiones mathematicae*, 146(2):219–257.

[Costa and Cover, 1984] Costa, M. and Cover, T. (1984). On the similarity of the entropy power inequality and the Brunn-Minkowski inequality. *IEEE Transactions on Information Theory*, 30(6):837–839.

[Costa, 1985] Costa, M. H. M. (1985). A New Entropy Power Inequality. *IEEE Transactions on Information Theory*, 31(6):751–760.

[Courtade, 2016] Courtade, T. A. (2016). Links between the logarithmic Sobolev inequality and the convolution inequalities for entropy and Fisher information. *arXiv preprint arXiv:1608.05431*.

[Courtade, 2017] Courtade, T. A. (2017). Concavity of entropy power: Equivalent formulations and generalizations. In *2017 IEEE International Symposium on Information Theory (ISIT)*, pages 56–60.

[Courtade, 2018] Courtade, T. A. (2018). A strong entropy power inequality. *IEEE Transactions on Information Theory*, 64(4):2173–2192.

[Courtade et al., 2017] Courtade, T. A., Fathi, M., and Pananjady, A. (2017). Wasserstein stability of the entropy power inequality for log-concave random vectors. In *2017 IEEE International Symposium on Information Theory (ISIT)*, pages 659–663.

Bibliography

[Cover and Thomas, 2012] Cover, T. M. and Thomas, J. A. (2012). *Elements of information theory.* John Wiley & Sons.

[Csiszár, 1967] Csiszár, I. (1967). Information-type measures of difference of probability distributions and indirect observation. *studia scientiarum Mathematicarum Hungarica*, 2:229–318.

[Cubitt et al., 2015] Cubitt, T., Kastoryano, M., Montanaro, A., and Temme, K. (2015). Quantum reverse hypercontractivity. *Journal of Mathematical Physics*, 56(10).

[Cubitt et al., 2012] Cubitt, T. S., Eisert, J., and Wolf, M. M. (2012). The complexity of relating quantum channels to master equations. *Communications in Mathematical Physics*, 310(2):383–418.

[Dai Pra et al., 2002] Dai Pra, P., Paganoni, A. M., and Posta, G. (2002). Entropy inequalities for unbounded spin systems. *The Annals of Probability*, 30(4):1959–1976.

[Dai Pra and Posta, 2013] Dai Pra, P. and Posta, G. (2013). Entropy decay for interacting systems via the bochner-bakry-Émery approach. *Electron. J. Probab.*, 18:21 pp.

[Daleckii and Krein, 1951] Daleckii, Y. L. and Krein, S. (1951). Formulas of differentiation according to a parameter of functions of Hermitian operators. In *Dokl. Akad. Nauk SSSR*, volume 76, pages 13–16.

[Daletskii and Krein, 1965] Daletskii, J. L. and Krein, S. (1965). Integration and differentiation of functions of Hermitian operators and applications to the theory of perturbations. *AMS Translations (2)*, 47(1-30):10–1090.

[Daneri and Savaré, 2008] Daneri, S. and Savaré, G. (2008). Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM Journal on Mathematical Analysis*, 40(3):1104–1122.

[Datta, 2009] Datta, N. (2009). Min- and max-relative entropies and a new entanglement monotone. *IEEE Transactions on Information Theory*, 55(6):2816–2826.

[Datta and Leditzky, 2014] Datta, N. and Leditzky, F. (2014). A limit of the quantum Rényi divergence. *Journal of Physics A: Mathematical and Theoretical*, 47(4):045304.

[Datta et al., 2017] Datta, N., Pautrat, Y., and Rouzé, C. (2017). Contractivity properties of a quantum diffusion semigroup. *Journal of Mathematical Physics*, 58(1):012205.

[Davies, 1977] Davies, E. (1977). Quantun Dynamical Semigroup and the Neutron Diffusion Equation. *Reports of mathematical Physics*, 11(2).

[Davies and Simon, 1984] Davies, E. and Simon, B. (1984). Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians. *Journal of Functional Analysis*, 59(2):335 – 395.

[Davies, 1974] Davies, E. B. (1974). Markovian master equations. *Communications in mathematical Physics*, 39(2):91–110.

[Davies, 1976] Davies, E. B. (1976). *Quantum theory of open systems.* Academic Press London.

[Davies, 1980] Davies, E. B. (1980). *One-parameter semigroups.* London ; New York : Academic Press.

[Davies, 1989] Davies, E. B. (1989). *Heat Kernels and Spectral Theory:.* Cambridge Tracts in Mathematics. Cambridge University Press.

[Davies et al., 1992] Davies, E. B., Gross, L., and Simon, B. (1992). Hypercontractivity: a bibliographic review. *Ideas and methods in quantum and statistical physics (Oslo, 1988)*, pages 370–389.

[Davies and Lindsay, 1992] Davies, E. B. and Lindsay, J. M. (1992). Non-commutative symmetric Markov semigroups. *Mathematische Zeitschrift*, 210(1):379–411.

[de Pagter and Sukochev, 2004] de Pagter, B. and Sukochev, F. (2004). Differentiation of operator functions in non-commutative L_p-spaces. *Journal of Functional Analysis*, 212(1):28 – 75.

[De Pagter et al., 2002] De Pagter, B., Witvliet, H., and Sukochev, F. (2002). Double operator integrals. *Journal of Functional Analysis*, 192(1):52–111.

[De Palma and Huber, 2018] De Palma, G. and Huber, S. (2018). The conditional entropy power inequality for quantum additive noise channels. *arXiv preprint arXiv:1803.00470*.

[De Palma et al., 2015a] De Palma, G., Mari, A., Giovannetti, V., and Holevo, A. S. (2015a). Normal form decomposition for Gaussian-to-Gaussian superoperators. *Journal of Mathematical Physics*, 56(5).

[De Palma et al., 2015b] De Palma, G., Mari, A., Lloyd, S., and Giovannetti, V. (2015b). Multimode quantum entropy power inequality. *Physical Review A*, 91(3):032320.

[De Palma and Trevisan, 2018] De Palma, G. and Trevisan, D. (2018). The conditional entropy power inequality for bosonic quantum systems. *Communications in Mathematical Physics*, 360(2):639–662.

[Delgosha and Beigi, 2014] Delgosha, P. and Beigi, S. (2014). Impossibility of local state transformation via hypercontractivity. *Communications in Mathematical Physics*, 332(1):449–476.

[Dembo, 1989] Dembo, A. (1989). Simple proof of the concavity of the entropy power with respect to added Gaussian noise. *IEEE Transactions on Information Theory*, 35(4):887–888.

[Dembo et al., 1991] Dembo, A., Cover, T. M., and Thomas, J. A. (1991). Information Theoretic Inequalities. *IEEE Transactions on Information Theory*, 37(6):1501–1518.

[den Hollander, 2008] den Hollander, F. (2008). *Large Deviations (Fields Institute Monographs)*. American Mathematical Society.

[Deschamps et al., 2016] Deschamps, J., Fagnola, F., Sasso, E., and Umanità, V. (2016). Structure of uniformly continuous quantum Markov semigroups. *Reviews in Mathematical Physics*, 28(01):1650003.

[Devetak, 2005] Devetak, I. (2005). The private classical capacity and quantum capacity of a quantum channel. *IEEE Transactions on Information Theory*, 51(1):44–55.

[Devetak et al., 2006] Devetak, I., Junge, M., King, C., and Ruskai, M. B. (2006). Multiplicativity of completely bounded p-norms implies a new additivity result. *Communications in Mathematical Physics*, 266(1):37–63.

[Dhahri et al., 2010] Dhahri, A., Fagnola, F., and Rebolledo, R. (2010). The decoherence-free subalgebra of a quantum Markov semigroup with unbounded generator. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 13(03):413–433.

[Diaconis, 1988] Diaconis, P. (1988). *Group representations in probability and statistics*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA.

Bibliography

[Diaconis and Saloff-Coste, 1993] Diaconis, P. and Saloff-Coste, L. (1993). Comparison theorems for reversible Markov chains. *The Annals of Applied Probability*, pages 696–730.

[Diaconis and Saloff-Coste, 1996a] Diaconis, P. and Saloff-Coste, L. (1996a). Logarithmic Sobolev inequalities for finite Markov chains. *The Annals of Applied Probability*, 6(3):695–750.

[Diaconis and Saloff-Coste, 1996b] Diaconis, P. and Saloff-Coste, L. (1996b). Nash inequalities for finite Markov chains. *Journal of Theoretical Probability*, 9(2):459–510.

[Diaconis et al., 1993] Diaconis, P., Saloff-Coste, L., et al. (1993). Comparison techniques for random walk on finite groups. *The Annals of Probability*, 21(4):2131–2156.

[Diaconis and Stroock, 1991] Diaconis, P. and Stroock, D. (1991). Geometric bounds for eigenvalues of Markov chains. *The Annals of Applied Probability*, pages 36–61.

[Dixmier, 1953] Dixmier, J. (1953). Formes linéaires sur un anneau d'opérateurs. *Bulletin de la Société Mathématique de France*, 81(9):39.

[Djellout et al., 2004] Djellout, H., Guillin, A., and Wu, L. (2004). Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *The Annals of Probability*, 32(3B):2702–2732.

[Dye, 1952] Dye, H. A. (1952). The Radon-Nikodým theorem for finite rings of operators. *Transactions of the American Mathematical Society*, 72(2):243–243.

[Erbar and Fathi, 2018] Erbar, M. and Fathi, M. (2018). Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature. *Journal of Functional Analysis*, 274(11):3056 – 3089.

[Erbar and Maas, 2012] Erbar, M. and Maas, J. (2012). Ricci curvature of finite Markov chains via convexity of the entropy. *Archive for Rational Mechanics and Analysis*, 206(3):997–1038.

[Evans, 1977] Evans, D. E. (1977). Irreducible quantum dynamical semigroups. *Communications in Mathematical Physics*, 54(3):293–297.

[Evans and Høegh-Krohn, 1978] Evans, D. E. and Høegh-Krohn, R. (1978). Spectral Properties of Positive Maps on C*-Algebras. *Journal of the London Mathematical Society*, s2-17(2):345–355. 00107.

[Fagnola, 1993] Fagnola, F. (1993). Characterization of isometric and unitarily weakly differentiable cocycles in Fock space, *Quantum Probability and Related Topics*.

[Fagnola, 1999] Fagnola, F. (1999). Quantum Markov semigroups and quantum flows. *Proyecciones*, 18(3):1–144.

[Fagnola and Pellicer, 2009] Fagnola, F. and Pellicer, R. (2009). Irreducible and periodic positive maps. *Communications on Stochastic Analysis*, 3(3). 00011.

[Fagnola and Rebolledo, 1998] Fagnola, F. and Rebolledo, R. (1998). The approach to equilibrium of a class of quantum dynamical semigroups. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 01(04):561–572.

[Fagnola and Rebolledo, 2008] Fagnola, F. and Rebolledo, R. (2008). Algebraic conditions for convergence of a quantum Markov semigroup to a steady state. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 11(03):467–474.

[Fagnola et al., 1994] Fagnola, F., Rebolledo, R., and Saavedra, C. (1994). Quantum flows associated to master equations in quantum optics. *Journal of Mathematical Physics*, 35(1):1–12.

[Fagnola and Umanitá, 2007] Fagnola, F. and Umanitá, V. (2007). Generators of detailed balance quantum Markov semigroups. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 10(03):335–363.

[Fathi and Maas, 2016] Fathi, M. and Maas, J. (2016). Entropic ricci curvature bounds for discrete interacting systems. *Ann. Appl. Probab.*, 26(3):1774–1806.

[Feller, 2008] Feller, W. (2008). *An introduction to probability theory and its applications*, volume 2.

[Frank and Lieb, 2013] Frank, R. L. and Lieb, E. H. (2013). Monotonicity of a relative Rényi entropy. *Journal of Mathematical Physics*, 54(12):122201.

[Frieden, 1998] Frieden, B. R. (1998). *Physics from Fisher Information: A Unification*. Cambridge University Press.

[Frigerio, 1978] Frigerio, A. (1978). Stationary states of quantum dynamical semigroups. *Communications in Mathematical Physics*, 63(3):269–276.

[Frigerio and Gorini, 1984] Frigerio, A. and Gorini, V. (1984). Markov dilations and quantum detailed balance. *Communications in Mathematical Physics*, 93(4):517–532.

[Frigerio and Verri, 1982] Frigerio, A. and Verri, M. (1982). Long-time asymptotic properties of dynamical semigroups on W*-algebras. *Mathematische Zeitschrift*, 180(3):275–286.

[Fuchs and van de Graaf, 2006] Fuchs, C. A. and van de Graaf, J. (2006). Cryptographic Distinguishability Measures for Quantum-mechanical States. *IEEE Transactions on Information Theory*, 45(4):1216–1227.

[G. De Palma et al., 2014] G. De Palma, A. Mari, and V. Giovannetti (2014). A generalization of the entropy power inequality to bosonic quantum systems. *Nature Photonics*, 8(12):958–964.

[Gao et al., 2017] Gao, L., Junge, M., and LaRacuente, N. (2017). Strong subadditivity inequality and entropic uncertainty relations. *arXiv preprint arXiv:1710.10038*.

[Gao et al., 2018a] Gao, L., Junge, M., and LaRacuente, N. (2018a). Capacity estimates via comparison with tro channels. *Communications in Mathematical Physics*, 364(1):83–121.

[Gao et al., 2018b] Gao, L., Junge, M., and LaRacuente, N. (2018b). Fisher information and logarithmic Sobolev inequality for matrix valued functions. *arXiv preprint arXiv:1807.08838*.

[Gardner, 2002] Gardner, R. (2002). The Brunn-Minkowski inequality. *Bulletin of the American Mathematical Society*, 39(3):355–405.

[Gentil, 2002] Gentil, I. (2002). Ultracontractive bounds on Hamilton–Jacobi solutions. *Bulletin des Sciences Mathématiques*, 126(6):507 – 524.

[Gigli and Maas, 2013] Gigli, N. and Maas, J. (2013). Gromov-Hausdorff convergence of discrete transportation metrics. *SIAM Journal on Mathematical Analysis*, 45(2):879–899.

[Gohberg and Kreĭn, 1969] Gohberg, I. C. and Kreĭn, M. G. (1969). *Introduction to the theory of linear non-selfadjoint operators*. Translated from the Russian by A. Feinstein. Translations of Mathematical Monographs, Vol. 18. American Mathematical Society, Providence, R.I.

Bibliography

[Goldstein and Lindsay, 1995] Goldstein, S. and Lindsay, J. M. (1995). KMS-symmetric Markov semigroups. *Mathematische Zeitschrift*, 219(1):591–608.

[Goldstein and Lindsay, 1999] Goldstein, S. and Lindsay, J. M. (1999). Markov semigroups KMS-symmetric for a weight. *Mathematische Annalen*, 313(1):39–67.

[Gorini et al., 1976] Gorini, V., Kossakowski, A., and Sudarshan, E. C. G. (1976). Complete positive dynamical semigroups of N-level systems. *Journal of Mathematical Physics*, 17(1976):821.

[Gozlan, 2009] Gozlan, N. (2009). A characterization of dimension free concentration in terms of transportation inequalities. *The Annals of Probability*, 37(6):2480–2498.

[Gromov and Milman, 1983] Gromov, M. and Milman, V. D. (1983). A topological application of the isoperimetric inequality. *American Journal of Mathematics*, 105(4):843–854.

[Gross, 1975a] Gross, L. (1975a). Hypercontractivity and logarithmic Sobolev inequalities for the Clifford-Dirichlet form. *Duke Mathematical Journal*, 42(3):383–396.

[Gross, 1975b] Gross, L. (1975b). Logarithmic Sobolev inequalities. *American Journal of Mathematics*, 97(4):1061–1083.

[Gross, 1993] Gross, L. (1993). *Logarithmic Sobolev inequalities and contractivity properties of semigroups*, pages 54–88. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Guido et al., 1996] Guido, D., Isola, T., and Scarlatti, S. (1996). Non-symmetric Dirichlet forms on semifinite von Neumann algebras. *Journal of Functional Analysis*, 135(1):50 – 75.

[Gupta and Wilde, 2015] Gupta, M. K. and Wilde, M. M. (2015). Multiplicativity of completely bounded p-norms implies a strong converse for entanglement-assisted capacity. *Communications in Mathematical Physics*, 334(2):867–887.

[Gurvits and Barnum, 2002] Gurvits, L. and Barnum, H. (2002). Largest separable balls around the maximally mixed bipartite quantum state. *Physical Review A*, 66(6). 00167.

[Haag and Kastler, 1964] Haag, R. and Kastler, D. (1964). An algebraic approach to quantum field theory. *Journal of Mathematical Physics*, 5(7):848–861.

[Haagerup, 1979] Haagerup, U. (1979). L_p-spaces associated with an arbitrary von Neumann algebra. In *Algebres d'opérateurs et leurs applications en physique mathématique (Proc. Colloq., Marseille, 1977)*, volume 274, pages 175–184.

[Hadwiger and Ohmann, 1956] Hadwiger, H. and Ohmann, D. (1956). Brunn-Minkowskischer satz und isoperimetrie. *Mathematische Zeitschrift*, 66(1):1–8.

[Hanson et al., 2018] Hanson, E. P., Joye, A., Pautrat, Y., and Raquépas, R. (2018). Landauer's Principle for Trajectories of Repeated Interaction Systems. *Annales Henri Poincaré*, 19(7):1939–1991. 00001.

[Hayashi, 2002] Hayashi, M. (2002). Two quantum analogues of Fisher information from a large deviation viewpoint of quantum estimation. *Journal of Physics A: Mathematical and General*, 35:7689–7727.

[Hayashi, 2005] Hayashi, M. (2005). *Asymptotic theory of quantum statistical inference: selected papers*. World Scientific.

[Hayashi, 2007] Hayashi, M. (2007). Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding. *Physical Review A*, 76(6):62301.

[Hayashi, 2016] Hayashi, M. (2016). *Quantum Information Theory*. Springer.

[Hayden et al., 2008] Hayden, P., Shor, P. W., and Winter, A. (2008). Random quantum codes from Gaussian ensembles and an uncertainty relation. *Open Systems & Information Dynamics*, 15(01):71–89.

[Hellmich, 2009] Hellmich, M. (2009). Decoherence in infinite quantum systems.

[Hellmich, 2011] Hellmich, M. (2011). Quantum Dynamical Semigroups and Decoherence. *Advances in Mathematical Physics*, 2011:1–16.

[Helstrom, 1967] Helstrom, C. (1967). Minimum mean-squared error of estimates in quantum statistics. *Physics Letters A*, 25(2):101 – 102.

[Henstock and Macbeath, 1953] Henstock, R. and Macbeath, A. M. (1953). On the measure of sum sets. (i) the theorems of Brunn, Minkowski, and Lusternik. *Proceedings of the London Mathematical Society*, s3-3(1):182–194.

[Hiai and Kosaki, 1999] Hiai, F. and Kosaki, H. (1999). Means for matrices and comparison of their norms. *Indiana University Mathematics Journal*, 48(3):899–936.

[Hiai et al., 2013] Hiai, F., Kosaki, H., Petz, D., and Ruskai, M. B. (2013). Families of completely positive maps associated with monotone metrics. *Linear Algebra and its Applications*, 439(7):1749–1791.

[Hiai and Mosonyi, 2017] Hiai, F. and Mosonyi, M. (2017). Different quantum f-divergences and the reversibility of quantum operations. *Reviews in Mathematical Physics*, 29(07):1750023.

[Hiai et al., 2011] Hiai, F., Mosonyi, M., Petz, D., and Bény, C. (2011). Quantum f-divergences and error correction. *Reviews in Mathematical Physics*, 23(07):691–747.

[Hiai and Petz, 1991] Hiai, F. and Petz, D. (1991). The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in Mathematical Physsics*, 143(1):99–114.

[Hille and Phillips, 1996] Hille, E. and Phillips, R. S. (1996). *Functional analysis and semi-groups*, volume 31. American Mathematical Soc.

[Holevo, 1993a] Holevo, A. (1993a). A note on covariant dynamical semigroups. *Reports on Mathematical Physics*, 32(2):211–216.

[Holevo, 1993b] Holevo, A. (1993b). On conservativity of covariant dynamical semigroups. *Reports on Mathematical Physics*, 33(1-2):95–110.

[Holevo, 1995a] Holevo, A. (1995a). On the Structure of Covariant Dynamical Semigroups. *Journal of Functional Analysis*, 131(2):255–278.

[Holevo, 2011] Holevo, A. (2011). *Probabilistic and Statistical Aspects of Quantum Theory*. Edizioni della Normale.

[Holevo, 1973] Holevo, A. S. (1973). Bounds for the quantity of information transmitted by a quantum communication channel. *Problems of Information Transmission*, 9(3):177–183.

Bibliography

[Holevo, 1995b] Holevo, A. S. (1995b). On translation-covariant quantum Markov equations. *Izvestiya: Mathematics*, 59:427.

[Holevo, 1996] Holevo, A. S. (1996). Covariant quantum Markovian evolutions. *Journal of Mathematical Physics*, 37(4):1812–1832.

[Holevo, 1998] Holevo, A. S. (1998). Covariant quantum dynamical semigroups: unbounded generators. *Irreversibility and Causality Semigroups and Rigged Hilbert Spaces*, 504(2):67.

[Holevo, 2006] Holevo, A. S. (2006). The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44(1):269–273.

[Holevo, 2012] Holevo, A. S. (2012). *Quantum systems, channels, information: a mathematical introduction*, volume 16. Walter de Gruyter.

[Horodecki et al., 2003] Horodecki, M., Shor, P. W., and Ruskai, M. B. (2003). Entanglement breaking channels. *Reviews in Mathematical Physics*, 15(06):629–641.

[Horodecki et al., 2009] Horodecki, R., Horodecki, P., Horodecki, M., and Horodecki, K. (2009). Quantum entanglement. *Reviews of Modern Physics*, 81:865–942.

[Hove, 1957] Hove, L. V. (1957). The approach to equilibrium in quantum statistics: A perturbation treatment to general order. *Physica*, 23(1):441 – 480.

[Huber et al., 2017] Huber, S., König, R., and Vershynina, A. (2017). Geometric inequalities from phase space translations. *Journal of Mathematical Physics*, 58(1):012206.

[Hudson and Parthasarathy, 1984] Hudson, R. L. and Parthasarathy, K. R. (1984). Quantum Ito's formula and stochastic evolutions. *Communications in Mathematical Physics*, 93(3):301–323.

[Hunt, 1956] Hunt, G. A. (1956). Semi-groups of measures on Lie groups. *Transactions of the American Mathematical Society*, 81(2):264–293.

[Itoh and Shishido, 2008] Itoh, M. and Shishido, Y. (2008). Fisher information metric and Poisson kernels. *Differential Geometry and its Applications*, 26(4):347 – 356.

[Jaksic et al., 2012] Jaksic, V., Ogata, Y., Pillet, C.-A., and Seiringer, R. (2012). Quantum hypothesis testing and non-equilibrium statistical mechanics. *Reviews in Mathematical Physics*, 24(06):1230002.

[Jenčová, 2018] Jenčová, A. (2018). Rényi relative entropies and noncommutative L_p-spaces. *Annales Henri Poincaré*.

[Jeong et al., 2018] Jeong, K., Lee, S., and Jeong, H. (2018). Conditional quantum entropy power inequality for d -level quantum systems. *Journal of Physics A: Mathematical and Theoretical*, 51(14):145303.

[Johnson, 2017] Johnson, O. (2017). A discrete log-Sobolev inequality under a Bakry–Émery type condition. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 53(4):1952–1970.

[Johnson and Yu, 2010] Johnson, O. and Yu, Y. (2010). Monotonicity, thinning, and discrete versions of the entropy power inequality. *IEEE Transactions on Information Theory*, 56(11):5387–5395.

[Joos and Zeh, 1985] Joos, E. and Zeh, H. D. (1985). The emergence of classical properties through interaction with the environment. *Zeitschrift für Physik B Condensed Matter*, 59(2):223–243.

[Joos et al., 2003] Joos, E., Zeh, H. D., Kiefer, C., Giulini, D., Kupsch, J., and Stamatescu, I.-O. (2003). *Decoherence and the Appearance of a Classical World in Quantum Theory*. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Jordan et al., 1998] Jordan, R., Kinderlehrer, D., and Otto, F. (1998). The variational formulation of the Fokker-Planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17.

[Jozsa and Schumacher, 1994] Jozsa, R. and Schumacher, B. (1994). A new proof of the quantum noiseless coding theorem. *Journal of Modern Optics*, 41(12):2343–2349.

[Junge et al., 2015] Junge, M., Palazuelos, C., Parcet, J., and Perrin, M. (2015). Hypercontractivity in finite-dimensional matrix algebras. *Journal of Mathematical Physics*, 56(2):023505.

[Junge and Parcet, 2010] Junge, M. and Parcet, J. (2010). *Mixed-norm inequalities and operator space L_p embedding theory*. Number 952-956. American Mathematical Soc.

[Kadison and Ringrose, 1983] Kadison, R. V. and Ringrose, J. R. (1983). *Fundamentals of the theory of operator algebras / Richard V. Kadison, John R. Ringrose*. Academic Press New York.

[Kadison and Ringrose, 2015] Kadison, R. V. and Ringrose, J. R. (2015). *Fundamentals of the Theory of Operator Algebras Vollume II: Advanced Theory*, volume 2. American Mathematical Soc.

[Kallenberg, 2006] Kallenberg, O. (2006). *Foundations of modern probability*. Springer Science & Business Media.

[Kastoryano and Temme, 2016] Kastoryano, M. and Temme, K. (2016). Non-commutative Nash inequalities. *Journal of Mathematical Physics*, 57(1).

[Kastoryano and Brandão, 2016] Kastoryano, M. J. and Brandão, F. G. S. L. (2016). Quantum gibbs samplers: The commuting case. *Communications in Mathematical Physics*, 344(3):915–957.

[Kastoryano and Temme, 2013] Kastoryano, M. J. and Temme, K. (2013). Quantum logarithmic Sobolev inequalities and rapid mixing. *Journal of Mathematical Physics*, 54(5).

[Kempe et al., 2001] Kempe, J., Bacon, D., Lidar, D. A., and Whaley, K. B. (2001). Theory of decoherence-free fault-tolerant universal quantum computation. *Physical Review A*, 63(4):042307.

[Kennedy et al., 2017] Kennedy, M., Manor, N. A., and Paulsen, V. I. (2017). Composition of PPT maps. *arXiv preprint arXiv:1710.08475*.

[Keyl et al., 2016] Keyl, M., Kiukas, J., and Werner, R. F. (2016). Schwartz operators. *Reviews in Mathematical Physics*, 28(03):1630001.

[King, 2003] King, C. (2003). Inequalities for trace norms of 2×2 block matrices. *Communications in Mathematical Physics*, 242(3):531–545.

[King, 2014] King, C. (2014). Hypercontractivity for semigroups of unital qubit channels. *Communications in Mathematical Physics*, 328(1):285–301.

[Knill et al., 2000] Knill, E., Laflamme, R., and Viola, L. (2000). Theory of quantum error correction for general noise. *Physical Review Letters*, 84(11):2525.

[Ko and Park, 2004] Ko, C. K. and Park, Y. M. (2004). Construction of a family of quantum Ornstein–Uhlenbeck semigroups. *Journal of Mathematical Physics*, 45(2):609–627.

Bibliography

[Koenig, 2015] Koenig, R. (2015). The conditional entropy power inequality for Gaussian quantum states. *Journal of Mathematical Physics*, 56(2):022201.

[Koenig and Smith, 2014] Koenig, R. and Smith, G. (2014). The entropy power inequality for quantum systems. *IEEE Transactions on Information Theory*, 60(3):1536–1548.

[Kosaki, 1984] Kosaki, H. (1984). Applications of the complex interpolation method to a von Neumann algebra: Non-commutative L_p-spaces. *Journal of Functional Analysis*, 56(1):29 – 78.

[Kossakowski, 1972] Kossakowski, A. (1972). On quantum statistical mechanics of non-Hamiltonian systems. *Reports on Mathematical Physics*, 3(4):247–274.

[Kossakowski et al., 1977] Kossakowski, A., Frigerio, A., Gorini, V., and Verri, M. (1977). Quantum detailed balance and KMS condition. *Communications in Mathematical Physics*, 57(2):97–110.

[Kraus, 1971] Kraus, K. (1971). General state changes in quantum theory. *Annals of Physics*, 64(2):311–335.

[Kubo, 1979] Kubo, Fumio, A. T. (1979). Means of positive linear operators. *Mathematische Annalen*, 246:205–224.

[Kullback, 1967] Kullback, S. (1967). A lower bound for discrimination information in terms of variation. *IEEE Transactions on Information Theory*, 13(1):126–127.

[Kümmerer and Maassen, 1987] Kümmerer, B. and Maassen, H. (1987). The essentially commutative dilations of dynamical semigroups on M_n. *Communications in Mathematical Physics*, 109(1):1–22.

[Kunze, 1958] Kunze, R. A. (1958). L_p Fourier transforms on locally compact unimodular groups. *Transactions of the American Mathematical Society*, 89(2):519–540.

[Kye, 2018] Kye, S.-H. (2018). Three qubit separable states of length ten with unique decompositions. *ArXiv e-prints, preprint arXiv:1807.06223*.

[Lami and Giovannetti, 2016] Lami, L. and Giovannetti, V. (2016). Entanglement-saving channels. *Journal of Mathematical Physics*, 57(3):032201.

[Lawler and Sokal, 1988] Lawler, G. F. and Sokal, A. D. (1988). Bounds on the L_2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality. *Transactions of the American Mathematical society*, 309(2):557–580.

[Ledoux, 1992] Ledoux, M. (1992). On an integral criterion for hypercontractivity of diffusion semigroups and extremal functions. *Journal of Functional Analysis*, 105(2):444 – 465.

[Leindler, 1972] Leindler, L. (1972). On a certain converse of Hölder's inequality. In *Linear Operators and Approximation/Lineare Operatoren und Approximation*, pages 182–184. Springer.

[Lesniewski and Ruskai, 1999] Lesniewski, A. and Ruskai, M. B. (1999). Monotone Riemannian metrics and relative entropy on non-commutative probability spaces. *Journal of Mathematical Physics*, 40(11):5702–5724.

[Li, 2014] Li, K. (2014). Second-order asymptotics for quantum hypothesis testing. *The Annals of Statistics*, 42(1):171–189.

[Lidar and Whaley, 2003] Lidar, D. and Whaley, K. (2003). Decoherence-free subspaces and subsystems. *Irreversible Quantum Dynamics*, pages 83–120.

[Lidar, 2014] Lidar, D. A. (2014). *Review of Decoherence-Free Subspaces, Noiseless Subsystems, and Dynamical Decoupling*, pages 295–354. Wiley-Blackwell.

[Lidar et al., 2001] Lidar, D. A., Bihary, Z., and Whaley, K. (2001). From completely positive maps to the quantum Markovian semigroup master equation. *Chemical Physics*, 268(1):35 – 53.

[Lidar et al., 1998] Lidar, D. A., Chuang, I. L., and Whaley, K. B. (1998). Decoherence-free subspaces for quantum computation. *Physical Review Letters*, 81(12):2594.

[Lieb, 1978] Lieb, E. H. (1978). Proof of an entropy conjecture of Wehrl. *Communications in Mathematical Physics*, 62(1):35–41.

[Lieb and Ruskai, 1973] Lieb, E. H. and Ruskai, M. B. (1973). Proof of the strong subadditivity of quantum-mechanical entropy. *Journal of Mathematical Physics*, 14(12):1938–1941.

[Lieb and Thirring, 1991] Lieb, E. H. and Thirring, W. E. (1991). Inequalities for the moments of the eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities. In *The Stability of Matter: From Atoms to Stars*, pages 135–169. Springer.

[Lindblad, 1976] Lindblad, G. (1976). On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48(2):119–130.

[Lindsay, 1990] Lindsay, J. (1990). Gaussian hypercontractivity revisited. *Journal of Functional Analysis*, 92(2):313 – 324.

[Liu et al., 2017] Liu, J., van Handel, R., and Verdú, S. (2017). Beyond the blowing-up lemma: Sharp converses via reverse hypercontractivity. In *2017 IEEE International Symposium on Information Theory (ISIT)*, pages 943–947.

[Lloyd, 1997] Lloyd, S. (1997). Capacity of the noisy quantum channel. *Physical Review A*, 55:1613–1622.

[Lo, 1995] Lo, H.-K. (1995). Quantum coding theorem for mixed states. *Optics Communications*, 119(5):552 – 556.

[Lott and Villani, 2007] Lott, J. and Villani, C. (2007). Weak curvature conditions and functional inequalities. *Journal of Functional Analysis*, 245(1):311 – 333.

[Lott and Villani, 2009] Lott, J. and Villani, C. (2009). Ricci curvature for metric-measure spaces via optimal transport. *Annals of Mathematics*, pages 903–991.

[Lugiewicz and Zegarlinski, 2007] Lugiewicz, P. and Zegarlinski, B. (2007). Coercive inequalities for Hörmander type generators in infinite dimensions. *Journal of Functional Analysis*, 247(2):438 – 476.

[Lusternik, 1935] Lusternik, L. (1935). Die Brunn-Minkowskische ungleiching für beliebige messbare mengen. *The Academy of Sciences of the USSR*, (8):55–58.

[Maas, 2011] Maas, J. (2011). Gradient flows of the entropy for finite Markov chains. *Journal of Functional Analysis*, 261(8):2250–2292.

[Majewski and Zegarlinski, 1996] Majewski, A. W. and Zegarlinski, B. (1996). On quantum stochastic dynamics and non-commutative L_p spaces. *Letters in Mathematical Physics*, 36(4):337–349.

[Martinelli, 1999] Martinelli, F. (1999). *Lectures on Glauber Dynamics for Discrete Spin Models*, pages 93–191. Springer Berlin Heidelberg, Berlin, Heidelberg.

Bibliography

[Marton, 1986] Marton, K. (1986). A simple proof of the blowing-up lemma. *IEEE Transactions on Information Theory*, 32(3):445–446.

[Marton, 1996a] Marton, K. (1996a). Bounding d-distance by informational divergence: A method to prove measure concentration. *The Annals of Probability*, 24(2):857–866.

[Marton, 1996b] Marton, K. (1996b). A measure concentration inequality for contracting Markov chains. *Geometric & Functional Analysis GAFA*, 6(3):556–571.

[Marton, 2015] Marton, K. (2015). Logarithmic Sobolev inequalities in discrete product spaces: a proof by a transportation cost distance. *arXiv preprint arXiv:1507.02803*.

[Masahito, 2005] Masahito, H. (2005). *Asymptotic theory of Quantum Statistical Inference*. World scientific.

[Matsumoto, 2010] Matsumoto, K. (2010). Reverse test and characterization of quantum relative entropy. *arXiv preprint arXiv:1010.1030*.

[Matsumoto, 2013] Matsumoto, K. (2013). A new quantum version of f-divergence. *arXiv preprint arXiv:1311.4722*.

[Maurey, 1991] Maurey, B. (1991). Some deviation inequalities. *Geometric & Functional Analysis GAFA*, 1(2):188–197.

[McCann, 1997] McCann, R. J. (1997). A convexity principle for interacting gases. *Advances in Mathematics*, 128(1):153 – 179.

[Meyer, 1993] Meyer, P.-A. (1993). *Quantum probability for probabilists*, volume 1538 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin.

[Milman, 2009a] Milman, E. (2009a). On the role of convexity in isoperimetry, spectral gap and concentration. *Inventiones mathematicae*, 177(1):1–43.

[Milman, 2009b] Milman, E. (2009b). On the role of convexity in isoperimetry,spectral gap and concentration. *Inventiones mathematicae*, 177(1):1–43.

[Monge, 1781] Monge, G. (1781). Mémoire sur la théorie des déblais et des remblais. *Histoire de l'Académie Royale des Sciences de Paris*, pages 666 – 704.

[Montanaro, 2012] Montanaro, A. (2012). Some applications of hypercontractive inequalities in quantum information theory. *Journal of Mathematical Physics*, 53(12):1–18.

[Montanaro and Osborne, 2010] Montanaro, A. and Osborne, T. J. (2010). Quantum boolean functions. *Chicago Journal of Theoretical Computer Science*, 2010(1).

[Moravčíková and Ziman, 2010] Moravčíková, L. and Ziman, M. (2010). Entanglement-annihilating and entanglement-breaking channels. *Journal of Physics A: Mathematical and Theoretical*, 43(27):275306.

[Morozova and Chentsov, 1991] Morozova, E. A. and Chentsov, N. N. (1991). Markov invariant geometry on manifolds of states. *Journal of Soviet Mathematics*, 56(5):2648–2669.

[Mosonyi and Ogawa, 2015] Mosonyi, M. and Ogawa, T. (2015). Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies. *Communications in Mathematical Physics*, 334(3):1617–1648.

[Mossel et al., 2006] Mossel, E., O'Donnell, R., Regev, O., Steif, J. E., and Sudakov, B. (2006). Non-interactive correlation distillation, inhomogeneous Markov chains, and the reverse Bonami-Beckner inequality. *Israel Journal of Mathematics*, 154(1):299–336.

[Mossel et al., 2013] Mossel, E., Oleszkiewicz, K., and Sen, A. (2013). On reverse hypercontractivity. *Geometric and Functional Analysis*, 23(3):1062–1097.

[Müller-Hermes and França, 2018] Müller-Hermes, A. and França, D. S. (2018). Sandwiched Rényi Convergence for Quantum Evolutions. *Quantum*, 2:55.

[Müller-Hermes et al., 2016] Müller-Hermes, A., França, D. S., and Wolf, M. M. (2016). Relative entropy convergence for depolarizing channels. *Journal of Mathematical Physics*, 57(2):022202.

[Müller-Hermes et al., 2016] Müller-Hermes, A., Stilck França, D., and Wolf, M. M. (2016). Entropy production of doubly stochastic quantum channels. *Journal of Mathematical Physics*, 57(2):022203.

[Müller-Lennert et al., 2013] Müller-Lennert, M., Dupuis, F., Szehr, O., Fehr, S., and Tomamichel, M. (2013). On quantum Rényi entropies: A new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203.

[Nagaoka, 2005] Nagaoka, H. (2005). On the parameter estimation problem for quantum statistical models, *Asymptotic Theory of Quantum Statistical Inference*. pages 125–132.

[Nagaoka, 2005] Nagaoka, H. (2005). On the Relation between Kullback Divergence and Fisher Information: From Classical Systems to Quantum Systems, *Asymptotic Theory of Quantum Statistical Inference: Selected Papers*. pages 399–419.

[Nagaoka, 2006] Nagaoka, H. (2006). The Converse Part of The Theorem for Quantum Hoeffding Bound. *eprint arXiv:quant-ph/0611289*.

[Nakajima, 1958] Nakajima, S. (1958). On quantum theory of transport phenomena: steady diffusion. *Progress of Theoretical Physics*, 20(6):948–959.

[Nakamura and Turumaru, 1954] Nakamura, M. and Turumaru, T. (1954). Expectations in an operator algebra. *Tohoku Mathematical Journal*, 6(2-3):182–188.

[Narnhofer and Thirring, 1985] Narnhofer, H. and Thirring, W. E. (1985). From relative entropy to entropy. *Fizika*, 17:257–265. [,149(1985)].

[Nash, 1958] Nash, J. (1958). Continuity of solutions of parabolic and elliptic equations. *American Journal of Mathematics*, 80(4):931–954.

[Nelson, 1966] Nelson, E. (1966). A quartic interaction in two dimensions. *Mathematical theory of elementary particles*, pages 69–73.

[Nelson, 1973a] Nelson, E. (1973a). The free Markoff field. *Journal of Functional Analysis*, 12(2):211 – 227.

[Nelson, 1973b] Nelson, E. (1973b). *Probability theory and Euclidean field theory*, pages 94–124. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Nielsen and Chuang, 2010] Nielsen, M. A. and Chuang, I. L. (2010). *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press.

Bibliography

[Nussbaum and Szkoła, 2009] Nussbaum, M. and Szkoła, A. (2009). The Chernoff lower bound for symmetric quantum hypothesis testing. *The Annals of Statistics*, 37(2):1040–1057.

[Ogata, 2010] Ogata, Y. (2010). Large deviations in quantum spin chains. *Communications in Mathematical Physics*, 296(1):35–68.

[Ogawa and Hayashi, 2004] Ogawa, T. and Hayashi, M. (2004). On error exponents in quantum hypothesis testing. *IEEE Transactions on Information Theory*, 50(6):1368–1372.

[Ogawa and Nagaoka, 1999] Ogawa, T. and Nagaoka, H. (1999). Strong converse to the quantum channel coding theorem. *IEEE Transactions on Information Theory*, 45(7):2486–2489.

[Ogawa and Nagaoka, 2000] Ogawa, T. and Nagaoka, H. (2000). Strong converse and Stein's lemma in quantum hypothesis testing. *IEEE Transactions on Information Theory*, 46(7):2428–2433.

[Ohya and Petz, 1993] Ohya, M. and Petz, D. (1993). *Quantum Entropy and Its Use*. Springer-Verlag.

[Olkiewicz, 1999] Olkiewicz, R. (1999). Environment-induced superselection rules in Markovian regime. *Communications in Mathematical Physics*, 208(1):245–265.

[Olkiewicz, 2000] Olkiewicz, R. (2000). Structure of the algebra of effective observables in quantum mechanics. *Annals of Physics*, 286(1):10 – 22.

[Olkiewicz and Zegarlinski, 1999] Olkiewicz, R. and Zegarlinski, B. (1999). Hypercontractivity in noncommutative Lp spaces. *Journal of Functional Analysis*, 161(1):246 – 285.

[Otto, 2001] Otto, F. (2001). The geometry of dissipative evolution equations: The porous medium equation. *Communications in Partial Differential Equations*, 26(1-2):101–174.

[Otto and Villani, 2000a] Otto, F. and Villani, C. (2000a). Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *Journal of Functional Analysis*, 173(2):361 – 400.

[Otto and Villani, 2000b] Otto, F. and Villani, C. (2000b). Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *Journal of Functional Analysis*, 173(2):361 – 400.

[Otto and Westdickenberg, 2005] Otto, F. and Westdickenberg, M. (2005). Eulerian calculus for the contraction in the Wasserstein distance. *SIAM Journal on Mathematical Analysis*, 37(4):1227–1255.

[Parthasarathy, 1992] Parthasarathy, K. (1992). *An Introduction to Quantum Stochastic Calculus*. Springer Basel.

[Paulsen, 2002] Paulsen, V. (2002). *Completely bounded maps and operator algebras*, volume 78. Cambridge University Press.

[Perez-Garcia et al., 2006] Perez-Garcia, D., Wolf, M. M., Petz, D., and Ruskai, M. B. (2006). Contractivity of positive and trace-preserving maps under L_p norms. *Journal of Mathematical Physics*, 47(8):083506.

[Petz, 1985] Petz, D. (1985). Quasientropies for states of a von Neumann algebra. *Publications of the Research Institute for Mathematical Sciences*, 21(4):787–800.

[Petz, 1986] Petz, D. (1986). Quasi-entropies for finite quantum systems. *Reports on Mathematical Physics*, 23(1):57–65.

[Petz, 1995] Petz, D. (1995). Information-geometry of quantum states, *Quantum Probability Communications.* pages 135–157.

[Petz and Ruskai, 1998] Petz, D. and Ruskai, M. (1998). Contraction of generalized relative entropy under stochastic mappings on matrices. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 1(1):83–89.

[Petz and Sudár, 1996] Petz, D. and Sudár, C. (1996). Geometries of quantum states. *Journal of Mathematical Physics*, 37(6):2662–2673.

[Petz and Sudár, 1999] Petz, D. and Sudár, C. (1999). Extending the Fisher metric to density matrices. *Geometry in Present Day Science*, 21.

[Pillet, 2006] Pillet, C.-A. (2006). Quantum Dynamical Systems, *Open Quantum Systems I.* 1880:107–182.

[Pinsker, 1960] Pinsker, M. S. (1960). Information and information stability of random variables and processes. *Izv. Akad. Nauk.*

[Pirandola et al., 2009] Pirandola, S., García-Patrón, R., Braunstein, S. L., and Lloyd, S. (2009). Direct and Reverse Secret-Key Capacities of a Quantum Channel. *Physical Review Letters*, 102(5).

[Pisier, 1993] Pisier, G. (1993). Non-commutative vector valued L_p -spaces and completely p-summing maps. *Comptes Rendus de l'Académie des Sciences*, 1(247).

[Pisier, 2003] Pisier, G. (2003). *Introduction to Operator Space Theory.* London Mathematical Society Lecture Note Series. Cambridge University Press.

[Pisier and Xu, 2003] Pisier, G. and Xu, Q. (2003). Non-commutative L_p-spaces. *Handbook of the geometry of Banach spaces*, 2:1459–1517.

[Potapov, D. and Sukochev, F., 2010] Potapov, D. and Sukochev, F. (2010). Double operator integrals and submajorization. *Mathematical Modelling of Natural Phenomena*, 5(4):317–339.

[Prékopa, 1973] Prékopa, A. (1973). On logarithmic concave measures and functions. *Acta Scientiarum Mathematicarum*, 34:335–343.

[Prigonine and Resibois, 1961] Prigonine, I. and Resibois, P. (1961). On the kinetics of the approach to equilibrium. *Physica*, 27(7):629–646.

[Ptapov and Sukochev, 2008] Ptapov, D. and Sukochev, F. (2008). Lipschitz and commutator estimates in symmetric operator spaces. *Journal of Operator Theory*, 59(1):211–234.

[Raginsky, 2016] Raginsky, M. (2016). Strong data processing inequalities and ϕ-Sobolev inequalities for discrete channels. *IEEE Transactions on Information Theory*, 62(6):3355–3389.

[Raginsky and Sason, 2014] Raginsky, M. and Sason, I. (2014). Concentration of Measure Inequalities in Information Theory, Communications, and Coding. *Foundations and Trends® in Communications and Information Theory*, 10(1-2):1–247.

[Rahaman et al., 2018] Rahaman, M., Jaques, S., and Paulsen, V. I. (2018). Eventually entanglement breaking maps. *Journal of Mathematical Physics*, 59(6):062201.

Bibliography

[Rajarama Bhat and Parthasarathy, 1995] Rajarama Bhat, B. V. and Parthasarathy, K. R. (1995). Markov dilations of nonconservative dynamical semigroups and a quantum boundary theory. *Annales de l'institut Henri Poincaré Probabilités et statistiques*, 31(4):601–651.

[Ram and Sason, 2016] Ram, E. and Sason, I. (2016). On Rényi entropy power inequalities. *IEEE Transactions on Information Theory*, 62(12):6800–6815.

[Rebolledo, 2005] Rebolledo, R. (2005). Decoherence of quantum Markov semigroups. *Annales de l'institut Henri Poincare (B) Probability and Statistics*, 41(3 SPEC. ISS.):349–373.

[Reed and Simon, 1972] Reed, M. and Simon, B. (1972). *Methods of Modern Mathematical Physics vol. I Functional analysis*. Princeton University. Academic Press.

[Rioul, 2017] Rioul, O. (2017). Yet another proof of the entropy power inequality. *IEEE Transactions on Information Theory*, 63(6):3595–3599.

[Robinson, 1982] Robinson, D. W. (1982). Strongly positive semigroups and faithful invariant states. *Communications in Mathematical Physics*, 85(1):129–142.

[Rothaus, 1981] Rothaus, O. (1981). Logarithmic Sobolev inequalities and the spectrum of Schrödinger operators. *Journal of Functional Analysis*, 42(1):110 – 120.

[Rothaus, 1986] Rothaus, O. (1986). Hypercontractivity and the Bakry-Émery criterion for compact Lie groups. *Journal of Functional Analysis*, 65(3):358 – 367.

[Rothschild and Stein, 1976] Rothschild, L. P. and Stein, E. M. (1976). Hypoelliptic differential operators and nilpotent groups. *Acta Mathematica*, 137:247–320.

[Rouzé and Datta, 2018] Rouzé, C. and Datta, N. (2018). Finite blocklength and moderate deviation analysis of hypothesis testing of correlated quantum states and application to classical-quantum channels with memory. *IEEE Transactions on Information Theory*, 64(1):593–612.

[Ruan, 1988] Ruan, Z. (1988). Subspaces of C*-algebras. *Journal of Functional Analysis*, 76(1):217–230.

[Saloff-Coste, 1994] Saloff-Coste, L. (1994). Precise estimates on the rate at which certain diffusions tend to equilibrium. *Mathematische Zeitschrift*, 217(4):641–677.

[Saloff-Coste, 1997] Saloff-Coste, L. (1997). *Lectures on finite Markov chains*, pages 301–413. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Sason, 2012] Sason, I. (2012). Moderate deviations analysis of binary hypothesis testing. In *2012 IEEE International Symposium on Information Theory Proceedings*, pages 821–825.

[Schumacher, 1995] Schumacher, B. (1995). Quantum coding. *Physical Review A*, 51:2738–2747.

[Schumacher and Westmoreland, 1997] Schumacher, B. and Westmoreland, M. D. (1997). Sending classical information via noisy quantum channels. *Physical Review A*, 56:131–138.

[Segal, 1947] Segal, I. E. (1947). Postulates for general quantum mechanics. *Annals of Mathematics*, pages 930–948.

[Segal, 1953] Segal, I. E. (1953). A Non-Commutative Extension of Abstract Integration. *The Annals of Mathematics*, 57(3):401.

[Shannon, 1948] Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423.

[Simon, 2010] Simon, B. (2010). *Trace ideals and their applications*. Number 120. American Mathematical Society.

[Simon and Høegh-Krohn, 1972] Simon, B. and Høegh-Krohn, R. (1972). Hypercontractive semigroups and two dimensional self-coupled Bose fields. *Journal of Functional Analysis*, 9(2):121 – 180.

[Spohn, 1977] Spohn, H. (1977). An algebraic condition for the approach to equilibrium of an open n-level system. *Letters in Mathematical Physics*, 2(1):33–38.

[Spohn and Lebowitz, 1978] Spohn, H. and Lebowitz, J. L. (1978). Irreversible thermodynamics for quantum systems weakly coupled to thermal reservoirs. x:109–142.

[Stafney, 1969] Stafney, J. D. (1969). The spectrum of an operator on an interpolation space. *Transactions of the American Mathematical Society*, 144:333–349.

[Stein, 1970] Stein, E. M. (1970). *Singular Integrals and Differentiability Properties of Functions (PMS-30)*. Princeton University Press.

[Stein and Weiss, 2016] Stein, E. M. and Weiss, G. (2016). *Introduction to Fourier analysis on Euclidean spaces*, volume 32. Princeton University press.

[Stinespring, 1955] Stinespring, W. F. (1955). Positive functions on C*-algebras. *Proceedings of the American Mathematical Society*, 6(2):211–216.

[Størmer, 1997] Størmer, E. (1997). *Conditional Expectations and Projection Maps of von Neumann Algebras, Operator Algebras and Applications*, pages 449–461. Springer Netherlands, Dordrecht.

[Strocchi, 2008] Strocchi, F. (2008). *An Introduction to the Mathematical Structure of Quantum Mechanics*. World Scientific, 2nd edition.

[Strunz, 2002] Strunz, W. T. (2002). *Decoherence in Quantum Physics, Coherent Evolution in Noisy Environments*, pages 199–233. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Sturm, 2006a] Sturm, K.-T. (2006a). On the geometry of metric measure spaces. *Acta Math.*, 196(1):65–131.

[Sturm, 2006b] Sturm, K.-T. (2006b). On the geometry of metric measure spaces II. *Acta Mathematica*, 196(1):133–177.

[Szehr et al., 2015] Szehr, O., Reeb, D., and Wolf, M. M. (2015). Spectral convergence bounds for classical and quantum Markov processes. *Communications in Mathematical Physics*, 333(2):565–595.

[Takesaki, 1972] Takesaki, M. (1972). Conditional expectations in von Neumann algebras. *Journal of Functional Analysis*, 9(3):306 – 321.

[Takesaki, 1979] Takesaki, M., editor (1979). *Theory of Operator Algebras I*. Springer New York.

[Takesaki, 2003] Takesaki, M. (2003). *Theory of Operator Algebras II*, volume 125 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag.

Bibliography

[Talagrand, 1996] Talagrand, M. (1996). Transportation cost for Gaussian and other product measures. *Geometric & Functional Analysis GAFA*, 6(3):587–600.

[Temme et al., 2010] Temme, K., Kastoryano, M. J., Ruskai, M., Wolf, M. M., and Verstraete, F. (2010). The $\chi 2$-divergence and mixing times of quantum Markov processes. *Journal of Mathematical Physics*, 51(12):122201.

[Temme et al., 2014] Temme, K., Pastawski, F., and Kastoryano, M. J. (2014). Hypercontractivity of quasi-free quantum semigroups. *Journal of Physics A: Mathematical and Theoretical*, 47(40):405303.

[Ticozzi and Viola, 2008] Ticozzi, F. and Viola, L. (2008). Quantum Markovian subsystems: Invariance, attractivity, and control. *IEEE Transactions on Automatic Control*, 53(9):2048–2063.

[Tomamichel et al., 2014] Tomamichel, M., Berta, M., and Hayashi, M. (2014). Relating different quantum generalizations of the conditional Rényi entropy. *Journal of Mathematical Physics*, 55(8):082206.

[Tomamichel and Hayashi, 2012] Tomamichel, M. and Hayashi, M. (2012). A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks. *Journal of Mathematical Physics*, 53(12):122205.

[Toscani, 2012] Toscani, G. (2012). An information-theoretic proof of Nash's inequality. *arXiv preprint arXiv 1206.2130*.

[Umegaki, 1954] Umegaki, H. (1954). Conditional expectation in an operator algebra, I. *Tohoku Mathematical Journal*, 6(2-3):177–181.

[Umegaki, 1956] Umegaki, H. (1956). Conditional expectation in an operator algebra, II. *Tohoku Mathematical Journal*, 8(1):86–100.

[Umegaki, 1959] Umegaki, H. (1959). Conditional expectation in an operator algebra. III. *Kodai Mathematical Seminar Reports*, 11(2):51–64.

[Umegaki, 1962] Umegaki, H. (1962). Conditional expectation in an operator algebra. IV. Entropy and information. *Kodai Mathematical Seminar Reports*, 14(2):59–85.

[van Horssen and Guţă, 2015] van Horssen, M. and Guţă, M. (2015). Sanov and central limit theorems for output statistics of quantum Markov chains. *Journal of Mathematical Physics*, 56(2):022109.

[van Hove, 1954] van Hove, L. (1954). Quantum-mechanical perturbations giving rise to a statistical transport equation. *Physica*, 21(1):517 – 540.

[Varopoulos, 1985] Varopoulos, N. (1985). Hardy-littlewood theory for semigroups. *Journal of Functional Analysis*, 63(2):240 – 260.

[Vedral and Plenio, 1998] Vedral, V. and Plenio, M. B. (1998). Entanglement measures and purification procedures. *Physical Review A*, 57(3):1619–1633.

[Vidal and Tarrach, 1999] Vidal, G. and Tarrach, R. (1999). Robustness of entanglement. *Physical Review A*, 59(1):141.

[Villani, 2000] Villani, C. (2000). A short proof of the "concavity of entropy power". *IEEE Transactions on Information Theory*, 46(4):1695–1696.

[Villani, 2008] Villani, C. (2008). *Optimal transport: old and new*, volume 338. Springer Science & Business Media.

[Viola et al., 2001] Viola, L., Knill, E., and Laflamme, R. (2001). Constructing qubits in physical systems. *Journal of Physics A: Mathematical and General*, 34(35):7067.

[Von Neumann, 2018] Von Neumann, J. (2018). *Mathematical Foundations of Quantum Mechanics: New Edition*. Princeton university press.

[Watrous, 2005] Watrous, J. (2005). Notes on super-operator norms induced by Schatten norms. *Quantum Information and Computation*, 5(1):58–68.

[Weidlich and Haake, 1965] Weidlich, W. and Haake, F. (1965). Master-equation for the statistical operator of solid state laser. *Zeitschrift für Physik*, 186(3):203–221.

[Werner, 1984] Werner, R. F. (1984). Quantum harmonic analysis on phase space. *Journal of Mathematical Physics*, 25(5):1404–1411.

[Werner and Holevo, 2002] Werner, R. F. and Holevo, A. S. (2002). Counterexample to an additivity conjecture for output purity of quantum channels. *Journal of Mathematical Physics*, 43(9):4353–4357.

[Wilde et al., 2017] Wilde, M. M., Tomamichel, M., and Berta, M. (2017). Converse bounds for private communication over quantum channels. *IEEE Transactions on Information Theory*, 63(3):1792–1817.

[Wilde et al., 2014] Wilde, M. M., Winter, A., and Yang, D. (2014). Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. *Communications in Mathematical Physics*, 331(2):593–622.

[Wilming et al., 2017] Wilming, H., Gallego, R., and Eisert, J. (2017). Axiomatic characterization of the quantum relative entropy and free energy. *Entropy*, 19(6):241.

[Winter, 1999] Winter, A. (1999). Coding theorem and strong converse for quantum channels. *IEEE Transactions on Information Theory*, 45(7):2481–2485.

[Wolf, 2012] Wolf, M. M. (2012). Quantum channels & operations: Guided tour. `http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf`. Lecture notes based on a course given at the Niels-Bohr Institute.

[Wolf and Perez-Garcia, 2010] Wolf, M. M. and Perez-Garcia, D. (2010). The inverse eigenvalue problem for quantum channels. *arXiv preprint arXiv:1005.4545*.

[Xu, 2007] Xu, Q. (2007). Operator spaces and non-commutative L_p. *The part on non-commutative Lp-spaces. Lectures in the Summer School on Banach spaces and Operator spaces, Nankai University-China*.

[Zegarlinski, 1992] Zegarlinski, B. (1992). Dobrushin uniqueness theorem and logarithmic Sobolev inequalities. *Journal of Functional Analysis*, 105(1):77 – 111.

[Zegarlinski, 1995] Zegarlinski, B. (1995). Ergodicity of Markov semigroups. *London Mathematical society lecture notes series*, pages 312–337.

[Zeh, 1970] Zeh, H. D. (1970). On the interpretation of measurement in quantum theory. *Foundations of Physics*, 1(1):69–76.

Bibliography

[Zurek, 1981] Zurek, W. H. (1981). Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse? *Physical Review D*, 24:1516–1525.

[Zurek, 1982] Zurek, W. H. (1982). Environment-induced superselection rules. *Physical Review D*, 26:1862–1880.

[Zwanzig, 1960] Zwanzig, R. (1960). Ensemble method in the theory of irreversibility. *The Journal of Chemical Physics*, 33(5):1338–1341.