Geometric inequalities from phase space translations

Stefan Huber [*1,2], Robert König[1,2] and Anna Vershynina[2,3]

[1]*Institute for Advanced Study, Technische Universität München, 85748 Garching, Germany*
[2]*Zentrum Mathematik, Technische Universität München, 85748 Garching, Germany*
[3]*BCAM - Basque Center for Applied Mathematics, 48009 Bilbao, Spain*

February 15, 2017

Abstract

We establish a quantum version of the classical isoperimetric inequality relating the Fisher information and the entropy power of a quantum state. The key tool is a Fisher information inequality for a state which results from a certain convolution operation: the latter maps a classical probability distribution on phase space and a quantum state to a quantum state. We show that this inequality also gives rise to several related inequalities whose counterparts are well-known in the classical setting: in particular, it implies an entropy power inequality for the mentioned convolution operation as well as the isoperimetric inequality, and establishes concavity of the entropy power along trajectories of the quantum heat diffusion semigroup. As an application, we derive a Log-Sobolev inequality for the quantum Ornstein-Uhlenbeck semigroup, and argue that it implies fast convergence towards the fixed point for a large class of initial states.

1 Introduction

The convolution operation $(X, Y) \mapsto X + Y$ between two real- (respectively vector-) valued independent random variables X and Y plays a central role in classical information theory. The operation is defined in terms of the action on probability density functions as

$$(f_X, f_Y) \mapsto f_{X+Y} \qquad \text{where} \qquad f_{X+Y}(z) := \int f_X(z-x) f_Y(x) dx . \qquad (1)$$

The convolution models a general class of additive noise channels, and thus provides a natural framework for the study of information processing and associated capacities. The operation (1) also is a central element in many functional analytic inequalities, most notably Young's inequality [1], the Brascamp-Lieb inequalities [2], de Bruijn's identity [3], the Fisher information inequality [3], and the entropy power inequality [4, 5]. Such inequalities have wide use in information theory, yielding bounds on communication capacities, as observed by Shannon [4]. They can also provide,

---

[*]stefan.huber@ma.tum.de

for example, bounds on the convergence rate in the central limit theorem [6]. Beyond these applications, these results are appealing from a conceptual, geometric viewpoint: many inequalities can be regarded as information-theoretic counterparts of related statements about convex bodies. For example, identifying entropy power with volumes reveals a formal similarity between the Brunn-Minkowski inequality and the entropy power inequality [7]. Indeed, there is even a proof of the latter guided by this intuition [8]. We refer to [9, 10] for detailed accounts of this wealth of inequalities and their interrelationships.

Our work is guided by the question of whether a similar array of inequalities exists in a quantum setting. A key first step in this direction – one which is directly relevant to our work – was taken by Werner [11]. He introduced the convolution operation

$$(f, \rho) \mapsto f \star_t \rho \qquad \text{where} \qquad f \star_t \rho = \int f(\xi) W(\sqrt{t}\xi) \rho W(\sqrt{t}\xi)^\dagger d\xi \,, \tag{2}$$

which involves a probability density function f on phase space as well as a state ρ (of a bosonic system). The operation results in the average state $f \star_t \rho$ when displacing ρ according to f using the (Weyl) displacement operators $W(\xi)$. Here we introduce the parameter $t \geq 0$ for convenience, the case $t = 1$ was considered in [11]. Treating the convolutions (1) and (2) on the same algebraic footing, Werner established (among other results characterizing (2)) a form of Young's inequality. It is worth mentioning that Carlen and Lieb have recently established generalizations of the latter in a fermionic context [12].

More recent work [13] has centered around a convolution operation of the form

$$(\rho_X, \rho_Y) \mapsto \rho_{X \boxplus_\lambda Y} = \mathsf{tr}_2(U_\lambda(\rho_X \otimes \rho_Y)U_\lambda^\dagger) \,, \tag{3}$$

where U_λ is a (d-mode) beamsplitter of transmissivity $\lambda \in [0, 1]$. It is worth pointing out that the action of this operation is formally similar to (1) when expressed in terms of the Wigner functions describing the quantum states ρ_X and ρ_Y. The map (3) describes a process where two states interact. It captures, in particular, the situation where one of the states is transmitted through an additive (bosonic) noise channel. In [13], the authors established an entropy power inequality of the form

$$e^{S(\rho_X \boxplus_\lambda \rho_Y)/d} \geq \lambda e^{S(\rho_X)/d} + (1 - \lambda)e^{S(\rho_Y)/d} \,, \tag{4}$$

for the convolution (3) and for $\lambda = 1/2$, where $S(\rho_X) = -\mathsf{tr}(\rho_X \log \rho_X)$ denotes the von Neumann entropy. Subsequent work [14] managed to lift the restriction on λ, and generalized this result to more general (Gaussian) unitaries in place of U_λ. A related inequality of the form $S(\rho_X \boxplus_\lambda \rho_Y) \geq \lambda S(\rho_X) + (1 - \lambda)S(\rho_Y)$ for $\lambda \in [0, 1]$ was also shown in [13], generalizing classical results [15] (see also [16] for a discussion of the relationship between the two). A generalization to conditional entropies was proposed in [17], and an application to channel capacities was discussed in [18].

A key tool in establishing these results is the quantum Fisher information $J(\rho)$, defined for a state ρ as the divergence-based Fisher information of the family $\{\rho^{(\theta)}\}_\theta$ obtained by displacing ρ along a phase space direction (see Section 3 for a precise definition). It was shown in [13] for $\lambda = 1/2$ and in [14] for general $\lambda \in [0, 1]$ that this quantity satisfies the Fisher information inequality [3]

$$J(\rho_X \boxplus_\lambda \rho_Y)^{-1} \geq \lambda J(\rho_Y)^{-1} + (1 - \lambda)J(\rho_X)^{-1} \,.$$

This identity is a consequence of the strong subadditivity (data processing) inequality for relative entropy, and lies at the heart of the proof of (4).

Following the theme of entropy power inequalities for quantum systems, Audenaert, Datta, and Ozols [19] have obtained strong majorization-type results for the finite-dimensional case. These center around an operation of the form (3), but with ρ_X and ρ_Y being states on a finite-dimensional Hilbert space, and a family $U_\lambda = e^{i\lambda H}$ of unitaries generated by a Hamiltonian H realizing the SWAP-operation of the two systems. As argued by Audenaert et al., these results imply several entropy-power-type inequalities. Very recently, Carlen, Lieb, and Loss [20] have provided an elegant short proof of these statements. Guha, Shapiro, and García-Patrón have discussed alternative definitions of the quantum entropy power and corresponding entropy power inequalities [21].

2   Our contribution

2.1   New geometric inequalities for bosonic systems

Here we focus on the convolution operation (2). We find that this operation satisfies similar properties as the convolutions (1) and (3). In particular, we establish a Stam inequality for the Fisher information of the form

$$J(f \star_t \rho)^{-1} \geq J(\rho)^{-1} + tJ(f)^{-1}, \qquad \text{for all } t \geq 0 . \tag{5}$$

This classical-quantum version of Stam's inequality has several immediate consequences, all of which follow the reasoning used in establishing classical results [9]. For example, we use (5) to establish an entropy power inequality of the form

$$\exp\left(S(f \star_t \rho)/d\right) \geq \exp\left(S(\rho)/d\right) + t\exp\left(H(f)/d\right) . \tag{6}$$

Taking $f = f_Z$ to be a unit-variance centered Gaussian, we find a quantum isoperimetric inequality of the form

$$\frac{d}{dt}\bigg|_{t=0} \left(\frac{1}{2d}J(f_Z \star_t \rho)\right)^{-1} \geq 1 .$$

Note that $f_Z \star_t \rho$ is the result of applying a classical noise channel to ρ, where the variance of the displacement goes to 0 in the limit $t \to 0$. This family of maps constitutes a semigroup, which we call the heat diffusion semigroup: it is generated by a Liouvillian $\mathcal{L}_{\mathsf{heat}}$ such that

$$f_Z \star_t \rho = e^{t\mathcal{L}_{\mathsf{heat}}}(\rho) \qquad \text{for an initial state } \rho \text{ and } t \geq 0 .$$

We find that the entropy power along trajectories generated by this Liouvillian is concave, i.e.,

$$\frac{d^2}{dt^2}\bigg|_{t=0} \exp\left(\frac{1}{d}S(e^{t\mathcal{L}_{\mathsf{heat}}}(\rho))\right) \leq 0 . \tag{7}$$

Eq. (7) generalizes a celebrated result [22, 23] concerning the classical heat equation. The entropy power inequality (6) for Gaussian f implies the lower bound $\exp\left(\frac{1}{d}S(e^{t\mathcal{L}_{\mathsf{heat}}}(\rho))\right) \geq \exp\left(\frac{1}{d}S(\rho)\right) +$

3

$(2\pi e)t$, and establishes the isoperimetric inequality for the Fisher information

$$J(\rho)\exp\left(\frac{1}{d}S(\rho)\right) \geq 4\pi e\, d \tag{8}$$

for states ρ of d bosonic modes. We find that for $d = 1$, this statement is tight: Gaussian thermal states achieve equality in (8) in the limit of large mean photon numbers.

2.2   Application to the Ornstein-Uhlenbeck semigroup

We apply our results, in particular Eq. (8), to the quantum Ornstein-Uhlenbeck (qOU) semigroup for a one-mode bosonic system. This is a one-parameter group of CPTP maps $\{e^{t\mathcal{L}_{\mu,\lambda}}\}_{t\geq 0}$ generated by a linear combination of Liouvillians of a quantum amplifier and an attenuator channel, respectively:

$$\mathcal{L}_{\mu,\lambda} = \mu^2\mathcal{L}_- + \lambda^2\mathcal{L}_+ \qquad \text{for } 0 < \lambda < \mu \;, \quad \text{where} \tag{9}$$

$$\mathcal{L}_+(\rho) = a^\dagger\rho a - \frac{1}{2}\{aa^\dagger,\rho\} \qquad \text{and} \qquad \mathcal{L}_-(\rho) = a\rho a^\dagger - \frac{1}{2}\{a^\dagger a,\rho\} \;,$$

where $a^\dagger$ and a are the creation and annihilation operators (i.e., $[a,a^\dagger] = \mathsf{id}$). The qOU semigroup (9) is a natural counterpart of the classical semigroup generated by the Fokker-Planck equation (see Appendix A). The unique fixed point of the semigroup $\{e^{t\mathcal{L}_{\mu,\lambda}}\}_{t\geq 0}$ is the state

$$\sigma_{\mu,\lambda} = (1-\nu)\sum_{n=0}^{\infty}\nu^n|n\rangle\langle n| \qquad \text{with} \qquad \nu = \lambda^2/\mu^2 \;,$$

i.e., it is diagonal in the number state basis $\{|n\rangle\}_{n\in\mathbb{N}_0}$ with a geometric distribution, hence a Gaussian thermal state.

We conjecture that for an arbitrary initial state ρ, this semigroup converges to the fixed point at an exponential rate given by the exponent $\mu^2 - \lambda^2$, that is,

$$D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \leq e^{-(\mu^2-\lambda^2)t}D(\rho\|\sigma_{\mu,\lambda}) \qquad \text{for all } t \geq 0 \;, \tag{10}$$

where $D(\rho\|\sigma) = \mathsf{tr}(\rho\log\rho - \rho\log\sigma)$ is the relative entropy. In Appendix D, we show that (10) holds for all Gaussian states ρ, and the exponent $\mu^2 - \lambda^2$ is optimal. In other words, our conjecture amounts to the statement that certain Gaussian thermal states converge "most slowly" to the fixed point, and the Log-Sobolev-1-constant, defined as the largest constant $\alpha_1 > 0$ such that $D(e^{t\mathcal{L}}(\rho)\|\sigma) \leq e^{-2\alpha_1 t}D(\rho\|\sigma)$ for any state ρ and all $t \geq 0$, is given by $\alpha_1 = \frac{1}{2}(\mu^2 - \lambda^2)$.

The quantum isoperimetric inequality (8) provides evidence for this conjecture: Taking as a specific example the case where $\lambda = 1$ and $\mu = \sqrt{2}$ (and thus $\mu^2 - \lambda^2 = 1$), we can show (see Example 2 in Section 6) that

$$\frac{d}{dt}\Big|_{t=0}D(e^{t\mathcal{L}_{\sqrt{2},1}}(\rho)\|\sigma_{\sqrt{2},1}) \leq -D(\rho\|\sigma_{\sqrt{2},1}) \qquad \text{for all states } \rho \text{ with } \mathsf{tr}(\rho a^\dagger a) \lesssim 0.67 \;. \tag{11}$$

We also show (see Example 1 in Section 6) that recent work by de Palma, Trevisan, and Giovannetti [24] similarly implies exponential convergence of the form (10) for initial states ρ having

4

large entropy: the isoperimetric inequality of [24] (which we discuss in more detail below) implies that

$$\frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}_{\sqrt{2},1}}(\rho)\|\sigma_{\sqrt{2},1}) \leq -D(\rho\|\sigma_{\sqrt{2},1}) \qquad \text{for all states } \rho \text{ with } S(\rho) \gtrsim 2.4 \ . \qquad (12)$$

While this does not establish the scaling (10) for all states ρ, it illustrates the use of the quantum isoperimetric inequality in a concrete context. We clarify the relationship between our case and the one in the classical setting (see Section 2.3), where the Log-Sobolev-1 constant can be obtained from the isoperimetric inequality for the classical Fisher information, following work by Carlen [25]. In fact, this is the main motivation for our conjecture: Gaussian distributions give the optimal convergence rate in the classical problem.

Whether conjecture (10) holds is a challenging open question. We remark that it would significantly strengthen known results about the qOU semigroup. Specifically, Carbone et al. [26] established that the qOU semigroup is hypercontractive. In [26, Proposition 4.2], the following inequality was shown for the Log-Sobolev-2 constant[1] α_2 of $\mathcal{L}_{\mu,\lambda}$:

$$\alpha_C^{-1} \leq \alpha_2^{-1} \leq \frac{4(5 - \log(1 - \nu))}{\mu^2(1 - \nu)} + (3\log 3)\alpha_C^{-1} \qquad \text{for} \qquad \nu = \lambda^2/\mu^2 \ .$$

In this expression, α_C is the Log-Sobolev-2 constant of the associated classical birth-and-death process (which is unknown), but for which the bounds

$$\frac{\log \nu^{-1}}{5\sqrt{5}\mu^2(1 - \nu)^{3/2}} \leq \alpha_C^{-1} \leq \frac{255}{4} \frac{(1 + \log 2)(1 - \nu) + \log \nu^{-1}}{\mu^2(1 - \nu)^3}$$

were shown (see [26, Proposition 4.1] where α_C is denoted α_0). Following [27, 28, 29], this implies that for all states ρ, we have $D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \leq e^{-2\alpha_2 t}D(\rho\|\sigma)$ (respectively, we have $D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \leq e^{-\alpha_2 t}D(\rho\|\sigma_{\mu,\lambda}))$, if the semigroup can be shown to be strongly (respectively weakly) L_p-regular.

The derivation of our fast convergence results from the isoperimetric inequality (8) follows, to some extent, a well-known line of reasoning considered in the classical context. Indeed, Carlen [25] has shown that the isoperimetric inequality gives rise to a Log-Sobolev inequality, which in turns provides bounds on the convergence of the classical Ornstein-Uhlenbeck (cOU) semigroup to the fixed point (see Appendix A). However, in our case we find that, while a Log-Sobolev inequality again easily follows from (8), the quantities appearing in it are not easily estimated. This concerns, in particular, the entropy rate of the attenuator

$$J_-(\rho) = 2\frac{d}{dt}\Big|_{t=0} S(e^{t\mathcal{L}_-}(\rho)) \ ,$$

given an arbitrary initial state ρ (the factor 2 is chosen for convenience only). This is in sharp contrast to the classical case: here the trivial identity $H(e^t X) = H(X) + t$, for a scalar $t > 0$, satisfied by the differential entropy $H(X)$ of a random variable X, is sufficient for the purpose of establishing fast convergence of the cOU semigroup to the fixed point (see Appendix A).

De Palma et al. [24] showed that the infimum $\inf_{\rho: S(\rho)=S} J_-(\rho)$ over all states ρ with a given entropy is achieved by a Gaussian thermal state. This statement, combined with a lower bound

---

[1]Here we follow the notation of [27, 28, 29]. In contrast, this is denoted α_1 in [26].

on the corresponding quantity $J_+(\rho)$ for the amplifier from [30], valid for all states ρ, immediately yields inequality (12).

To establish (11), we prove another lower bound on $J_-(\rho)$: more precisely, we show that the infimum

$$\inf_{\rho:\, \mathsf{tr}(\rho a^\dagger a) \leq \mathbf{n}} J_-(\rho)$$

over all states ρ with mean photon number bounded by $\mathbf{n}$ is achieved by a Gaussian thermal state. The proof proceeds by reduction to the classical case using recent majorization-type results [31, 32]. The latter can be treated using the results from [24]. This, then, provides the required lower bound and yields statement (11). It is not, however, tight enough to establish Conjecture (10).

Understanding the relationship between entropy production rates along trajectories of the qOU semigroup, i.e., different quantities of the form $J_{\mu,\lambda}(\rho) = 2\frac{d}{dt}\Big|_{t=0} S(e^{t\mathcal{L}_{\mu,\lambda}}(\rho))$ for $(\mu,\lambda) \neq (1,0)$, remains an open problem. We believe that progress in this direction could help provide further evidence for (or indeed lead to a proof of) the validity of conjecture (10).

2.3 Prior work in the classical setting

Our work in the quantum setting follows a long sequence of well-known existing arguments applicable to classical probability distributions. All geometric inequalities established here have classical counterparts, and their proofs are inspired by (and directly generalize) corresponding classical proofs. This raises the question of whether other analogues of classical results exist: for example, one may conjecture that there is a quantum counterpart of Young's inequality for the convolution operation (3).

It is not our intention to provide a complete review of this assortment of classical results: it is hardly possible to do justice to the many important developments in this area. We refer to the article [9] for a survey of many known connections. Instead, we briefly review some of the basic definitions and seminal results which are directly relevant to our work.

In the classical setting we assume to have $\mathbb{R}^d$-valued random variables with absolutely continuous density functions. For such a random variable X with density function f (which we often assume to be non-vanishing everywhere for simplicity), a fundamental information measure of interest is the *Shannon (differential) entropy*

$$H(f) = -\int_{\mathbb{R}^d} f(\boldsymbol{x}) \log f(\boldsymbol{x}) d\boldsymbol{x} \ ,$$

also denoted as $H(X)$. The *entropy power* is given by

$$N(f) = \exp\left(2H(f)/d\right) \ ,$$

and also written as $N(X)$. Up to a factor, this quantity coincides with the variance of a Gaussian distribution having the same entropy. The *divergence*, or *relative entropy*, of two density functions f, g is defined as

$$D\left(f \| g\right) = \int_{\mathbb{R}^d} f(\boldsymbol{x}) \log \frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} \, d\boldsymbol{x} \ . \tag{13}$$

The *Fisher information* of a random variable X with density function f is defined as the following quantity[2]

$$J(f) = \int (\nabla f(\boldsymbol{x}))^T \cdot (\nabla f(\boldsymbol{x})) \cdot \frac{1}{f(\boldsymbol{x})} d\boldsymbol{x} , \tag{15}$$

which is associated with the family of translated probability density functions

$$f^{(\boldsymbol{\theta})}(\boldsymbol{x}) = f(\boldsymbol{x} - \boldsymbol{\theta}) \qquad \text{for } \boldsymbol{\theta} \in \mathbb{R}^d . \tag{16}$$

The quantity $J(f)$ is also often denoted as $J(X)$, to emphasize that it is the Fisher information of a random variable X which has density function f.

For two densities f and g describing random variables X and Y, the convolution operation of interest is given by (1), and describes the addition of the two random variables. Of particular interest is the case where one of the random variables is a centered normal distribution with unit variance. Such a random variable is denoted as Z below. Key results in this setting are

the classical de Bruijn identity:

$$\frac{\partial}{\partial t} H \left(X + \sqrt{t} Z \right) = \frac{1}{2} J \left(X + \sqrt{t} Z \right) . \tag{17}$$

This result was established by de Bruijn and gives an important relation between the Fisher information and the entropy when a random variable X is perturbed under an additive Gaussian noise channel. It is a key ingredient in proofs of many information-theoretic inequalities. A simple proof can be found in [7].

the Fisher information inequality:

$$J \left(\sqrt{\lambda} X + \sqrt{1 - \lambda} Y \right) \le \lambda J(X) + (1 - \lambda) J(Y), \text{ for } \lambda \in [0, 1] , \tag{18}$$

as well as the related inequality

$$J \left(X + Y \right)^{-1} - J(X)^{-1} - J(Y)^{-1} \ge 0 .$$

Proofs of these inequalities are given in [5], [7], and [33]. Zamir [33] gives a particularly useful proof which relies on the information-processing inequality. This inequality states that the application of a channel cannot increase the Fisher information. Zamir's proof of the Fisher information inequality can be generalized to the quantum case.

---

[2]In more general terms, (15) is the trace of the Fisher information matrix

$$\left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \Big|_{\boldsymbol{\theta} = \boldsymbol{\theta_0}} D \left(f^{(\boldsymbol{\theta_0})} \Big\| f^{(\boldsymbol{\theta})} \right) \right)^m_{i,j=1} \tag{14}$$

corresponding to (16). The Fisher information matrix provides a lower bound on the variance of an unbiased estimate of the unknown parameter $\boldsymbol{\theta}$ according to the Cramer-Rao inequality. For our purposes, it is sufficient to consider the family (16).

the Fisher information isoperimetric inequality

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0}\left[\frac{1}{d}J(X+\sqrt{\epsilon}Z)\right]^{-1}\geq 1 \ .$$

This inequality implies that for Gaussian states, the inverse of the Fisher information has minimal sensitivity to additive Gaussian noise [34].

the concavity of the entropy power

$$\frac{d^2}{d\epsilon^2}\bigg|_{\epsilon=0}N(X+\sqrt{\epsilon}Z)\leq 0 \ .$$

This celebrated result establishes that the entropy power is a concave function along trajectories of the heat flow semigroup. A proof is given in [22], and some shorter ones are presented in [34, 23].

the entropy power inequality

$$N(X+Y)\geq N(X)+N(Y) \ .$$

Stam [3] gave a proof of the entropy power inequality which relies on the de Bruijn identity and the Fisher information inequality. The proof was later simplified by Blachman [5] and others [6, 9].

the isoperimetric inequality for entropies

$$\frac{1}{d}J(X)N(X)\geq 2\pi e \ ,$$

as given in [7]. The isoperimetric inequality for entropies implies that Gaussians have minimal entropy power among random variables with fixed Fisher information, and can be used to derive Log-Sobolev inequalities for the Ornstein-Uhlenbeck semigroup [25], as we review in Appendix A.

3 Preliminaries

3.1 States and information measures of interest

We consider a d-mode bosonic system with "position" and "momentum" operators (Q_k, P_k) of the k-th mode satisfying the canonical commutation relations $[Q_j, P_k] = i\delta_{j,k}I$. Denoting the vector of position- and momentum-operators by $\boldsymbol{R} = (Q_1, P_1, \ldots, Q_d, P_d)$, the Weyl displacement operators are defined as

$$W(\boldsymbol{\xi}) = e^{i\sqrt{2\pi}\boldsymbol{\xi}\cdot(\sigma\boldsymbol{R})} \qquad \text{for } \boldsymbol{\xi}\in\mathbb{R}^{2d} \ . \tag{19}$$

Here $\sigma = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{\oplus d}$ is the matrix defining the symplectic inner product. The factor $\sqrt{2\pi}$ in the definition (19) is for convenience only. From the commutation relations of position and momentum

operators and the Campbell-Baker-Hausdorff formula, it is straightforward to check that the Weyl operators satisfy

$$W(\boldsymbol{\xi})W(\boldsymbol{\eta}) = e^{-i\pi\boldsymbol{\xi}\cdot(\sigma\boldsymbol{\eta})}W(\boldsymbol{\xi}+\boldsymbol{\eta}) \qquad \text{for } \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{R}^{2d} . \tag{20}$$

Consider a state ρ on d modes. Quantities of interest are the von Neumann entropy $S(\rho) = -\mathsf{tr}(\rho\log\rho)$, as well as the relative entropy $D(\rho\|\sigma) = \mathsf{tr}\left(\rho\log\rho - \rho\log\sigma\right)$. The latter expression is defined for positive operators ρ, σ, and we will assume without further comments that the states ρ, σ have full rank.

For a multi-parameter family $\{\rho^{(\boldsymbol{\theta})}\}_{\boldsymbol{\theta}\in\mathbb{R}^D}$ of states depending smoothly on the parameters $\boldsymbol{\theta}$, the divergence-based quantum Fisher information is defined as the trace of the Fisher information matrix

$$J\left(\{\rho^{(\boldsymbol{\theta})}\};\boldsymbol{\theta}\right)\big|_{\boldsymbol{\theta}=\boldsymbol{\theta_0}} = \left(\frac{\partial^2}{\partial\theta_j\partial\theta_k}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta_0}} D\left(\rho^{(\boldsymbol{\theta_0})}\big\|\,\rho^{(\boldsymbol{\theta})}\right)\right)^D_{j,k=1} .$$

This definition quantifies the dependence of the states on the parameter $\boldsymbol{\theta}$ in the neighborhood of $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

In the following, we will apply this definition to the family $\{\rho^{(\boldsymbol{\theta})}\}_{\boldsymbol{\theta}\in\mathbb{R}^{2d}}$ of states obtained by translating a given d-mode state ρ: Analogously to (16), where we translated a given probability density function f, we define the translated states

$$\rho^{(\boldsymbol{\theta})} := W(\boldsymbol{\theta})\rho W(\boldsymbol{\theta})^\dagger \qquad \text{for } \boldsymbol{\theta} \in \mathbb{R}^{2d} . \tag{21}$$

Here translation by the parameter $\boldsymbol{\theta}$ on phase space is achieved by means of the Weyl operators. The corresponding quantity

$$J(\rho) = \mathsf{tr}\left(J\left(\{\rho^{(\boldsymbol{\theta})}\};\boldsymbol{\theta}\right)\big|_{\boldsymbol{\theta}=0}\right) , \tag{22}$$

will simply be called the Fisher information of ρ. Note that this definition matches that of the classical Fisher information (cf. (14)). We emphasize that the concept of quantum Fisher information is non-unique (see [35]), but the use of (22) is sufficient for our purposes.

3.2   The quantum diffusion semigroup and the de Brujin identity

Consider the Liouvillian defined on d modes as

$$\mathcal{L}_{\mathsf{heat}}(\rho) = -\pi \sum_{j=1}^{2d} [R_j, [R_j, \rho]] . \tag{23}$$

(The factor π differs from the convention used in [13], but turns out to be convenient in the proof of Theorem 3, as explained later.) The one-parameter semigroup $\{e^{t\mathcal{L}_{\mathsf{heat}}}\}_{t\geq 0}$ of completely positive trace-preserving maps (CPTPMs) generated by $\mathcal{L}_{\mathsf{heat}}$ will be called the quantum (heat) diffusion semigroup. It has various nice properties: for example, as shown in [13], a quantum version of the de Bruijn identity (17) reads

$$\frac{d}{dt}\bigg|_{t=0} S\left(e^{t\mathcal{L}_{\mathsf{heat}}}(\rho)\right) = \frac{1}{2}J(\rho) . \tag{24}$$

9

We remark that the proof of (24), which has subsequently been applied, e.g., in [21] and generalized in [14], involves certain formal manipulations whose rigorous justification remains an interesting mathematical problem: as a quantum counterpart of partial integration, for example, arguments under the trace need to be cyclically permuted. It is clear that such manipulations should be valid for sufficiently regular families of states (and indeed, are established for Gaussian states), but corresponding conditions are currently unknown. We believe that the recent introduction of Schwartz operators in [36] provides an appropriate framework to shed light on this aspect. In the following, we will assume that our states under consideration satisfy the required regularity assumptions. We hope that this issue will eventually be resolved in a similar manner as in the classical setting, where initial work by Shannon [4] was followed by a long sequence of papers with increasing rigor. In the case of the classical de Bruijn identity (17), Barron [6], based on Stam's work [3], has shown validity for all random variables X with finite variance.

The map $e^{t\mathcal{L}_{\text{heat}}}$ can be explicitly written as (see [13, 37])

$$e^{t\mathcal{L}_{\text{heat}}}(\rho) = \frac{1}{(2\pi)^d} \int e^{-\|\boldsymbol{\xi}\|^2/2} W(\sqrt{t}\boldsymbol{\xi}) \, \rho \, W(\sqrt{t}\boldsymbol{\xi})^\dagger d\boldsymbol{\xi} \ . \tag{25}$$

We may interpret this as the result of applying a certain convolution operation to a Gaussian distribution and a quantum state. More precisely, for $t \geq 0$, we define a convolution operation $\star_t$ between a probability density function $f : \mathbb{R}^{2d} \to \mathbb{R}$ and a d-mode state ρ by

$$(f, \rho) \mapsto f \star_t \rho := \int f(\boldsymbol{\xi}) W(\sqrt{t}\boldsymbol{\xi}) \, \rho \, W(\sqrt{t}\boldsymbol{\xi})^\dagger d\boldsymbol{\xi} \ . \tag{26}$$

In this terminology, Eq. (25) becomes

$$f_Z \star_t \rho = e^{t\mathcal{L}_{\text{heat}}}(\rho) \ ,$$

where f_Z is a unit-variance centered Gaussian distribution, that is,

$$f_Z(\boldsymbol{\xi}) = (2\pi)^{-d} e^{-\|\boldsymbol{\xi}\|^2/2} \ . \tag{27}$$

To close this section, we list two elementary properties of the convolution (26) which can be checked by straightforward calculation. If $f = f_X$ is the probability density function of a random variable X and $t \geq 0$, then

$$f_X \star_t \rho = f_{\sqrt{t}X} \star_1 \rho \ , \tag{28}$$

where the probability density function of the rescaled random variable $\sqrt{t}X$ is given by $f_{\sqrt{t}X}(\boldsymbol{\xi}) = f(\boldsymbol{\xi}/\sqrt{t})/\sqrt{t}^{2d}$. Addition of random variables corresponds to convolution in the following sense:

$$f_{X_1} \star_1 (f_{X_2} \star_1 \rho) = f_{X_1+X_2} \star_1 \rho \ , \tag{29}$$

where $f_{X_1+X_2}$ is defined as in (1).

4 Quantum geometric inequalities

In this section, we present several statements about the convolution operation (26) and the quantum Fisher information (22). The key idea in establishing these results is the fact that the convolution operation (26) constitutes data processing, and hence provides an inequality because of the monotonicity of relative entropy. This is expressed in the following lemma. In the classical setting, the analogous argument for obtaining the Fisher information inequality (18) was first emphasized by Zamir [33].

Lemma 1. *(Data processing inequality for convolution) Let $f, g : \mathbb{R}^{2d} \to \mathbb{R}$ be probability density functions with full support. Then*

$$D\left(f \star_t \rho \,\middle\|\, g \star_t \sigma\right) \leq D\left(f \,\middle\|\, g\right) + D\left(\rho \,\middle\|\, \sigma\right) \tag{30}$$

for any states ρ, σ.

Proof. The quantum relative entropy satisfies the following scaling property for scalars $\lambda, \mu > 0$:

$$D\left(\lambda \rho \,\middle\|\, \mu \sigma\right) = \lambda D\left(\rho \,\middle\|\, \sigma\right) - \lambda \mathsf{tr}\left(\rho\right) \log \frac{\mu}{\lambda} \ . \tag{31}$$

Defining $h(\boldsymbol{\xi}) = \frac{g(\boldsymbol{\xi})}{f(\boldsymbol{\xi})}$, we obtain (using the translated states defined in Eq. (21))

$$
\begin{aligned}
D\left(f \star_t \rho \,\middle\|\, g \star_t \sigma\right) &= D\left(f \star_t \rho \,\middle\|\, (f \cdot h) \star_t \sigma\right) \\
&= D\left(\int f(\boldsymbol{\xi}) \rho^{(\sqrt{t}\boldsymbol{\xi})} \, \mathrm{d}\boldsymbol{\xi} \,\middle\|\, \int f(\boldsymbol{\xi}) h(\boldsymbol{\xi}) \sigma^{(\sqrt{t}\boldsymbol{\xi})} \, \mathrm{d}\boldsymbol{\xi}\right) \\
&\leq \int f(\boldsymbol{\xi}) D\left(\rho^{(\sqrt{t}\boldsymbol{\xi})} \,\middle\|\, h(\boldsymbol{\xi}) \sigma^{(\sqrt{t}\boldsymbol{\xi})}\right) \mathrm{d}\boldsymbol{\xi} \\
&= \int f(\boldsymbol{\xi}) \left(D\left(\rho^{(\sqrt{t}\boldsymbol{\xi})} \,\middle\|\, \sigma^{(\sqrt{t}\boldsymbol{\xi})}\right) - \mathsf{tr}\left(\rho^{(\sqrt{t}\boldsymbol{\xi})}\right) \log h(\boldsymbol{\xi})\right) \mathrm{d}\boldsymbol{\xi} \\
&= D\left(\rho \,\middle\|\, \sigma\right) - \int f(\boldsymbol{\xi}) \log \frac{g(\boldsymbol{\xi})}{f(\boldsymbol{\xi})} \\
&= D\left(\rho \,\middle\|\, \sigma\right) + D\left(f \,\middle\|\, g\right) \ .
\end{aligned}
$$

Here the inequality we used is the joint convexity of the relative entropy (see [38, Theorem 1]). For the third equality we used property (31), and for the fourth equality we used unitary invariance of the relative entropy and the trace as well as the fact that f is a probability distribution. The last equality follows from the definition (13) of the divergence. $\qquad\square$

To convert Lemma 1 into a statement about Fisher information, we need the following covariance property of the convolution operation (26): it breaks down translations of the state $f \star_t \rho$ into translations of the function f and the state ρ, respectively.

Lemma 2. *Let $\omega_q, \omega_c > 0$ and $t \geq 0$. Then*

$$(f \star_t \rho)^{(\omega \boldsymbol{\theta})} = f^{(\omega_c \boldsymbol{\theta})} \star_t \rho^{(\omega_q \boldsymbol{\theta})} \qquad \text{for all } \boldsymbol{\theta} \in \mathbb{R}^{2d} \ , \tag{32}$$

where $\omega = \omega_q + \sqrt{t}\omega_c$.

11

Proof. According to Definition (26) and (20) we have

$$(f \star_t \rho)^{(\omega \boldsymbol{\theta})} = \int f(\boldsymbol{\xi}) W(\omega \boldsymbol{\theta}) W(\sqrt{t} \boldsymbol{\xi}) \rho W(\sqrt{t} \boldsymbol{\xi})^\dagger W(\omega \boldsymbol{\theta})^\dagger d\boldsymbol{\xi}$$

$$= \int f(\boldsymbol{\xi}) W(\omega \boldsymbol{\theta} + \sqrt{t} \boldsymbol{\xi}) \rho W(\omega \boldsymbol{\theta} + \sqrt{t} \boldsymbol{\xi})^\dagger d\boldsymbol{\xi} \ . \tag{33}$$

On the other hand, we similarly have

$$f^{(\omega_c \boldsymbol{\theta})} \star_t \rho^{(\omega_q \boldsymbol{\theta})} = \int f(\boldsymbol{\xi} - \omega_c \boldsymbol{\theta}) W(\sqrt{t} \boldsymbol{\xi}) W(\omega_q \boldsymbol{\theta}) \rho W(\omega_q \boldsymbol{\theta})^\dagger W(\sqrt{t} \boldsymbol{\xi})^\dagger d\boldsymbol{\xi}$$

$$= \int f(\boldsymbol{\xi} - \omega_c \boldsymbol{\theta}) W(\omega_q \boldsymbol{\theta} + \sqrt{t} \boldsymbol{\xi}) \rho W(\omega_q \boldsymbol{\theta} + \sqrt{t} \boldsymbol{\xi})^\dagger d\boldsymbol{\xi} \ . \tag{34}$$

Through a simple change of variables and recalling that $\omega = \omega_q + \sqrt{t} \omega_c$, the claim (32) follows from (33) and (34). $\qquad \square$

Combining the data processing inequality of Lemma 1 with Lemma 2, we prove an inequality which may be seen as a classical-quantum version of the Stam inequality. It relates the Fisher information of the state $f \star_t \rho$ to the Fisher informations of f and ρ, respectively.

Theorem 1. *(Quantum Stam inequality) Let $\omega_q, \omega_c \in \mathbb{R}$, and $t \geq 0$. Then*

$$\omega^2 J(f \star_t \rho) \leq \omega_q^2 J(\rho) + \omega_c^2 J(f) \ , \tag{35}$$

where $\omega = \omega_q + \sqrt{t} \omega_c$. In particular,

$$J(f \star_t \rho)^{-1} - J(\rho)^{-1} - t J(f)^{-1} \geq 0 \ . \tag{36}$$

Proof. Let $\boldsymbol{\theta_0} = (\theta_0^{(1)}, \dots, \theta_0^{(2d)}) \in \mathbb{R}^{2d}$. For $j = 1, \dots 2d$ and $\theta_j \in \mathbb{R}$, introduce the vector

$$\tilde{\boldsymbol{\theta}}_{\boldsymbol{j}} = \tilde{\boldsymbol{\theta}}_{\boldsymbol{j}}(\theta_j) = (\theta_0^{(1)}, \dots, \theta_0^{(j-1)}, \theta_j, \theta_0^{(j+1)}, \dots, \theta_0^{(2d)}) \ .$$

Define the functions

$$f(\theta_j) := D\left(f^{(\omega_c \boldsymbol{\theta_0})} \big\| f^{(\omega_c \tilde{\boldsymbol{\theta}}_{\boldsymbol{j}})} \right) + D\left(\rho^{(\omega_q \boldsymbol{\theta_0})} \big\| \rho^{(\omega_q \tilde{\boldsymbol{\theta}}_{\boldsymbol{j}})} \right) \ ,$$

and

$$g(\theta_j) := D\left(f^{(\omega_c \boldsymbol{\theta_0})} \star_t \rho^{(\omega_q \boldsymbol{\theta_0})} \big\| f^{(\omega_c \tilde{\boldsymbol{\theta}}_{\boldsymbol{j}})} \star_t \rho^{(\omega_q \tilde{\boldsymbol{\theta}}_{\boldsymbol{j}})} \right) \ .$$

From the definition of the relative entropy and the data processing inequality (30), for every θ_j, we have

$$0 \leq g(\theta_j) \leq f(\theta_j)$$
$$0 = f(\theta_0^{(j)}) = g(\theta_0^{(j)}) \ .$$

The second derivative of g can be written as the limit

$$\frac{d^2}{d\theta_j^2}\bigg|_{\theta_j = \theta_0^{(j)}} g(\theta_j) = \lim_{\epsilon \to 0} \frac{g(\theta_0^{(j)} + \epsilon) - 2g(\theta_0^{(j)}) + g(\theta_0^{(j)} - \epsilon)}{\epsilon^2} \ ,$$

and, therefore, is it bounded

$$0 \leq \left.\frac{d^2}{d\theta_j^2}\right|_{\theta_j = \theta_0^{(j)}} g(\theta_j) \leq \left.\frac{d^2}{d\theta_j^2}\right|_{\theta_j = \theta_0^{(j)}} f(\theta_j) \ .$$

Since

$$\text{tr}\left(J(\{f^{(\omega_c\boldsymbol{\theta})}\}; \boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta_0}}\right) + \text{tr}\left(J(\{\rho^{(\omega_q\boldsymbol{\theta})}\}; \boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta_0}}\right) = \sum_{j=1}^{2d} \left.\frac{d^2}{d\theta_j^2}\right|_{\theta_j = \theta_0^{(j)}} f(\theta_j) \ ,$$

and

$$\text{tr}\left(J(\{f^{(\omega_c\boldsymbol{\theta})} \star_t \rho^{(\omega_q\boldsymbol{\theta})}\}; \boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta_0}}\right) = \sum_{j=1}^{2d} \left.\frac{d^2}{d\theta_j^2}\right|_{\theta_j = \theta_0^{(j)}} g(\theta_j) \ ,$$

we conclude that

$$\text{tr}\left(J(\{f^{(\omega_c\boldsymbol{\theta})} \star_t \rho^{(\omega_q\boldsymbol{\theta})}\}; \boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta_0}}\right) \leq \text{tr}\left(J(\{f^{(\omega_c\boldsymbol{\theta})}\}; \boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta_0}}\right) + \text{tr}\left(J(\{\rho^{(\omega_q\boldsymbol{\theta})}\}; \boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta_0}}\right) \ . \tag{37}$$

We remark that above inequality can also be derived as a matrix inequality without tracing both sides. However, this is not required for our purposes and our definition of the Fisher information.

Using Lemma 2, the left-hand side of this equation for $\boldsymbol{\theta_0} = \boldsymbol{0}$ can be written as

$$\text{tr}\left(J(\{f^{(\omega_c\boldsymbol{\theta})} \star_t \rho^{(\omega_q\boldsymbol{\theta})}\}; \boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{0}}\right) = \text{tr}\left(J(\{(f \star_t \rho)^{(\omega\boldsymbol{\theta})}\}; \boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{0}}\right) = \omega^2 J(f \star_t \rho) \ . \tag{38}$$

The right-hand side of Eq. (37) can be simplified by noticing that both the classical and quantum Fisher information matrices satisfy reparametrization formulas (see [13, Lemma IV.1])

$$J(\{f^{(\omega_c\boldsymbol{\theta})}\}; \boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta_0}} = \omega_c^2 J(\{f^{(\boldsymbol{\theta})}\}; \boldsymbol{\theta}) \qquad \text{and} \qquad J(\{\rho^{(\omega_q\boldsymbol{\theta})}\}; \boldsymbol{\theta}) = \omega_q^2 J(\{\rho^{(\boldsymbol{\theta})}\}; \boldsymbol{\theta}) \ .$$

Therefore, taking $\boldsymbol{\theta} = \boldsymbol{0}$ leads to

$$\text{tr}\left(J(\{f^{(\omega_c\boldsymbol{\theta})}\}; \boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{0}}\right) + \text{tr}\left(J(\{\rho^{(\omega_q\boldsymbol{\theta})}\}; \boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{0}}\right) = \omega_c^2 J(f) + \omega_q^2 J(\rho) \ . \tag{39}$$

With (37), (38), and (39), we arrive at the desired result (35). Finally, setting $\omega_q = \frac{J(\rho)^{-1}}{J(\rho)^{-1}+tJ(f)^{-1}}$ and $\omega_c = \frac{\sqrt{t}J(f)^{-1}}{J(\rho)^{-1}+tJ(f)^{-1}}$, we obtain (36).

$\square$

In the next lemma we show how the quantum Stam inequality implies an isoperimetric inequality for the quantum Fisher information.

Lemma 3. *(Quantum Fisher information isoperimetric inequality) The following inequality holds:*

$$\left.\frac{d}{dt}\right|_{t=0} \left[\frac{1}{2d} J(e^{t\mathcal{L}_{\text{heat}}}(\rho))\right]^{-1} \geq 1 \ . \tag{40}$$

Proof. Recall that we have $e^{t\mathcal{L}_{\text{heat}}}(\rho) = f_Z \star_t \rho$ for a Gaussian random variable Z (27). In the quantum Stam inequality (36), take $f = f_Z$, then

$$\frac{1}{t}\left(J(f_Z \star_t \rho)^{-1} - J(\rho)^{-1}\right) \geq J(f_Z)^{-1} = (2d)^{-1} .$$

Taking the limit $t \to 0$, we arrive at the desired inequality. □

The isoperimetric inequality is tight for one mode ($d = 1$) and saturated by the Gaussian thermal state

$$\omega_{\mathbf{n}} = \frac{1}{\mathbf{n}+1} \sum_{j=0}^{\infty} \left(\frac{\mathbf{n}}{\mathbf{n}+1}\right)^j |j\rangle\langle j| \tag{41}$$

with mean-photon number $\mathbf{n}$: As shown in Appendix B, we have

$$\frac{d}{dt}\bigg|_{t=0} \left[\frac{1}{2}J(e^{t\mathcal{L}_{\text{heat}}}(\omega_{\mathbf{n}}))\right]^{-1} = \frac{1}{\mathbf{n}(\mathbf{n}+1)} \log^{-2}\left(1 + \frac{1}{\mathbf{n}}\right) \to 1 \qquad \text{as } \mathbf{n} \to \infty .$$

As in classical information theory, the isoperimetric inequality for the quantum Fisher information implies concavity of the entropy power under diffusion as an immediate consequence. We define the entropy power as

$$N(\rho) = \exp\left(S(\rho)/d\right) . \tag{42}$$

Theorem 2. *(Concavity of the quantum entropy power) The entropy power along trajectories of the diffusion semigroup* (23) *is concave, i.e.*

$$\frac{d^2}{dt^2}\bigg|_{t=0} N(e^{t\mathcal{L}_{\text{heat}}}(\rho)) \leq 0 .$$

Proof. Two applications of de Bruijn identity (24) yield

$$\frac{d^2}{dt^2}\bigg|_{t=0} N(e^{t\mathcal{L}_{\text{heat}}}(\rho)) = N(\rho)\left(\left[\frac{1}{2d}J(\rho)\right]^2 + \frac{1}{2d}\frac{d}{dt}\bigg|_{t=0} J(e^{t\mathcal{L}_{\text{heat}}}(\rho))\right) .$$

The quantum Fisher information isoperimetric inequality (40) is equivalent to

$$\frac{1}{2d}J(\rho)^2 + \frac{d}{dt}\bigg|_{t=0} J(e^{t\mathcal{L}_{\text{heat}}}(\rho)) \leq 0 .$$

This completes the proof. □

To proceed, we establish bounds on the asymptotic scaling of the entropy power for large times. The following lemma follows directly from [5] and [13, Corollary III.4] (see also [39]).

Lemma 4 (Asymptotic scaling of the entropy power under the heat flow). *Let*

$$\mathcal{C}_c = \frac{1}{2}\sum_{j=1}^{2d} \frac{\partial^2}{\partial \xi_j^2} \tag{43}$$

14

be the generator of the classical heat diffusion semigroup on $\mathbb{R}^{2d}$. In the limit $t \to \infty$, we have

$$\exp\left(H(e^{t\mathcal{C}_c}(f))/d\right) = (2\pi e)t + O(1) \ ,$$
$$\exp\left(S(e^{t\mathcal{L}_{\text{heat}}}(\rho))/d\right) = (2\pi e)t + O(1)$$

independent of the probability density function f on $\mathbb{R}^{2d}$ and the d-mode state ρ, respectively.

Having the same scaling for classical and quantum heat flows motivated the choice of constants in (19) and (23).

Note that if X is a random variable with probability density function f, then $e^{t\mathcal{C}_c}(f)$ is the probability density function of the random variable $X + \sqrt{t}Z$ obtained by adding a centered Gaussian random variable with unit variance, see Eq. (27).

Lemma 5. *We have*

$$e^{\xi\mathcal{L}_{\text{heat}}}(f \star_t \rho) = e^{\nu\mathcal{C}_c}(f) \star_t e^{\mu\mathcal{L}_{\text{heat}}}(\rho)$$

whenever $\xi = \mu + t\nu$.

Proof. Observe that writing $\star$ for $\star_1$, we get

$$\begin{aligned}
e^{\xi\mathcal{L}_{\text{heat}}}(f_X \star_t \rho) &= f_Z \star_\xi (f_{\sqrt{t}X} \star \rho) \\
&= f_{\sqrt{\xi}Z} \star (f_{\sqrt{t}X} \star \rho) \\
&= f_{\sqrt{\xi}Z + \sqrt{t}X} \star \rho \ .
\end{aligned}$$

On the other hand,

$$\begin{aligned}
e^{\nu\mathcal{C}_c}(f) \star_t e^{\mu\mathcal{L}_{\text{heat}}}(\rho) &= f_{X+\sqrt{\nu}Z_1} \star_t \left(e^{\mu\mathcal{L}_{\text{heat}}}(\rho)\right) \\
&= f_{\sqrt{t}(X+\sqrt{\nu}Z_1)} \star \left(f_{\sqrt{\mu}Z_2} \star \rho\right) \\
&= f_{\sqrt{t}X + \sqrt{t\nu}Z_1 + \sqrt{\mu}Z_2} \star \rho \\
&= f_{\sqrt{t}X + \sqrt{\mu+t\nu}Z} \star \rho \\
&= f_{\sqrt{t}X + \sqrt{\xi}Z} \star \rho \ .
\end{aligned}$$

We have used properties (28) and (29). In the penultimate step we have used that for independent unit-variance centered Gaussian random variables Z_1 and Z_2, we have $aZ_1 + bZ_2 = \sqrt{a^2 + b^2}Z$. Hence the two expressions are equal and the statement follows.

$\square$

The next theorem presents the entropy power inequality for both the convolution operation (26) and the heat diffusion semigroup (23).

Theorem 3. *(Entropy power inequality) For $t \geq 0$, the following inequality holds*

$$N(f \star_t \rho) \geq N(\rho) + tN(f) \ .$$

In particular, choosing $f = f_Z$ as the distribution of a unit-variance centered Gaussian defined in Eq. (27), we have

$$N(e^{t\mathcal{L}_{\text{heat}}}(\rho)) \geq N(\rho) + t\,2\pi e \ . \tag{44}$$

Proof. The proof is inspired by the proof of the entropy power inequality in [13], which itself is inspired by the proof for classical random variables by Blachman [5]. Here we provide all necessary details modified to the present situation. For $\mu, \nu, \xi \geq 0$, define functions

$$\mu \to E_A(\mu) := \exp\left(S(e^{\mu \mathcal{L}_{\text{heat}}}(\rho))/d\right) \ ,$$
$$\nu \to E_B(\nu) := \exp\left(H(e^{\nu \mathcal{L}_{\text{heat,cl}}}(f))/d\right) \ ,$$
$$\xi \to E_C(\xi) := \exp\left(S(e^{\xi \mathcal{L}_{\text{heat}}}(f \star_t \rho)))/d\right) \ ,$$

where $\mathcal{C}_c$ is the generator of the classical heat semigroup as defined in Eq. (43).

The initial value problems

$$\begin{aligned}
\dot{\mu}(s) &= E_A(\mu(s)) \ , \quad \mu(0) = 0 \ , \\
\dot{\nu}(s) &= E_B(\nu(s)) \ , \quad \nu(0) = 0
\end{aligned} \tag{45}$$

have solutions $\mu(\cdot), \nu(\cdot)$. Fix a pair of such solutions $(\mu(\cdot), \nu(\cdot))$ and $t \geq 0$. Define

$$\xi(s) := \mu(s) + t\nu(s) \ . \tag{46}$$

These functions diverge, i.e.

$$\lim_{s \to \infty} \mu(s) = \lim_{s \to \infty} \nu(s) = \lim_{s \to \infty} \xi(s) = \infty \ , \tag{47}$$

because of (45) and $E_{A,B} \geq 1$.

Consider the function

$$\delta(s) := \frac{E_A(\mu(s)) + t E_B(\nu(s))}{E_C(\xi(s))} \ .$$

With the initial conditions (45), it follows that the claim of the theorem is equivalent to

$$\delta(0) \leq 1 \ .$$

This inequality follows from two facts: first, the fact that

$$\lim_{s \to \infty} \delta(s) = 1 \ ,$$

as follows from the asymptotic scaling shown in Lemma 4, the divergence (47), and the choice (46) of $\xi(s)$; second, the fact that

$$\dot{\delta}(s) \geq 0 \qquad \text{for all } s \geq 0 \ . \tag{48}$$

It remains to show identity (48). Computing the derivative of δ leads to the following equality

$$\dot{\delta}(s) = \frac{\dot{E}_A(\mu(s))\dot{\mu}(s) + t\dot{E}_B(\nu(s))\dot{\nu}(s)}{E_C(\xi(s))} - \frac{E_A(\mu) + t E_B(\nu)}{E_C(\xi)^2}\dot{E}_C(\xi(s))\dot{\xi}(s) \ . \tag{49}$$

Define the Fisher informations

$$J_A(\mu) := J(e^{\mu \mathcal{L}_{\text{heat}}}(\rho)) \ ,$$
$$J_B(\nu) := J(e^{\nu \mathcal{C}_{\text{c}}}(f)) \ ,$$
$$J_C(\xi) := J(e^{\xi \mathcal{L}_{\text{heat}}}(f \star_t \rho)) \ ,$$

for $\mu, \nu, \xi \geq 0$. From the quantum de Bruijn identity (24) and the classical de Bruijn identity (17) we obtain

$$\dot{E}_V(\zeta) = \frac{1}{2d} E_V(\zeta) J_V(\zeta) \qquad \text{where } V \in \{A, B, C\} \ .$$

With these identities and Eq. (45), Eq. (49) is equivalent to

$$2d\dot{\delta}(s) = \frac{E_A^2 J_A + t E_B^2 J_B}{E_C} - \frac{(E_A + t E_B)^2}{E_C^2} E_C J_C \ , \tag{50}$$

where we have used the shorthand notation $E_A = E_A(\mu(s))$, $E_B = E_B(\nu(s))$, and $E_C = E_C(\xi(s))$, and similarly for $J_{A,B,C}$.

Recall that by Lemma 5, $e^{\xi \mathcal{L}_{\text{heat}}}(f \star_t \rho) = e^{\nu \mathcal{C}_{\text{c}}}(f) \star_t e^{\mu \mathcal{L}_{\text{heat}}}(\rho)$, and by the quantum Stam inequality (36), we have the bound

$$J_C \leq \frac{J_A J_B}{t J_A + J_B} \ .$$

Inserting this upper bound into (50), we obtain

$$2d\dot{\delta} E_C \geq E_A^2 J_A + t E_B^2 J_B - (E_A + t E_B)^2 \frac{J_A J_B}{t J_A + J_B} = \frac{t(E_A J_A - E_B J_B)^2}{t J_A + J_B} \geq 0 \ .$$

This proves (48). $\qquad \square$

As the last statement in this section, we derive an isoperimetric inequality for entropies from the entropy power inequality.

Theorem 4. *(Isoperimetric inequality for entropies)* We have

$$\frac{1}{d} J(\rho) N(\rho) \geq 4\pi e \ . \tag{51}$$

Proof. Applying the de Bruijn identity (24) to the definition (42) of the entropy power $N(\rho)$, we obtain

$$\left. \frac{d}{dt} \right|_{t=0} N(e^{t \mathcal{L}_{\text{heat}}}(\rho)) = \frac{1}{2d} J(\rho) N(\rho) \ .$$

On the other hand, for $t \geq 0$, the entropy power inequality (44) reduces to

$$\frac{1}{t} [N(e^{t \mathcal{L}_{\text{heat}}}(\rho)) - N(\rho)] \geq 2\pi e \ .$$

Therefore, taking the limit $t \to 0$, we obtain the desired bound. $\qquad \square$

17

To conclude this section, we remark that the isoperimetric inequality for entropies (51) is tight in the one-mode case, $d = 1$: For a Gaussian thermal state $\omega_{\mathbf{n}}$ (41) with mean photon number $\mathbf{n}$ we obtain

$$J(\omega_{\mathbf{n}})N(\omega_{\mathbf{n}}) = 4\pi \left(\frac{\mathbf{n}+1}{\mathbf{n}}\right)^{\mathbf{n}} \log\left(\frac{\mathbf{n}+1}{\mathbf{n}}\right)^{\mathbf{n}+1} \to 4\pi e \qquad \text{for } \mathbf{n} \to \infty .$$

Detailed calculations are provided in Appendix C.

5 Gaussian optimality for energy-constrained entropy rates

In this section we show that Gaussian thermal states minimize the entropy rate for the one-mode attenuator semigroup among states with bounded mean photon number. The semigroup is defined by its generator

$$\mathcal{L}_-(\rho) = a\rho a^\dagger - \frac{1}{2}\{a^\dagger a, \rho\} . \tag{52}$$

We prove the following theorem:

Theorem 5. *For any $\mathbf{n} > 0$, the infimum $\inf_{\rho:\mathrm{tr}(\hat{n}\rho)\leq\mathbf{n}} \frac{d}{dt}\big|_{t=0} S(e^{t\mathcal{L}_-}(\rho))$ over states ρ with mean photon number $\mathbf{n}$ is achieved by the Gaussian thermal state $\omega_{\mathbf{n}}$ defined in (41). In particular,*

$$\inf_{\rho:\mathrm{tr}(\hat{n}\rho)\leq\mathbf{n}} \frac{d}{dt}\bigg|_{t=0} S(e^{t\mathcal{L}_-}(\rho)) = \begin{cases} -\mathbf{n}\log\left(1 + \frac{1}{\mathbf{n}}\right) & \text{if } \mathbf{n} > 0 , \\ 0 & \text{if } \mathbf{n} = 0 . \end{cases}$$

The proof proceeds by reduction to a recent result by De Palma, Trevisan, and Giovannetti [24], where it is shown that Gaussian thermal states minimize the entropy rate of the quantum attenuator among all states with a given input entropy[3]. Our argument additionally uses the recently introduced concept of Fock majorization and associated results by Jabbour, Garcia-Patron and Cerf [31], as well as the (classical) Gaussian maximum entropy principle.

In more detail, we first show that the problem of minimizing the entropy rate reduces to the study of properties of a classical semigroup describing a pure-death process. This connection was used previously in [26] (see Section 2.2) and is also an essential first step in [24]. More explicitly, in Section 5.1 (Theorem 6) we prove the identity

$$\inf_{\rho:\mathrm{tr}(\hat{n}\rho)\leq\mathbf{n}} \frac{d}{dt}\bigg|_{t=0} S(e^{t\mathcal{L}_-}(\rho)) = \inf_{p:\mathbb{E}_p[N]\leq\mathbf{n}} \frac{d}{dt}\bigg|_{t=0} H(e^{t\mathcal{C}_-}(p)) .$$

Here the infimum is over all probability distributions p on $\mathbb{N}_0$ with expectation value $\mathbb{E}_p[N]$ bounded by $\mathbf{n}$, the quantity $H(p)$ is the Shannon entropy of the distribution p, and $\mathcal{C}_-$ is the generator of a semigroup describing a classical pure-death process (see (53) for a precise definition).

---

[3]More precisely, [24, Theorem 6] states that the entropy rate of a state with finite support is lower bounded by that of a Gaussian state with the same entropy. Since we are interested in a statement applicable to arbitrary states, we use their Theorem 24 instead: the latter gives an analogous statement for the corresponding classical process but without a finiteness assumption.

Using the results of [24], we then show that the entropy rate for the classical process is optimized by a geometric distribution: In Section 5.2 (Theorem 6) we prove that for $\mathbf{n} > 0$

$$\inf_{p:\mathbb{E}_p[N]\leq\mathbf{n}} \frac{d}{dt}\Big|_{t=0} H(e^{t\mathcal{C}_-}(p)) = -\mathbf{n}\log\left(1+\frac{1}{\mathbf{n}}\right) .$$

Finally, in Section 5.3, we calculate the entropy rate for the Gaussian thermal state $\omega_\mathbf{n}$ with mean photon number at most $\mathbf{n}$ and find that

$$\frac{d}{dt}\Big|_{t=0} S(e^{t\mathcal{L}_-}(\omega_\mathbf{n})) = -\mathbf{n}\log\left(1+\frac{1}{\mathbf{n}}\right).$$

5.1 Connection to a classical pure-death process

For an initial state $\rho = \sum_n p_n |n\rangle\langle n|$ which is diagonal in the number state basis $\{|n\rangle = \frac{\left(a^\dagger\right)^n}{\sqrt{n!}}|0\rangle\}_{n\in\mathbb{N}_0}$ (where $\hat{n}|0\rangle = 0$), the time-evolved state has the same form, i.e., $\rho(t) = e^{t\mathcal{L}_-}(\rho) = \sum_n p_n(t)|n\rangle\langle n|$. Thus the attenuator semigroup $\{e^{t\mathcal{L}_-}\}_{t\geq 0}$ gives rise to a semigroup $\{e^{t\mathcal{C}_-}\}_{t\geq 0}$ on classical probability distributions by $p(t) = e^{t\mathcal{C}_-}(p)$. Its generator $\mathcal{C}_-$ describes the dynamics of a classical pure-death process. It can be obtained from (52) by inserting a number state $|n\rangle$: it is straightforward to check that

$$\mathcal{L}_-(|n\rangle\langle n|) = \begin{cases} n(|n-1\rangle\langle n-1| - |n\rangle\langle n|) & \text{for } n > 0 , \\ 0 & \text{for } n = 0 . \end{cases}$$

In particular, the coefficients $\{p_n(t)\}_{n\in\mathbb{N}_0}$ satisfy the system of differential equations

$$\dot{p}_n(t) = -np_n(t) + (n+1)p_{n+1}(t) \qquad \text{for all } n \in \mathbb{N}_0 , \tag{53}$$

with initial condition $p_n(0) = p_n$ for $n \in \mathbb{N}_0$. The expression on the right-hand side of (53) defines the generator $\mathcal{C}_-$, that is, we have

$$(\mathcal{C}_-(p))_n = -np_n + (n+1)p_{n+1} \quad \text{for all } n \in \mathbb{N}_0 . \tag{54}$$

The following theorem reduces the problem of minimizing the entropy rate for the quantum attenuator semigroup to the classical problem of minimizing the entropy rate for this pure-death process.

Theorem 6 (Correspondence to classical problem). *We have the identity*

$$\inf_{\rho:\text{tr}(\hat{n}\rho)\leq\mathbf{n}} \frac{d}{dt}\Big|_{t=0} S(e^{t\mathcal{L}_-}(\rho)) = \inf_{p:\mathbb{E}_p[N]\leq\mathbf{n}} \frac{d}{dt}\Big|_{t=0} H(e^{t\mathcal{C}_-}(p)) ,$$

where $\mathbb{E}_p[N] = \sum_{n=0}^\infty np_n$ and $H(p) = -\sum_{n=0}^\infty p_n\log p_n$ are the expectation value and entropy of the distribution p, respectively.

The proof of Theorem 6 relies on results obtained in [32] and [31]. Here we review the necessary definitions and results, and specialize them to our situation.

Definition 1 (Majorization). *Let p and q be decreasing summable sequences of positive numbers. Then p weakly sub-majorizes q, $q \prec_w p$, if and only if*

$$\sum_{i=0}^{n} q_i \leq \sum_{i=0}^{n} p_i \qquad \text{for all } n \in \mathbb{N}.$$

Definition 2. *Let P and Q be positive trace-class operators with eigenvalues $\{p_n\}_{n \in \mathbb{N}}$ and $\{q_n\}_{n \in \mathbb{N}}$, arranged in decreasing order. Then P weakly majorizes Q, i.e., $Q \prec_w P$, if and only if $q \prec_w p$. Also, P majorizes Q, i.e., $Q \prec P$, if the traces of P and Q are identical and $Q \prec_w P$.*

Definition 3 (Fock rearrangement). *Let X be a positive trace-class operator with eigenvalues $\{x_n\}_{n \in \mathbb{N}_0}$ in decreasing order. The Fock rearrangement (or passive rearrangement) is defined as*

$$X^{\downarrow} := \sum_{n=0}^{\infty} x_n \, |n\rangle \langle n| .$$

Our restriction on the mean photon number forces us to consider an additional majorization relation. In contrast, in [24], where the authors restrict the input entropy instead, Fock majorization does not need to be considered.

Definition 4 (Fock majorization). *The Fock majorization relation, denoted $\prec_F$, was introduced in [31] as follows:*

$$\sigma \prec_F \rho \qquad \Leftrightarrow \qquad \mathsf{tr}(\Pi_n \sigma) \leq \mathsf{tr}(\Pi_n \rho) \quad \text{for all } n \in \mathbb{N}_0 , \tag{55}$$

where

$$\Pi_n = \sum_{j=0}^{n} |j\rangle\langle j|. \tag{56}$$

Theorem 7. *([32]) For any state ρ and all $t \geq 0$ we have*

$$e^{t\mathcal{L}_-}(\rho) \prec e^{t\mathcal{L}_-}(\rho^{\downarrow}) . \tag{57}$$

This implies that

$$S\left(e^{t\mathcal{L}_-}(\rho)\right) \geq S\left(e^{t\mathcal{L}_-}(\rho^{\downarrow})\right) .$$

Proof. The first statement is shown in [32, Eq. (VI.10)], and the second statement then follows from [32, Theorem III.3]. $\qquad\square$

Lemma 6. *([32, Lemma IV.9]) Suppose X, Y, Z are positive trace-class operators with*

$$Y \prec_w Z \qquad \text{and} \qquad Z^{\downarrow} = Z .$$

Then

$$\mathsf{tr}(XY) \leq \mathsf{tr}(X^{\downarrow} Z) .$$

We begin proving Theorem 6 by investigating the change of the mean photon number under the Fock rearrangement procedure.

Lemma 7. *For any state ρ, the Fock rearrangement does not increase the mean photon number*

$$\mathsf{tr}(\hat{n}\rho^{\downarrow}) \leq \mathsf{tr}(\hat{n}\rho) \ ,$$

where $\hat{n} = a^{\dagger}a$.

Proof. Let $n \in \mathbb{N}_0$ be arbitrary, and set $X = \Pi_n$ with Π_n defined in (56). Then clearly $X^{\downarrow} = X = \Pi_n$. Setting $Y = \rho$, $Z = \rho^{\downarrow}$, we have $Z^{\downarrow} = Z$, and $Y \prec Z$, according to (57) for $t = 0$, i.e.

$$\rho \prec \rho^{\downarrow} \ .$$

Thus, Lemma 6 leads to $\mathsf{tr}(\Pi_n\rho) \leq \mathsf{tr}(\Pi_n\rho^{\downarrow})$. According to Definition (55), this is equivalent to

$$\rho \prec_F \rho^{\downarrow}.$$

It was shown in ([31, Eq. (5)]) that

$$\rho \prec_F \sigma \qquad \Rightarrow \qquad \mathsf{tr}(\hat{n}\sigma) \leq \mathsf{tr}(\hat{n}\rho) \ .$$

The claim follows by taking $\sigma = \rho^{\downarrow}$ in the last statement. $\qquad\square$

Proof of Theorem 6. Since $S(\rho^{\downarrow}) = S(\rho)$, Theorem 7 implies

$$\frac{d}{dt}\bigg|_{t=0} S\left(e^{t\mathcal{L}_-}(\rho)\right) \geq \frac{d}{dt}\bigg|_{t=0} S\left(e^{t\mathcal{L}_-}(\rho^{\downarrow})\right) \ .$$

With Lemma 7, we therefore obtain

$$\inf_{\rho:\mathsf{tr}(\hat{n}\rho)\leq \mathbf{n}} \frac{d}{dt}\bigg|_{t=0} S\left(e^{t\mathcal{L}_-}(\rho)\right) = \inf_{\rho=\rho^{\downarrow}:\mathsf{tr}(\hat{n}\rho^{\downarrow})\leq \mathbf{n}} \frac{d}{dt}\bigg|_{t=0} S\left(e^{t\mathcal{L}_-}(\rho^{\downarrow})\right) \ .$$

Since $\rho^{\downarrow}$ is a passive, Fock-rearranged state, the right-hand side is a classical problem related to the pure-death process (53), and, therefore, can be replaced with the infimum of the entropy rate of a probability distribution evolving under a pure-death process. Thus the claim follows. $\qquad\square$

5.2 Geometric distributions optimize entropy rates of the classical death process under energy constraint

For a probability distribution p on $\mathbb{N}_0$, let

$$J_-(p) := 2\frac{d}{dt}\bigg|_{t=0} H(e^{t\mathcal{C}_-}(p))$$

denote the entropy rate when p evolves under the classical death-process $\mathcal{C}_-$ (cf. (54)). We are interested in distributions p with a fixed expectation value $\mathbb{E}_p[N] = \sum_{n=0}^{\infty} np_n$. The main result we use here is [24, Theorem 24]: it states that for any probability distribution p on $\mathbb{N}_0$, the quantity $J_-(p)$ is bounded by

$$\inf_{p:H(p)\leq H} J_-(p) \geq 2 \inf_{p:H(p)\leq g(\mathbf{n})} f(H(p)) \tag{58}$$

where $f(H) = -g^{-1}(H)g'(g^{-1}(H))$ and where $g(\mathbf{n}) = (\mathbf{n}+1)\log(\mathbf{n}+1) - \mathbf{n}\log\mathbf{n}$ is the entropy of a geometric distribution with expectation value $\mathbf{n}$ (or equivalently the entropy of a Gaussian state with mean photon number $\mathbf{n}$). We use this to show the following:

Theorem 8. *The infimum* $\inf_{p:\mathbb{E}[N]\leq\mathbf{n}} J_-(p)$ *is achieved by the geometric distribution* $p_k^{\mathrm{geo},\mathbf{n}} = (1-r)r^k$ *with* $r = \frac{\mathbf{n}}{\mathbf{n}+1}$. *In particular, for* $\mathbf{n} > 0$,

$$\inf_{p:\mathbb{E}[N]\leq\mathbf{n}} J_-(p) = -2\mathbf{n}\log\left(1 + \frac{1}{\mathbf{n}}\right) .$$

Proof. We show that

$$\inf_{p:\mathbb{E}[N]=\mathbf{n}} J_-(p) = J(p^{\mathrm{geo},\mathbf{n}}) = -2\mathbf{n}\log\left(1 + \frac{1}{\mathbf{n}}\right) . \tag{59}$$

Since the right-hand side is monotonically decreasing with $\mathbf{n}$, Eq. (59) implies the claim of the theorem.

The geometric distribution $p^{\mathrm{geo},\mathbf{n}} = (1-r)r^k$ with $r = \frac{\mathbf{n}}{\mathbf{n}+1}$ has expectation value $\mathbb{E}_{p^{\mathrm{geo},\mathbf{n}}}[N] = \mathbf{n}$ and entropy $H(p^{\mathrm{geo},\mathbf{n}}) = g(\mathbf{n})$. By the maximum entropy principle [40, Chapter 12], we know that geometric distributions are the distributions with maximal entropy among all distributions with a fixed expectation value $\mathbb{E}_p[N]$. Therefore we have

$$\mathbb{E}_p[N] = \mathbf{n} \qquad \Rightarrow \qquad H(p) \leq g(\mathbf{n}) .$$

Combining this with (58) we obtain

$$\inf_{p:\mathbb{E}_p[N]=\mathbf{n}} J_-(p) \geq \inf_{p:H(p)\leq g(\mathbf{n})} J_-(p) \geq 2\inf_{p:H(p)\leq g(\mathbf{n})} f(H(p)) ,$$

Since f is decreasing [24, Lemma 5], it follows that

$$\inf_{p:\mathbb{E}[N]=\mathbf{n}} J_-(p) \geq 2f(g(\mathbf{n})) = -2\mathbf{n}\log\left(1 + \frac{1}{\mathbf{n}}\right) . \tag{60}$$

However, since $\mathbb{E}_{p^{\mathrm{geo},\mathbf{n}}}[N] = \mathbf{n}$ and $J_-(p^{\mathrm{geo},\mathbf{n}}) = -2\mathbf{n}\log\left(1 + \frac{1}{\mathbf{n}}\right)$, we have equality in (60). $\qquad\square$

5.3 Gaussian optimality of entropy rates for the quantum attenuator semigroup

Proof of Theorem 5. From Theorem 6 and Theorem 8 we have

$$\inf_{\rho:\mathrm{tr}(\hat{n}\rho)\leq\mathbf{n}} \frac{d}{dt}\Big|_{t=0} S(e^{t\mathcal{L}_-}(\rho)) = -\mathbf{n}\log\left(1 + \frac{1}{\mathbf{n}}\right).$$

Let $\omega_{\mathbf{n}}$ be the Gaussian thermal state with mean photon number $\mathbf{n}$ as defined in (41). We have $S(\omega_{\mathbf{n}}) = g(\mathbf{n}) = (\mathbf{n}+1)\log(\mathbf{n}+1) - \mathbf{n}\log\mathbf{n}$ for $\mathbf{n} > 0$. Under the map $e^{t\mathcal{L}_-}$, the state $\omega_{\mathbf{n}}$ evolves into the thermal state $\omega_{\mathbf{n}_t}$ according to

$$e^{t\mathcal{L}_-}(\omega_{\mathbf{n}}) = \omega_{\mathbf{n}_t} , \qquad \text{where} \qquad \mathbf{n}_t = e^{-t}\mathbf{n} .$$

In particular,

$$\frac{d}{dt}\Big|_{t=0} S(e^{t\mathcal{L}_-}(\omega_{\mathbf{n}})) = g'(\mathbf{n})\mathbf{n}'_t\Big|_{t=0} = -\mathbf{n}\log\left(1 + \frac{1}{\mathbf{n}}\right) .$$

Therefore the considered infimum is achieved by the Gaussian thermal state $\omega_{\mathbf{n}}$.

$$\square$$

6 Application to fast convergence of the Ornstein-Uhlenbeck semigroup

In this section we consider a one-parameter group of CPTP maps $\{e^{\mathcal{L}_{\mu,\lambda}}\}_{t\geq 0}$ generated by the linear combination

$$\mathcal{L}_{\mu,\lambda} = \mu^2 \mathcal{L}_- + \lambda^2 \mathcal{L}_+ \qquad \text{for } \mu > \lambda > 0 .$$

where $\mathcal{L}_-$ is defined by (52), and $\mathcal{L}_+$ is defined by

$$\mathcal{L}_+(\rho) = a^\dagger \rho a - \frac{1}{2}\{aa^\dagger, \rho\} .$$

In the following, we use the entropy production rates

$$J_\pm(\rho) := 2\frac{d}{dt}S(e^{t\mathcal{L}_\pm}(\rho)) .$$

The factor 2 here is for convenience to match de Bruijn identity (24) for $J(\rho)$. These quantities are related to the Fisher information $J(\rho)$ by

$$J(\rho) = 2\pi(J_-(\rho) + J_+(\rho)) , \tag{61}$$

because of de Bruijn's identity and the fact that $\mathcal{L}_{\text{heat}} = 2\pi\mathcal{L}_- + 2\pi\mathcal{L}_+$.

Lemma 8. *Let $\mu > \lambda > 0$. Then*

$$-\zeta D(\rho\|\sigma_{\mu,\lambda}) - \frac{d}{dt}\bigg|_{t=0} D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) = \frac{\mu^2}{2}J_-(\rho) + \frac{\lambda^2}{2}J_+(\rho) + \zeta S(\rho)$$
$$+ \lambda^2 \log\nu + \zeta\log(1-\nu) \tag{62}$$

for any state ρ, where $\nu = \frac{\lambda^2}{\mu^2}$, $\zeta = \mu^2 - \lambda^2$, and $\sigma_{\mu,\lambda}$ is the fixed point of $\mathcal{L}_{\mu,\lambda}$.

Proof. As mentioned in Section 2.2, the unique fixed point of the semigroup $\{e^{t\mathcal{L}_{\mu,\lambda}}\}_{t\geq 0}$ is the state

$$\sigma_{\mu,\lambda} = (1-\nu)\sum_{n=0}^{\infty}\nu^n|n\rangle\langle n| = (1-\nu)\nu^{\hat{n}},$$

where $\nu = \lambda^2/\mu^2$. In particular, this implies that for any state ρ

$$D(\rho\|\sigma_{\mu,\lambda}) = -S(\rho) - \text{tr}(\rho\log\sigma_{\mu,\lambda}) = -S(\rho) - (\log\nu)\text{tr}(\rho\hat{n}) - \log(1-\nu) . \tag{63}$$

Straightforward calculations show that the mean photon number of ρ converges to the mean photon number of the fixed point $\sigma_{\mu,\lambda}$ with an exponential rate

$$\mathbf{n}_t = \text{tr}(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\hat{n}) = \text{tr}(\rho e^{t\mathcal{L}_{\mu,\lambda}^\dagger}(\hat{n})) = e^{-(\mu^2-\lambda^2)t}\text{tr}(\rho\hat{n}) + (1 - e^{-(\mu^2-\lambda^2)t})\mathbf{n}_\infty ,$$

where $\mathbf{n}_\infty = \text{tr}(\sigma_{\mu,\lambda}\hat{n}) = \frac{\lambda^2}{\mu^2-\lambda^2} = \frac{\nu}{1-\nu}$. Therefore

$$\frac{d}{dt}\bigg|_{t=0}\text{tr}(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\hat{n}) = -(\mu^2 - \lambda^2)\left(\text{tr}(\rho\hat{n}) - \mathbf{n}_\infty\right).$$

Combining the last equality with (63), we find that

$$-\frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) = \frac{\mu^2}{2}J_-(\rho) + \frac{\lambda^2}{2}J_+(\rho) - (\log\nu)(\mu^2 - \lambda^2)\mathrm{tr}(\rho\hat{n}) + \lambda^2\log\nu \ . \quad (64)$$

With (63) and (64), and setting $\zeta = \mu^2 - \lambda^2$, we obtain the desired equality. $\qquad\square$

The choice of $\zeta = \mu^2 - \lambda^2$ in the lemma is motivated by Gaussian states: in Appendix D we show that for any Gaussian state ρ the following inequality holds

$$\frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \leq -\zeta D(\rho\|\sigma_{\mu,\lambda}) \qquad \text{with} \qquad \zeta = \mu^2 - \lambda^2 > 0 \ .$$

Furthermore, for any $\epsilon > 0$ there exists a Gaussian state ρ such that

$$\frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \geq -(\zeta + \epsilon)D(\rho\|\sigma_{\mu,\lambda}) \ .$$

Let us now consider a specific example of a quantum Ornstein-Uhlenbeck process.

Example 1. *Consider $\mu^2 = 2$, $\lambda^2 = 1$. Then*

$$\frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}_{\sqrt{2},1}}(\rho)\|\sigma_{\sqrt{2},1}) \leq -D(\rho\|\sigma_{\sqrt{2},1})$$

for any state ρ with $S(\rho) \gtrsim 2.06$. In comparison, the entropy of the fixed point is $S(\sigma_{\sqrt{2},1}) = 2\log(2) \approx 1.39$.

In Lemma 8 we can bound $J_+(\rho) \geq 2$ [30, Eq. (43)], and the linear combination of $J_-(\rho)$ and $S(\rho)$ can be bounded by the result of de Palma et al. [24]. That is, for any state ρ with $S(\rho) \geq S_0$,

$$\frac{\mu^2}{2}J_-(\rho) + \zeta S(\rho) \geq \inf_{S \geq S_0} \left(\mu^2 f(S) + \zeta S\right) \ ,$$

where $f(S) = -g^{-1}(S)g'(g^{-1}(S))$, and $g(\mathbf{n}) = (\mathbf{n}+1)\log(\mathbf{n}+1) - \mathbf{n}\log(\mathbf{n})$. Substituting $S = g(\mathbf{n})$ gives

$$\frac{\mu^2}{2}J_-(\rho) + \zeta S(\rho) \geq \inf_{\mathbf{n} \geq g^{-1}(S_0)} \left(\mu^2(-\mathbf{n}g'(\mathbf{n})) + \zeta g(\mathbf{n})\right) =: F(S_0) \ .$$

That is, for any state with $S(\rho) \geq S_0$, we have

$$-\zeta D(\rho\|\sigma_{\mu,\lambda}) - \frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \geq F(S_0) + \lambda^2 + \lambda^2\log\nu + \zeta\log(1-\nu) \ .$$

With the choice $\mu^2 = 2$, $\lambda^2 = 1$, the function $F(S)$ is monotonically increasing for $S \gtrsim 0.5$. For $S_0 \gtrsim 2.06$, the right-hand side of the last inequality is non-negative. $\qquad\square$

The isoperimetric inequality for entropies (51) for one mode ($d = 1$) can be written as

$$-S(\rho) \leq \log\left(\frac{1}{4\pi e}J(\rho)\right) \ .$$

24

For any $A > 0$, using $\log x \le x - 1$, we get

$$
\begin{aligned}
\log\Big(\frac{1}{4\pi e}J(\rho)\Big) &= \log\Big(\frac{AJ(\rho)}{4\pi eA}\Big) \\
&= \log\Big(\frac{1}{4\pi eA}\Big) + \log\Big(AJ(\rho)\Big) \\
&\le AJ(\rho) - 2 - \log(4\pi A).
\end{aligned}
$$

Therefore

$$
-S(\rho) \le AJ(\rho) - (2 + \log(4\pi A)) \quad \text{for } A > 0 . \tag{65}
$$

Lemma 9 (Log-Sobolev inequality for the qOU semigroup)**.** *Let $\mu > \lambda > 0$ and $\zeta > 0$. Then*

$$
-\zeta D(\rho\|\sigma_{\mu,\lambda}) - \frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \ge \alpha_- J_-(\rho) + \alpha_+ J_+(\rho) + \gamma\,\mathsf{tr}(\rho\hat{n}) + \delta \tag{66}
$$

for all states ρ, where

$$
\begin{aligned}
\alpha_- &= \mu^2/2 - 2\pi A\zeta \\
\alpha_+ &= \lambda^2/2 - 2\pi A\zeta \\
\gamma &= (\log\nu)\left(\zeta - (\mu^2 - \lambda^2)\right) \\
\delta &= \zeta\left(\log(1-\nu) + 2 + \log(4\pi A)\right) + \lambda^2 \log\nu
\end{aligned}
$$

for any $A > 0$. Here $\nu = \frac{\lambda^2}{\mu^2}$, and $\sigma_{\mu,\lambda}$ is the fixed point of $\mathcal{L}_{\mu,\lambda}$. In particular, choosing $\zeta = \mu^2 - \lambda^2$ and $A = \frac{\lambda^2}{4\pi(\mu^2-\lambda^2)}$, we have

$$
-\zeta D(\rho\|\sigma_{\mu,\lambda}) - \frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \ge -\zeta\mathbf{n}\log(1 + 1/\mathbf{n}) + \delta , \tag{67}
$$

where $\mathbf{n} = \mathsf{tr}(\rho\hat{n})$.

Proof. Combining (61) and (65) leads to

$$
S(\rho) \ge -2\pi A J_-(\rho) - 2\pi A J_+(\rho) + 2 + \log(4\pi A) .
$$

Using this inequality in Lemma 8 we obtain the result (66).

For the second part of the statement, Theorem 5 shows that

$$
J_-(\rho) \ge J_-(\omega_{\mathbf{n}}) = -2\mathbf{n}\log(1 + 1/\mathbf{n}) , \tag{68}
$$

where $\omega_{\mathbf{n}}$ is the Gaussian thermal state with mean photon number $\mathbf{n}$. Eq. (67) then follows directly from the first part of the statement with the choice $A = \frac{\lambda^2}{4\pi(\mu^2-\lambda^2)}$, inequality (68), and the choice of $\zeta = \mu^2 - \lambda^2$. $\qquad\square$

Considering the same specific qOU process as in Example 1 we obtain the same rate of convergence to its fixed point, but now only for states with low mean photon number instead of large initial entropy as in Example 1.

Example 2. *Consider $\mu^2 = 2$ and $\lambda^2 = 1$. Then*

$$\frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}_{\sqrt{2},1}}(\rho)\|\sigma_{\sqrt{2},1}) \leq -D(\rho\|\sigma_{\sqrt{2},1})$$

for any state ρ with $\mathsf{tr}(\rho\hat{n}) \lesssim 0.67$. In comparison, the mean photon number of the fixed point is $\mathsf{tr}(\sigma_{\sqrt{2},1}\hat{n}) = 1$.

Choosing $\mu^2 = 2$ and $\lambda^2 = 1$ in Lemma 9, we obtain

$$-\zeta D(\rho\|\sigma_{\mu,\lambda}) - \frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \geq -\mathbf{n}\log(1+1/\mathbf{n}) + 2 - 2\log(2) \ .$$

The right-hand side is monotonically decreasing and for $\mathbf{n} \lesssim 0.67$, it is non-negative. $\qquad\square$

7   Discussion

We have established new information-theoretic inequalities for bosonic quantum systems. Our inequalities are motivated by and directly generalize well-known existing results concerning the sum of two real-valued random variables. They also complement recent results concerning the "addition" of two bosonic quantum states by means of a beamsplitter: we consider a hybrid operation which amounts to a certain way of combining a classical probability distribution on phase space with a quantum state. Mathematically, our work thus makes progress towards a unified view of three types of convolution operations: the convolution of two classical probability density functions, of two quantum Wigner functions, and of a pair consisting of a classical probability density function and a quantum Wigner function. Operationally, our results extend entropic characterizations of classical additive noise and quantum additive noise to so-called classical noise in bosonic systems. Indeed, the proofs of our main inequalities, which include hybrid versions of the Fisher information and entropy power inequalities, are formally very similar to existing proofs in the fully classical as well as fully quantum settings.

The consideration of the hybrid classical-quantum convolution operation (2) brings additional advantages, however: it allows for the study of infinitesimal Gaussian perturbations to a given quantum state. In contrast, existing fully quantum entropy power inequalities are not directly amenable to such arguments (at least not in an obvious way) since basic uncertainty relations prevent us from making sense of e.g., a Gaussian state with infinitesimal variance. Mirroring the derivation of the isoperimetric inequality from the Brunn-Minkowski inequality (where a given set is perturbed by adding an infinitesimally small ball), we obtain a quantum isoperimetric inequality relating Fisher information and entropy power. A striking simple consequence of the latter is the statement that the entropy power is a concave function of time for the quantum heat diffusion semigroup: again, this provides a quantum generalization of a fundamental result about the classical heat equation.

Let us conclude by mentioning a few potential directions for future work, as well as some open problems. For concreteness and simplicity, we have considered the simplest non-trivial definition of a hybrid convolution operation defined on multiple modes, as studied in [11]. One may generalize our convolution (2) and the associated results by considering additional (linear) operations along

the lines of [14]. From the point of view of information theory, it is also interesting to examine the implications of our results for the capacity of the classical noise channels similarly to [18].

On a more speculative side, one may wonder whether alternative quantum generalizations of the results considered here exist, especially related to the conjectured photon number inequality by Guha, Erkmen and Shapiro [41]. These authors (and subsequent work such as [21]) suggest replacing the quantum entropy power $e^{S(\rho)}$ by the arguably more natural expression $g^{-1}(S(\rho))$. This is the mean photon number of a Gaussian state with identical entropy as ρ. It appears that at least for our isoperimetric inequality, such a generalization would require more than a naïve substitution as there is no meaningful lower bound on the product $g^{-1}(S(\rho))J(\rho)$ even for Gaussian states. While this may be considered as additional mathematical justification for our formulation of these inequalities, we believe that progress in this direction could be helpful in resolving, e.g., our conjecture concerning the convergence rate to the fixed point of the quantum Ornstein-Uhlenbeck (qOU) semigroup. For the latter problem, apart from proving our conjecture, it would also be interesting to obtain multi-mode generalizations. This concerns, in particular, the entropy production rates for the qOU semigroup. Here, a resolution of the conjecture of [24] for the multi-mode attenuator would likely provide important insights.

Finally, we mention some challenging mathematical problems resulting from our work. For example, while our isoperimetric inequality is tight for Gaussian states, necessary and sufficient conditions for equality in most of our statements are currently unknown. Finally, a rigorous discussion of the family of states for which the de Bruijn identity (24) holds, possibly using the framework of Schwartz operators [36], would be an interesting task for future work.

8 Remark

After posting our paper to the arxiv, we were made aware of concurrent related work by Rouzé, Datta, and Pautrat. Their paper has now been made available [42].

Acknowledgements

A   A Log-Sobolev inequality and the classical Ornstein-Uhlenbeck process

In this appendix we discuss known classical results for the reader's convenience: we briefly review the relationship between the isoperimetric inequality for classical Fisher information, the Log-Sobolev inequality, and the rate of convergence to the fixed point for the classical Ornstein-Uhlenbeck semigroup. These arguments were given by Carlen [25] for a particular element of the two-parameter family of Ornstein-Uhlenbeck processes. Here we specialize to real-valued random variables, but allow arbitrary parameters in order to illustrate the parallels to the qOU semigroup. We also explicitly discuss the convergence to the fixed point, which is only implicit in [25] but appears to be folklore.

Let f_0 be a probability density on $\mathbb{R}$ of a real-valued random variable X. The classical Ornstein-Uhlenbeck (cOU) process, for $\theta > 0$ and $\sigma > 0$ is given by the Fokker-Planck equation

$$\frac{\partial f}{\partial t} = \theta \frac{\partial}{\partial x} [xf] + \frac{\sigma^2}{2} \frac{\partial^2 f}{\partial x^2} = \mathcal{A}_{\theta,\sigma^2}(f) \ .$$

(Carlen considers the case where $\theta = 1$ and $\sigma^2 = 1/\pi$.) The solution to this equation can be written (in terms of random variables) as

$$X_t = e^{-\theta t} X_0 + \frac{\sigma}{\sqrt{2\theta}} \sqrt{1 - e^{-2\theta t}} Z \ , \tag{69}$$

where $Z \sim \mathcal{N}(0,1)$ is an independent centered Gaussian random variable with unit variance. The stationary solution f_∞ therefore is a centered Gaussian distribution with variance $\sigma^2/(2\theta)$. In particular, (69) implies that the second moments satisfy

$$\mathbb{E}[X_t^2] = e^{-2\theta t} \mathbb{E}[X_0^2] + (1 - e^{-2\theta t}) \frac{\sigma^2}{2\theta} \ . \tag{70}$$

We use that the relative entropy between a random variable X and a centered normal variable $Z_{\sigma^2} \sim \mathcal{N}(0, \sigma^2)$ is given by $D(X \| Z_{\sigma^2}) = -H(X) + \frac{1}{2} \log 2\pi\sigma^2 + \frac{1}{2\sigma^2} \mathbb{E}[X^2]$ as can be verified easily. In particular, the relative entropy between the solution at time t and the fixed point $Z_{\sigma^2/(2\theta)}$ is given by

$$D(X_t \| Z_{\sigma^2/(2\theta)}) = -H(X_t) + \frac{1}{2} \log \pi\sigma^2/\theta + \frac{\theta}{\sigma^2} \mathbb{E}[X_t^2] \ . \tag{71}$$

Furthermore, we have the following de Bruijn-type identity:

Lemma 10. *Let $\{X_t\}_{t \geq 0}$ be of the form (69), i.e., a solution to the cOU process. Let $J(X) = \int \frac{|\frac{\partial f(x)}{\partial x}|^2}{f(x)} dx$ denote the Fisher information of a random variable X with distribution function f. Then*

$$\frac{d}{dt}\Big|_{t=0} H(X_t) = \frac{\sigma^2}{2} J(X_0) - \theta \ . \tag{72}$$

Observe that the second summand essentially stems from the fact that entropies transform very simply under rescaling of random variables, namely

$$H(e^{-\theta t}X) = H(X) - \theta t . \tag{73}$$

Eq. (73) significantly simplifies the analysis. Such a property does not hold in the quantum case: as a consequence, we do not have a simple expression in terms of $J(X_0)$ only.

Proof. The derivative of the entropy along a semigroup with generator $\mathcal{A}$ is given by the expression $\frac{d}{dt}\big|_{t=0}H(e^{t\mathcal{A}}(f)) = -\int \mathcal{A}(f)(x)\log f(x)dx$. Using this fact gives

$$\frac{d}{dt}\bigg|_{t=0}H(X_t) = -\theta \int \left(\frac{\partial}{\partial x}[xf(x)]\right)\log f(x)dx - \frac{\sigma^2}{2}\int \frac{\partial^2 f(x)}{\partial x^2}\log f(x)dx .$$

Denoting $f'(x) = \frac{\partial}{\partial x}f(x)$, we obtain

$$\int (xf)'\log f dx = -\int (xf)f'/f dx = -\int xf'dx = \int f dx = 1 ,$$

where we have used partial integration and the fact that boundary terms vanish twice. Similarly, we have

$$\int f''\log f dx = -\int (f')^2/f dx$$

by partial integration. Combining these statements gives the claim. $\qquad\square$

Following [25], we can write the isoperimetric inequality $1/N(X) \leq \frac{J(X)}{2\pi e}$ as

$$-H(X) \leq \frac{1}{2}\log\left(\frac{J(X)}{2\pi e}\right)$$
$$= \frac{1}{2}\log\left(\frac{J(X)}{2\pi}\right) - \frac{1}{2} .$$

In particular, using $\log x \leq x - 1$, we get

$$\log\left(\frac{J(X)}{2\pi e}\right) = \log\left(\frac{J(X)\cdot A}{2\pi e A}\right)$$
$$= \log\frac{1}{2\pi e A} + \log(AJ(X))$$
$$\leq AJ(X) - 2 - \log 2\pi A , \tag{74}$$

for any $A > 0$. Using inequality (74), it is straightforward to show the following.

Theorem 9 (Fast convergence of the cOU semigroup [25]). *Let $\{X_t\}_{t\geq 0}$ be of the form (69), i.e., a solution to the cOU process with parameters $\theta > 0, \sigma > 0$. Then*

$$\frac{d}{dt}\bigg|_{t=0}D(X_t\|Z_{\sigma^2/(2\theta)}) \leq -2\theta D(X_0\|Z_{\sigma^2/(2\theta)}) .$$

29

Note that because we are considering a semigroup, this result immediately implies that

$$D(X_t \| Z_{\sigma^2/(2\theta)}) \leq e^{-2\theta t} D(X_0 \| Z_{\sigma^2/(2\theta)}) \qquad \text{for all } t \geq 0 .$$

Also, this result is tight: if $X_0 \sim \mathcal{N}(0, \sigma_{X_0}^2)$ is a centered Gaussian random variable with variance $\sigma_{X_0}^2$, then (69) implies $X_t \sim \mathcal{N}(0, \sigma_t^2)$ where the variance of X_t is

$$\sigma_t^2 = e^{-2\theta t} \sigma_{X_0}^2 + \frac{\sigma^2}{2\theta}(1 - e^{-2\theta t}) .$$

Inserting into (71), using $H(X_t) = \frac{1}{2}(1 + \log(2\pi\sigma_t^2))$ yields

$$D(X_t \| Z_{\sigma^2/(2\theta)}) = -\frac{1}{2}(1 + \log(2\pi\sigma_t^2)) + \frac{1}{2}\log \pi\sigma^2/\theta + \frac{\theta}{\sigma^2}\sigma_t^2 .$$

In particular, we obtain

$$\lim_{\sigma_{X_0}^2 \to \infty} \left(\frac{d}{dt}\Big|_{t=0} D(X_t \| Z_{\sigma^2/(2\theta)}) \right) / D(X_0 \| Z_{\sigma^2/(2\theta)}) = -2\theta .$$

Proof. According to Eqs. (70), (71) and (72), we have

$$\frac{d}{dt}\Big|_{t=0} D(X_t \| Z_{\sigma^2/(2\theta)}) = -\frac{\sigma^2}{2} J(X_0) - \frac{2\theta^2}{\sigma^2}\mathbb{E}[X^2] + 2\theta . \tag{75}$$

Combining (74) with (71) yields

$$D(X_0 \| Z_{\sigma^2/(2\theta)}) \leq \frac{AJ(X)}{2} - (1 + \frac{1}{2}\log 2\pi A) + \frac{1}{2}\log \pi\sigma^2/\theta + \theta/\sigma^2 \mathbb{E}[X_0^2] . \tag{76}$$

Combining (76) with (75) yields

$$-2\theta D(X_0 \| Z_{\sigma^2/(2\theta)}) - \frac{d}{dt}\Big|_{t=0} D(X_t \| Z_{\sigma^2/(2\theta)}) \geq \left(-\theta A + \frac{\sigma^2}{2} \right) J(X_0)$$

$$+ 2\theta \left[\left(1 + \frac{1}{2}\log 2\pi A - \frac{1}{2}\log \pi\sigma^2/\theta \right) - 1 \right] .$$

The claim then follows by choosing $A = \frac{\sigma^2}{2\theta}$. □

B   Tightness of the quantum Fisher information isoperimetric inequality

Consider a one-mode Gaussian thermal state $\omega_{\mathbf{n}}$ with mean photon number $\mathbf{n} > 0$. Its entropy is

$$S(\omega_{\mathbf{n}}) = g(\mathbf{n}) = (\mathbf{n} + 1)\log(\mathbf{n} + 1) - \mathbf{n}\log \mathbf{n} . \tag{77}$$

Under the diffusion semigroup, the state $\omega_{\mathbf{n}}$ evolves as

$$e^{t\mathcal{L}_{\text{heat}}}(\omega_{\mathbf{n}}) = \omega_{\mathbf{n}_t} , \qquad \text{where} \qquad \mathbf{n}_t = \mathbf{n} + 2\pi t .$$

In particular, by the de Bruijn identity

$$J(\omega_{\mathbf{n}}) = 2\frac{d}{dt}S(e^{t\mathcal{L}_{\text{heat}}}(\omega_{\mathbf{n}}))\Big|_{t=0} = 2g'(\mathbf{n})\mathbf{n}_t'\Big|_{t=0} = 4\pi\log\left(\frac{\mathbf{n}+1}{\mathbf{n}}\right) . \tag{78}$$

Also,

$$J(e^{t\mathcal{L}_{\text{heat}}}(\omega_{\mathbf{n}})) = 4\pi\log\left(1 + \frac{1}{\mathbf{n}+2\pi t}\right) .$$

Calculating the right-hand side of the quantum Fisher information isoperimetric inequality (40), we obtain

$$\frac{d}{dt}\Big|_{t=0}\left[\frac{1}{2}J(e^{t\mathcal{L}_{\text{heat}}}(\omega_{\mathbf{n}}))\right]^{-1} = \frac{1}{\mathbf{n}(\mathbf{n}+1)}\log^{-2}\left(1 + \frac{1}{\mathbf{n}}\right) \to 1 \qquad \text{as } \mathbf{n} \to \infty .$$

C   Tightness of the isoperimetric inequality

Consider a one-mode Gaussian thermal state $\omega_{\mathbf{n}}$ with mean photon number $\mathbf{n}$. From Eq. (78) we know that

$$J(\omega_{\mathbf{n}}) = 4\pi\log\left(\frac{\mathbf{n}+1}{\mathbf{n}}\right) ,$$

and that the entropy power of $\omega_{\mathbf{n}}$ is given by (cf. (77))

$$N(\omega_{\mathbf{n}}) = \exp(S(\omega_{\mathbf{n}})/1) = \frac{(\mathbf{n}+1)^{\mathbf{n}+1}}{\mathbf{n}^{\mathbf{n}}} .$$

Combining these two expressions, we see that the left-hand side of (51) is

$$J(\omega_{\mathbf{n}})N(\omega_{\mathbf{n}}) = 4\pi\left(\frac{\mathbf{n}+1}{\mathbf{n}}\right)^{\mathbf{n}}\log\left(\frac{\mathbf{n}+1}{\mathbf{n}}\right)^{\mathbf{n}+1} \to 4\pi e \qquad \text{for } \mathbf{n} \to \infty .$$

D   On the convergence rate for Gaussian initial states

In this section we show that for a one-mode Gaussian state ρ the following inequality holds:

$$\frac{d}{dt}\Big|_{t=0}D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \le -\zeta D(\rho\|\sigma_{\mu,\lambda}) \qquad \text{with } \zeta = \mu^2 - \lambda^2 > 0 . \tag{79}$$

Furthermore, this statement is optimal: for any $\epsilon > 0$ there exists a Gaussian state ρ such that

$$\frac{d}{dt}\Big|_{t=0}D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \ge -(\zeta + \epsilon)D(\rho\|\sigma_{\mu,\lambda}) .$$

First, we note that the entropy of a Gaussian state does not depend on its first moments. Hence it follows that for a Gaussian state the right-hand side of Eq. (62) does not depend on its first moments, and it is suffices to consider centered Gaussian states.

Calculating the right-hand side of (62), we focus on calculating $J_{\pm}(\rho)$. Recall that the covariance matrix M of a centered state ρ is defined as $M_{jk} = \mathsf{tr}(\rho\{R_j, R_k\})$.

Lemma 11. *Let ρ be a one-mode centered Gaussian state with mean-photon number $\mathbf{n}$ and covariance matrix given by $M = \kappa S^T S$, where $\kappa = 2\mathbf{n} + 1$, and $S = O_1 \begin{pmatrix} z & 0 \\ 0 & 1/z \end{pmatrix} O_2^T$ with $O_i \in Sp(2) \cap O(2)$ and $z \geq 1$. Denote*

$$J_{\pm}(\rho) = 2\frac{d}{dt}S\left(e^{t\mathcal{L}_{\pm}}(\rho)\right)\Big|_{t=0} .$$

Then we have

$$J_{\pm}(\rho) = \left(\frac{1}{2}\left(1/z^2 + z^2\right) \pm \kappa\right)\left(\log\left(\frac{\kappa+1}{2}\right) - \log\left(\frac{\kappa-1}{2}\right)\right) . \tag{80}$$

Proof. The covariance matrix of the time-evolved state $e^{t\mathcal{L}_{\pm}}(\rho)$ is

$$M_{\pm}(t) = c_1^{\pm}(t)M(0) + c_2^{\pm}(t)\,\mathrm{id} ,$$

where

$$\begin{aligned}
c_1^-(t) &= e^{-t} &&\text{and} & c_2^-(t) &= 1 - e^{-t} , & (81) \\
c_1^+(t) &= e^t &&\text{and} & c_2^+(t) &= e^t - 1 . & (82)
\end{aligned}$$

Therefore writing $S[M]$ for the entropy of a Gaussian state with covariance matrix M, we have

$$\begin{aligned}
S\left[M_{\pm}(t)\right] &= S\left[c_1^{\pm}(t)\kappa S^T S + c_2^{\pm}(t)\,\mathrm{id}\right] = S\left[c_1^{\pm}(t)\kappa O_2^T K^T O_1^T O_1 K O_2 + c_2^{\pm}(t)O_2^T O_2\right] \\
&= S\left[c_1^{\pm}(t)\kappa K^T O_1^T O_1 K + c_2^{\pm}(t)\,\mathrm{id}\right] \\
&= S\left[\begin{pmatrix} c_1^{\pm}(t)\kappa z^2 + c_2^{\pm}(t) & 0 \\ 0 & c_1^{\pm}(t)\kappa/z^2 + c_2^{\pm}(t) \end{pmatrix}\right] . \tag{83}
\end{aligned}$$

The symplectic eigenvalue of the matrix argument in (83) is the square root of its determinant:

$$\kappa_{\pm}(t) = \sqrt{\left(c_1^{\pm}(t)\kappa + c_2^{\pm}(t)z^2\right)\left(c_1^{\pm}(t)\kappa + c_2^{\pm}(t)/z^2\right)} . \tag{84}$$

The entropy of the time evolved state is

$$S(e^{t\mathcal{L}_{\pm}}(\rho)) = g(\mathbf{n}(\kappa_{\pm}(t))) ,$$

where $g(x) = (x + 1)\log(x + 1) - x\log x$ and $\mathbf{n}(\kappa) = \frac{1}{2}(\kappa - 1)$. By the chain rule, we have that

$$\frac{d}{dt}S\left(e^{t\mathcal{L}_{\pm}}(\rho)\right) = \frac{1}{2}g'(\mathbf{n}(\kappa_{\pm}(t)))\,\kappa_{\pm}'(t) .$$

Since $g'(x) = \log(x + 1) - \log x$, we only need to find $\kappa_{\pm}'(t)$.

Combining (84) and (81) we obtain

$$\begin{aligned}
\kappa_-(t) &= \sqrt{\left(e^{-t}\kappa + \left(1 - e^{-t}\right)z^2\right)\left(e^{-t}\kappa + \left(1 - e^{-t}\right)/z^2\right)} \\
&= \kappa + t\left(\frac{1}{2}\left(z^2 + 1/z^2\right) - \kappa\right) + O\left(t^2\right) .
\end{aligned}$$

Therefore $\kappa_-'(t)|_{t=0} = \frac{1}{2}\left(1/z^2 + z^2\right) - \kappa$ and finally

$$J_-(\rho) = \left(\frac{1}{2}\left(1/z^2 + z^2\right) - \kappa\right)\log\frac{\kappa+1}{\kappa-1} \ .$$

Similarly, using (84) together with (82), we obtain

$$\kappa_+(t) = \sqrt{\left(e^t\kappa + \left(e^t - 1\right)z^2\right)\left(e^t\kappa + \left(e^t - 1\right)/z^2\right)}$$

$$= \kappa + t\left(\frac{1}{2}\left(z^2 + 1/z^2\right) + \kappa\right) + O\left(t^2\right) \ .$$

Thus $\kappa_+'(t)|_{t=0} = \frac{1}{2}\left(1/z^2 + z^2\right) + \kappa$ and

$$J_+(\rho) = \left(\frac{1}{2}\left(1/z^2 + z^2\right) + \kappa\right)\log\frac{\kappa+1}{\kappa-1} \ .$$

$\square$

From Lemma 8 it is clear that we are interested in minimizing $J_\pm(\rho)$. Both expressions in (80) have a minimum at $z = 1$. Therefore, with $\zeta = \mu^2 - \lambda^2$ and $\mathbf{n} = \frac{\kappa-1}{2}$, from Lemma 8 we obtain

$$-\zeta D(\rho\|\sigma_{\mu,\lambda}) - \frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \geq \mu^2\log(\mathbf{n}+1) - \lambda^2\log\mathbf{n} + \lambda^2\log\lambda^2$$

$$- \mu^2\log\mu^2 + (\mu^2 - \lambda^2)\log(\mu^2 - \lambda^2)$$

$$=: h_{\mu,\lambda}(\mathbf{n}) \ .$$

For fixed $\mu^2 > \lambda^2$, the function $h_{\mu,\lambda}$ satisfies $\lim_{\mathbf{n}\to 0} h_{\mu,\lambda}(\mathbf{n}) = \lim_{\mathbf{n}\to\infty} h_{\mu,\lambda}(\mathbf{n}) = \infty$. Since $h_{\mu,\lambda}$ is differentiable (in fact smooth) for $\mathbf{n} > 0$, we can find the global minimum by finding the zeros of the derivative

$$\frac{d}{d\mathbf{n}} h_{\mu,\lambda}(\mathbf{n}) = \frac{\mu^2}{\mathbf{n}+1} - \frac{\lambda^2}{\mathbf{n}} = 0 \ .$$

The only solution is $\mathbf{n} = \frac{\lambda^2}{\mu^2 - \lambda^2}$, and since

$$\frac{d^2}{d\mathbf{n}^2}\Big|_{\mathbf{n}=\frac{\lambda^2}{\mu^2-\lambda^2}} h_{\mu,\lambda}(\mathbf{n}) = (\mu^2 - \lambda^2)\left(\frac{1}{\lambda^2} - \frac{1}{\mu^2}\right) > 0 \ ,$$

it is the minimum. The value of the minimum is $h_{\mu,\lambda}\left(\frac{\lambda^2}{\mu^2-\lambda^2}\right) = 0$, hence it follows that $h_{\mu,\lambda}(\mathbf{n}) \geq 0$ for all $\mathbf{n} > 0$ and we have

$$\frac{d}{dt}\Big|_{t=0} D(e^{t\mathcal{L}_{\mu,\lambda}}(\rho)\|\sigma_{\mu,\lambda}) \leq -\zeta D(\rho\|\sigma_{\mu,\lambda}) \ .$$

Moreover, the last inequality becomes equality for $\rho = \sigma_{\mu,\lambda} = \omega_{\mathbf{n}_\infty}$, with $\mathbf{n}_\infty = \frac{\lambda^2}{\mu^2 - \lambda^2}$.

33

It remains to show that $\zeta = \mu^2 - \lambda^2$ is optimal. Let $\epsilon > 0$ and $\zeta' = \zeta + \epsilon$. Then for the Gaussian thermal state $\omega_{\mathbf{n}}$ we have

$$-\zeta' D(\omega_{\mathbf{n}} \| \sigma_{\mu,\lambda}) - \frac{d}{dt}\bigg|_{t=0} D(e^{t\mathcal{L}_{\mu,\lambda}}(\omega_{\mathbf{n}}) \| \sigma_{\mu,\lambda}) = h_{\mu,\lambda}(\mathbf{n}) + \epsilon(\mathbf{n}+1)\log(\mathbf{n}+1) - \epsilon\,\mathbf{n}\log\mathbf{n}$$

$$+ \epsilon\,\mathbf{n}\log\left(\frac{\lambda^2}{\mu^2}\right) + \epsilon\log\left(1 - \frac{\lambda^2}{\mu^2}\right)$$

$$= \log\left(\left(\frac{\mathbf{n}+1}{\mathbf{n}}\right)^{\mu^2+\epsilon(\mathbf{n}+1)} \mathbf{n}^{\mu^2-\lambda^2+\epsilon}\left(\frac{\lambda^2}{\mu^2}\right)^{\epsilon\mathbf{n}}\right)$$

$$+ c(\mu,\lambda) \to -\infty \qquad \text{for } \mathbf{n} \to \infty\,,$$

where $c(\mu,\lambda) = \log\left((\mu^2)^{-\mu^2-\epsilon}(\lambda^2)^{\lambda^2}(\mu^2-\lambda^2)^{\mu^2-\lambda^2+\epsilon}\right)$. Therefore, for any $\epsilon > 0$, there exists $\mathbf{n}$ such that

$$\frac{d}{dt}\bigg|_{t=0} D(e^{t\mathcal{L}_{\mu,\lambda}}(\omega_{\mathbf{n}}) \| \sigma_{\mu,\lambda}) > -(\zeta + \epsilon) D(\omega_{\mathbf{n}} \| \sigma_{\mu,\lambda})\,.$$

This shows that the constant $\zeta = \mu^2 - \lambda^2$ is optimal in the inequality (79).

References

[1] W. Beckner, "Inequalities in Fourier analysis," *The Annals of Mathematics*, vol. 102, no. 1, pp. pp. 159–182, 1975.

[2] H. Brascamp and E. Lieb, "Best constants in Young's inequality, its converse and its generalization to more than three functions," *Adv. in Math.*, vol. 20, pp. 151–172, 1976.

[3] A. J. Stam, "Some inequalities satisfied by the quantities of information of Fisher and Shannon," *Information and Control*, vol. 2, no. 2, pp. 101 – 112, 1959.

[4] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 623–656, October 1948.

[5] N. Blachman, "The convolution inequality for entropy powers," *IEEE Transactions on Information Theory*, vol. 11, pp. 267 – 271, Apr 1965.

[6] A. R. Barron, "Entropy and the central limit theorem," *Ann. Probab.*, vol. 14, no. 1, pp. 336–342, 1986.

[7] M. H. M. Costa and T. M. Cover, "On the similarity of the entropy power inequality and the Brunn–Minkowski inequality," *IEEE Transactions on Information Theory*, vol. 30, no. 6, pp. 837–839, 1984.

[8] S. J. Szarek and D. Voiculescu, *Shannon's entropy power inequality via restricted Minkowski sums*, vol. 1745 of *Lecture Notes in Math.*, pp. 257–262. Springer Berlin Heidelberg, 2000.

[9] A. Dembo, T. M. Cover, and J. A. Thomas, "Information theoretic inequalities," *IEEE Transactions on Information Theory*, vol. 37, pp. 1501 –1518, Nov 1991.

[10] F. Barthe, "The Brunn–Minkowski theorem and related geometric and functional inequalities," in *Proceedings of the International Congress of Mathematicians*, (Madrid), p. pp. 1529–1546, European Mathematical Society, August 2006.

[11] R. Werner, "Quantum harmonic analysis on phase space," *Journal of Mathematical Physics*, vol. 25, no. 5, pp. 1404–1411, 1984.

[12] E. A. Carlen and E. H. Lieb, "Brascamp-Lieb Inequalities for Non-Commutative Integration," *Documenta Mathematica*, vol. 13, pp. 553–584, 2008.

[13] R. König and G. Smith, "The entropy power inequality for quantum systems," *IEEE Transactions on Information Theory*, vol. 60, pp. 1536–1548, March 2014.

[14] G. D. Palma, A. Mari, S. Lloyd, and V. Giovannetti, "Multimode quantum entropy power inequality," *Phys. Rev. A*, vol. 91, p. 032320, Mar 2015.

[15] E. H. Lieb, "Proof of an entropy conjecture of Wehrl," *Comm. Math. Phys.*, vol. 62, no. 1, pp. 35–41, 1978.

[16] S. Verdu and D. Guo, "A simple proof of the entropy-power inequality," *IEEE Transactions on Information Theory*, vol. 52, pp. 2165 –2166, May 2006.

[17] R. König, "The conditional entropy power inequality for Gaussian quantum states," *Journal of Mathematical Physics*, vol. 56, p. 022201, February 2015.

[18] R. König and G. Smith, "Limits on classical communication from quantum entropy power inequalities," *Nature Photonics*, vol. 7, pp. 142–146, February 2012.

[19] K. Audenaert, N. Datta, and M. Ozols, "Entropy power inequalities for qudits," *Journal of Mathematical Physics*, vol. 57, p. 052202, 2016.

[20] E. A. Carlen, E. H. Lieb, and M. Loss, "On a quantum entropy power inequality of Audenaert, Datta and Ozols," *Journal of Mathematical Physics*, vol. 57, p. 062203, 2016.

[21] S. Guha, J. H. Shapiro, and R. G.-P. Sanchez, "Thinning, photonic beam splitting, and a general discrete entropy power inequality," in *IEEE International Symposium on Information Theory (ISIT), Barcelona, 2016*, pp. 705–709, 2016.

[22] M. Costa, "A new entropy power inequality," *IEEE Transactions on Information Theory*, vol. 31, pp. 751–760, Nov 1985.

[23] C. Villani, "A short proof of the concavity of entropy power," *IEEE Transactions on Information Theory*, vol. 46, pp. 1695–1696, July 2000.

[24] G. D. Palma, D. Trevisan, and V. Giovannetti, "Gaussian states minimize the output entropy of the one-mode quantum attenuator," *IEEE Transactions on Information Theory*, vol. 63, pp. 728–737, 2017.

[25] E. A. Carlen, "Superadditivity of Fisher's information and logarithmic Sobolev inequalities," *Journal of Functional Analysis*, vol. 101, no. 1, pp. 194 – 211, 1991.

[26] R. Carbone and E. Sasso, "Hypercontractivity for a quantum Ornstein–Uhlenbeck semigroup," *Probability Theory and Related Fields*, vol. 140, no. 3-4, pp. 505–522, 2008.

[27] R. Olkiewicz and B. Zegarlinski, "Hypercontractivity in noncommutative lpspaces," *Journal of Functional Analysis*, vol. 161, no. 1, pp. 246 – 285, 1999.

[28] M. J. Kastoryano and K. Temme, "Quantum logarithmic Sobolev inequalities and rapid mixing," *Journal of Mathematical Physics*, vol. 54, no. 5, 2013.

[29] A. Müller-Hermes, D. S. França, and M. M. Wolf, "Relative entropy convergence for depolarizing channels," *Journal of Mathematical Physics*, vol. 57, no. 2, p. 022202, 2016.

[30] F. Buscemi, S. Das, and M. M. Wilde, "Approximate reversibility in the context of entropy gain, information gain, and complete positivity," *Physical Review A*, vol. 93, p. 062314, 2016.

[31] M. Jabbour, R. Garcia-Patron, and N. Cerf, "Majorization preservation of gaussian bosonic channels," *New J. Phys.*, vol. 18, p. 073047, 2016.

[32] G. D. Palma, D. Trevisan, and V. Giovannetti, "Passive states optimize the output of bosonic Gaussian quantum channels," *IEEE Transactions on Information Theory*, vol. 62, pp. 2895–2906, 2016.

[33] R. Zamir, "A proof of the Fisher information inequality via a data processing argument," *Information Theory, IEEE Transactions on*, vol. 44, pp. 1246 –1250, May 1998.

[34] A. Dembo, "Simple proof of the concavity of the entropy power with respect to added Gaussian noise," *IEEE Transactions on Information Theory*, vol. 35, no. 4, pp. 887–888, 1989.

[35] D. Petz and C. Ghinea, "Introduction to quantum Fisher information," *QP - PQ: Quantum Probab. White Noise Anal.*, vol. 27, pp. 261–281, 2011.

[36] M. Keyl, J. Kiukas, and R. F. Werner, "Schwartz operators," *Rev. Math. Phys.*, p. 1630001, 2016.

[37] J. Eisert and M. M. Wolf, *Gaussian quantum channels*, pp. 23–42. Quantum Information with Continuous Variables of Atoms and Light, Imperial College Press, London, 2007.

[38] E. H. Lieb, "Convex trace functions and the Wigner-Yanase-Dyson conjecture," *Advances in Mathematics*, vol. 11, no. 3, pp. 267 – 288, 1973.

[39] R. König and G. Smith, "Erratum for "The entropy power inequality for quantum systems" , IEEE Trans. Inf. Th., vol. 60, no. 3, p. 1536-1548 (2014)," *IEEE Transactions on Information Theory*, vol. 62, pp. 4358–4359, May 2016.

[40] T. M. Cover and J. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.

[41] S. Guha, B. Erkmen, and J. Shapiro, "The entropy photon-number inequality and its consequences," in *Information Theory and Applications Workshop, 2008*, pp. 128–130, IEEE, 2008.

[42] N. Datta, Y. Pautrat, and C. Rouzé, "Contractivity properties of a quantum diffusion semigroup," *Journal of Mathematical Physics*, vol. 58, p. 012205, 2017.