ICFO–The Institute of Photonic Sciences
Castelldefels (Barcelona), Spain

Ph.D. thesis

Quantum information outside quantum information

Alejandro Pozas-Kerstjens

October 15, 2019

Thesis Advisor: Prof. Dr. Antonio Acín

ICFO
BIST Barcelona Institute of Science and Technology
UPC UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Cover picture: Artistic impression of Figure 1.1.
Author: Héctor Lobo Rodríguez

Acknowledgments

Si yo pudiera enumerar cuánto debo a mis antecesores y contemporáneos, no me quedaría mucho en propiedad

—*Johann Wolfgang von Goethe*

This thesis is the result of a four-year journey. Many people have contributed to it in one way or another, and I want to take this opportunity to thank them sincerely. This thesis belongs to them as well.

First and foremost, I am deeply grateful to Toni Acín for enabling the journey, for his mentoring and guidance, and for his trust and support. You have created a great environment by putting together excellent people at both the professional and the personal level. My only shame is that, in four full years, I have not been able to save you a single penalty kick.

Also, I will always be thankful to Frank Koppens and Mark Lundeberg, Rafael Pérez del Real and María Ángeles Hernández Vozmediano, Luis Garay, and Eduardo Martín-Martínez. Without any of you, the journey wouldn't have even started. I want to specially thank Eduardo for supervising my Master's thesis and for helping me, with passion and care, in my first steps as a researcher.

These four years have also comprised a physical journey, of which I feel blessed for having learned from the best. Thanks to Miguel Navascués for hosting me in Vienna, but most importantly for having patience with a young freshman and for beginning to shape this thesis. And to Elie Wolfe and Dennis Rosset for accepting to host me in Waterloo. Because of you, this was the most stimulating and productive period of my Ph.D.

ICFO is a great place for many reasons. The most important is its people, excellent in all aspects. I want to highlight the labor of María José Martínez, Manuela Furkert, Mery Gil, and María Jesús Moreno, without whom everything would simply collapse. Thank you for your patience, your constant readiness, and your hard work.

And speaking about people, I cannot be grateful enough to all the wonderful ones that shared part of the journey in the QIT group and its surroundings. The good thing is that it is a lot of people with which I have shared discussions, coffees, and games of futbolín. The bad thing is that it is a lot of names to remember, and I hope I don't forget anyone. Thanks to Alejandro, Alexia,

Arnau, Bogna, Boris, Christian, Chung-Yun, Dani, Dario, Elisa, Eric, Erik, Felix, Flavio, Flo, Gabi, Ivan, Jacopo, Janek, Joe, Karen, Binz, Leo, Markus, Martí, Matteo Uno, Matteo Due, Michał, Moha, Osvaldo, Paolo, Patrick, Peter, Senaida, Stefan, Victoria, Xavier, and Zahra. Also, thanks to Albert, Andreu, Giulia, Gorka, Iris, Joan, Jordi, Maciej, María, Miguel Ángel, Pei-Sheng, Przemek, Sarah and Simona.

A special shout-out goes to Alexia Salavrakos and Alejandro Máttar. Office 368 will always be the Alexes office. And to those who made a special impact: Senaida–You are the strongest person I have met. Thank you for caring when times were dark. Flavio–I love your work. I love the way you work. And I love your humor. Peter–You have an energy and determination that makes possible the impossible. Thanks to you I have done things I only dreamed of, and others I could never have imagined. You are a great inspiration. Moha–Every day you came to work was a day in which I had a laugh. Thanks for not giving up in convincing me for going for coffee. Felix–I have enjoyed each and every one of the conversations we have had, and I admire your will to always dig deeper. I am still amazed of how many things you know about.

Many thanks to all those who have been present despite the distance. To 31 Louisa et al. (*You're just the best. Period.*). To Mariana Carrillo, Perseas Christodoulidis, Tom O'Brien and Miroslav Rapčák (*This time I got your last name right*). A Sandra, Mónica y Quique (*Los cotilleos no saben de distancia*). Y a Kike, Fran y Lobo (*Que nadie nos quite las cañas*).

No puc oblidar-me de Vall d'Hebron i Electrons Tropicals. Gràcies pels bons moments, per la vostra alegria, i per aguantar un tipus que no ha deixat de jugar amb la samarreta del Reial Madrid.

Esta tesis, y el trabajo detrás de cada una de sus páginas, están dedicados a mis cinco pilares. A Rafael, Mari Paz y Guille. Gracias por enseñarme a vivir con fuerza, inteligencia, y valentía. A Trasto, por recordarme periódicamente lo que realmente es importante en esta vida. Y, sin duda, a Erika. Solo nosotros dos conocemos todo lo que contiene este Gracias.

Abstract

Quantum theory, as counter-intuitive as a theory can get, has turned out to make predictions of the physical world that match observations so precisely that it has been described as the most accurate physical theory ever devised. Viewing quantum entanglement, superposition and interference not as undesirable necessities but as interesting resources paved the way to the development of quantum information science. This area studies the processing, transmission and storage of information when one accounts for the fact that information is physical and subjected to the laws of nature that govern the systems it is encoded in. The development of the consequences of this idea, along with the great advances experienced in the control of individual quantum systems, has led to what is now known as the *second quantum revolution* or the era of *quantum technologies*, in which quantum information science has emerged as a fully-grown field of physics. As such, ideas and tools developed within the framework of quantum information theory begin to permeate to other fields of research.

This Ph.D. dissertation is devoted to the use of concepts and methods akin to the field of quantum information science in other areas of research. In the same way, it also considers how encoding information in quantum degrees of freedom may allow further development of well-established research fields and industries. That is, this thesis aims to the study of quantum information outside the field of quantum information. Four different areas are visited.

A first question posed is that of the role of quantum information in quantum field theory, with a focus in the quantum vacuum. It is known that the quantum vacuum contains entanglement, but it remains unknown whether this entanglement can be accessed and exploited in experimental setups. We give crucial steps in this direction by studying the extraction of vacuum entanglement in realistic models of light-matter interaction, and by giving strict mathematical conditions of general applicability that must be fulfilled for extraction to be possible at all.

Another field where quantum information methods can offer great insight is that of quantum thermodynamics, where the idealizations made in macroscopic thermodynamics break down. Making use of a quintessential framework of quantum information and quantum optics, we provide the first model of quantum thermodynamic interactions that does not make any of the assumptions per-

formed in macroscopic thermodynamics. We make use of this model to study the cyclic operation of a microscopic heat engine composed by a single particle reciprocating between two finite-size baths, and we study on the consequences of the removal of the macroscopic idealizations.

One more step down the stairs to applications in society, we analyze the impact that encoding information in quantum systems and processing it in quantum computers may have in the field of machine learning. A great desideratum in this area, largely obstructed by computational power, is that of explainable models, which not only make predictions but also provide information about the decision process that triggers them. We develop an algorithm to train neural networks using Bayesian techniques that makes use of entanglement and superposition to execute efficiently in quantum computers, in contrast with classical counterparts. Furthermore, we run it in state-of-the-art quantum computers with the aim of assessing the viability of realistic implementations.

Lastly, and encompassing all the above, we explore the notion of causality and cause-effect relations in quantum mechanics from an information-theoretic point of view. While it is known since the works of John S. Bell in the decade of 1960 that, for a same causal pattern, quantum systems can generate correlations between variables that are impossible to obtain employing only classical systems, there is an important lack of tools to study complex causal effects whenever a quantum behavior is expected. We fill this gap by providing general methods for the characterization of the quantum correlations achievable in complex causal patterns. Closing the circle, we make use of these tools to find phenomena of fundamental and experimental relevance back in quantum information.

List of publications

This thesis is based on the following publications:

- *Entanglement harvesting from the electromagnetic vacuum with hydrogenlike atoms*, Alejandro Pozas-Kerstjens, and Eduardo Martín-Martínez, Physical Review D **94**, 064074 (2016)

- *Degenerate detectors are unable to harvest spacelike entanglement*, Alejandro Pozas-Kerstjens, Jorma Louko, and Eduardo Martín-Martínez, Physical Review D **95**, 105009 (2017)

- *A quantum Otto engine with finite heat baths: energy, correlations, and degradation*, Alejandro Pozas-Kerstjens, Eric G. Brown, and Karen V. Hovhannisyan, New Journal of Physics **20**, 043034 (2018)

- *Bayesian deep learning on a quantum computer*, Zhikuan Zhao, Alejandro Pozas-Kerstjens, Patrick Rebentrost, and Peter Wittek, Quantum Machine Intelligence **1**, 41–51 (2019)

- *Bounding the sets of classical and quantum correlations in networks*, Alejandro Pozas-Kerstjens, Rafael Rabelo, Łukasz Rudnicki, Rafael Chaves, Daniel Cavalcanti, Miguel Navascués, and Antonio Acín, preprint at arXiv:1904.08943 (2019). Accepted in *Physical Review Letters*.

- *Quantum inflation: a general approach to quantum causal compatibility*, Elie Wolfe, Alejandro Pozas-Kerstjens, Matan Grinberg, Denis Rosset, Antonio Acín, and Miguel Navascués. Submitted to *Physical Review X*.

Other works of the author not included in this thesis are:

- *Efficient training of energy-based models through frustration reduction*, Alejandro Pozas-Kerstjens, Gorka Muñoz-Gil, Miguel Angel Garcia-March, Antonio Acín, Maciej Lewenstein, and Przemysław R. Grzybowski. Submitted to *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

List of computational contributions

This PhD thesis has given rise to the following computational tools:

- *gausianotto*: simulations of quantum thermodynamics using Gaussian quantum mechanics, Alejandro Pozas-Kerstjens, Eric G. Brown, and Karen V. Hovhannisyan, Zenodo 847182 (2017)
- *bayesian-dl-quantum*: implementation of the quantum linear systems algorithm in state-of-the-art quantum computers, Alejandro Pozas-Kerstjens, and Peter Wittek, GitLab apozas/bayesian-dl-quantum (2018)
- *quantum-networks-scalar-extension*: the scalar extension method for quantum causal compatibility, Alejandro Pozas-Kerstjens, Zenodo 2646262 (2019)

Other computational tools developed, not included in this thesis:

- *NGrad*: numerical gradient of multivariate functions, Alejandro Pozas-Kerstjens, Wolfram Library Archive 9547 (2016)
- *ebm-torch*: energy-based machine learning models in PyTorch, Alejandro Pozas-Kerstjens, GitHub apozas/ebm-torch (2018)
- *qaoa-color*: constrained optimization in QAOA, Alejandro Pozas-Kerstjens, GitHub apozas/qaoa-color (2019)

Contents

1 | Introduction

Two discoveries at the beginning of the XX century shook the pillars of science by setting fundamental limits on our ability to describe nature. On one hand, Kurt Gödel proved in 1931 that mathematics was incomplete, in the sense that in any consistent theory of arithmetics there exist propositions which cannot be proven nor refuted [Göd31]. On the other, five years earlier, physicist Max Born had proven a similar statement about the physical world: despite having a complete characterization of a physical system, there exist measurements on it whose outcomes could never be predicted with certainty [Bor26]. This "incompleteness" of the physical world was further demonstrated in 1927 with Werner Heisenberg's uncertainty principle, by which two properties of a same system could not be determined simultaneously with arbitrary precision [Hei27].

Born's and Heisenberg's results, although deeply shocking, are just two of many counter-intuitive features of the theory of quantum mechanics. This theory, that aims to describe the interactions of the fundamental constituents of matter, initiated its development in 1900 in an attempt to accurately describe the radiation spectrum of a black body. From the very beginning, the assumptions required and the consequences derived from them pointed to an extremely challenging vision of reality, where properties of systems were not defined before measurements were conducted and particles and waves were the same entity. This made the theory face hard criticisms[1], and to date, 120 years after its initial developments, its interpretation is still object of debate[2].

Despite this, quantum mechanics has turned out to be the most successful theory ever devised, by achieving the most accurate predictions of physical phenomena. This is why, nicely summarized in David Mermin's *Shut up and calculate!*[3], generations of physicists began using quantum properties such as entanglement, superposition and interference to explain fundamental observa-

[1] After solving the black-body radiation problem, Max Planck said to have formulated the quantum hypothesis in what he called an "act of desperation".

[2] The most prominent interpretations to date include the Copenhagen interpretation, de Broglie-Bohm's pilot wave theory, Everett's many-worlds interpretation, or Fuchs and Schack's quantum Bayesianism.

[3] N. David Mermin, Physics Today, April 1989, page 9, DOI:10.1063/1.2810963

tions and make precise predictions in particle, atomic, nuclear and solid-state physics, without strongly focusing on the ontological implications of the theory.

An important realization that took place in the decade of 1980 was the fact that information was not an ethereal concept but a physical entity, subjected to the laws of nature that govern the systems where it is stored and processed. Scientists began to wonder about how information encoded in quantum particles would be affected by the nonclassical phenomena. Quantum effects, which challenged the conceptions of nature and reality, became powerful resources to improve information-processing technologies. Examples of this are Bennett and Brassard's protocol to securely establish cryptographic keys [BB84], Bennett and Weisner's method to transmit increased amounts of information [BW92], or Shor's algorithm for factoring numbers [Sho94].

Nowadays, developments both in the fundamentals of the theory and in its applications have led to quantum information science gaining the reputation of a field in its own. As a theory, it has pushed quantum mechanics to its boundary establishing fundamental limits on its information-processing power, and to go beyond them by exploring post-quantum theories and signatures that could, in principle, be observed experimentally [NGHA15]. On the practical side, it has both confirmed deep insights about physical reality, such as its nonlocal nature [HBD$^+$15, GVW$^+$15, SMSC$^+$15], and has led to the development of information-processing systems that will have—and are already having—a very strong impact in society [Rib15, PAAM17]. It is now also a moment of expansion, in which the concepts developed and methods employed in quantum information are beginning to permeate other areas of research. Exemplary cases are that of the study of information transmission in processes in biological systems [SIFW10], or obtaining deeper insights about how elementary particles interact to form the primary constituents of matter [TKU19].

This thesis is devoted to the exploration and exploitation of quantum information outside the boundaries of the field of quantum information science, both at the fundamental and applied levels. What we mean by the concept of *quantum information outside quantum information* is shown in Fig. 1.1. On one hand, we employ quantum information techniques to understand and characterize both fundamental concepts such as those of vacuum and causality, and applied scenarios such as microscopic thermal engines. On the other, we address the ways in which quantum effects may bring improvements in specific areas of technological and industrial relevance, as it is the case of machine learning.

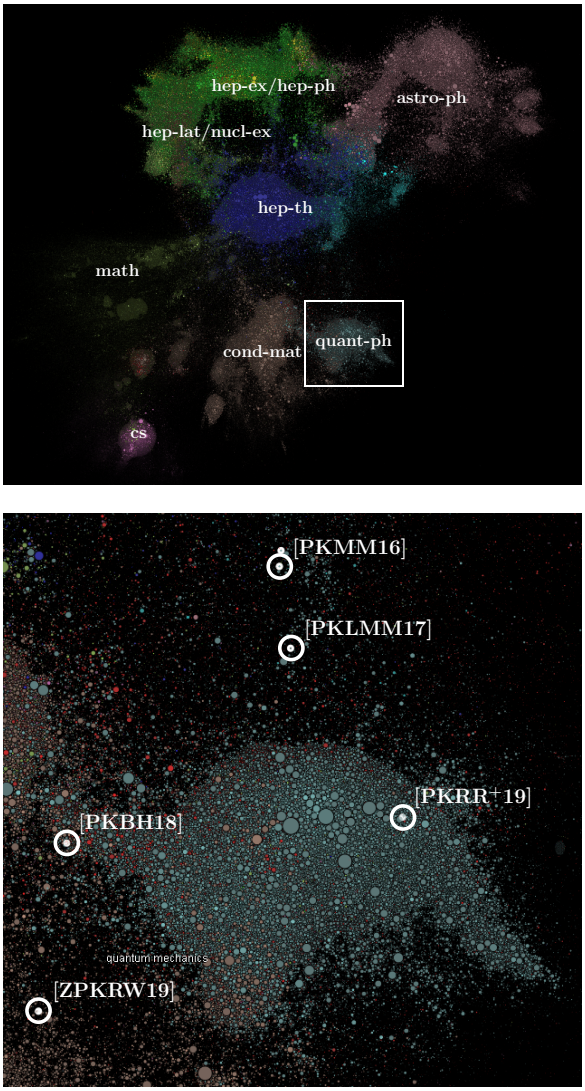

Figure 1.1: Quantum information outside quantum information. The top figure is a pictorial representation of the arXiv preprint repository, where manuscripts (represented by circles) are clustered together according to how they reference each other, as a measure of the proximity in topics. Different colors correspond to different main categories. The white square represents the area most related to quant-ph, which is associated to the turquoise color. Zooming in the white square, we obtain the image in the bottom. The studies conducted in this thesis and available on arXiv are surrounded by white circles. As it can be seen, most of them are placed in the boundary of quant-ph, connecting to other areas of physics. Images credit: www.paperscape.org

1.1 Motivation and contributions

Quantum information in quantum field theory

The XX century will pass to the annals of science as the century of physics. This was marked by two major breakthroughs: on one hand, the development of the Special and General theories of Relativity allowed to understand the nature of space, time and gravity, and with it the large-scale behavior of the universe. On the other, quantum mechanics developed into quantum electrodynamics and, later on, into quantum field theory, which explained the laws that nature follows at extremely small scales, identifying the ultimate constituents of matter and the ways in which they interact. The ultimate development of quantum field theory was the Standard Model of elementary particles, of which every part has been now confirmed in experiments. This is the reason why many experts argue that quantum mechanics, and its further developments, constitute the most accurate physical theory devised to date.

One of the most challenging concepts introduced by quantum field theory is that of the vacuum. It turns out that, in a spacetime devoid of any matter, pairs of particles and antiparticles are continuously created and annihilated due to Heisenberg's uncertainty principle. Aside from philosophical interpretations, the vacuum is a physical entity, that can be characterized and exploited in applications. In fact, vacuum fluctuations are harnessed everyday in laboratories around the world as drivers of nonlinear processes in optics such as high-order harmonic generation. They also constitute the key element to explain phenomena such as the evaporation of black holes, or the origin of the universe.

A particularly interesting feature about the vacuum is that it contains both classical and quantum correlations. This is, the vacuum fluctuations in different regions of spacetime are correlated. After this discovery, the first question asked was whether these correlations could be accessed, and exploited for information transmission or processing. In 1991 it was shown that the correlations in the vacuum can be extracted to quantum particles, that did not need to directly interact—in fact, they do not need to have knowledge of the existence of the other—in order to become entangled. The correlations present in the vacuum were indeed useful, and in recent years protocols have been developed that utilize them, for instance, to teleport energy between distant parties, to perform covert communication, or to obtain information about global properties of spacetime. Also, analyses were made that studied the impact of the quantum vacuum in information-processing tasks such as random number generation.

The exploitation of vacuum correlations remains largely a theoretical field. In fact, these correlations have been directly measured only at the beginning of

2019 and, in order to develop real applications, it remains to be understood if the extraction of correlations is possible under realistic assumptions, and under which conditions an efficient extraction is possible. Motivated by these questions, we develop two different projects on the extraction of entanglement from the quantum vacuum. On one hand, we analyze the extraction of correlations using realistic models of light-matter interaction. On the other hand, we aim to understand which are the mathematical conditions that must be satisfied in order for the extraction to be possible.

Main results

We study how fully-featured hydrogenlike atoms harvest entanglement from the electromagnetic field vacuum, even when the atoms are spacelike separated. We compare this case—qualitatively and quantitatively—with previous results that used scalar fields and featureless, idealized atomic models. We show that, under certain circumstances, relaxing previous idealizations makes vacuum entanglement harvesting more efficient.

On a different project, we show that pairs of identical particle detectors in spacelike separation can only harvest entanglement from the vacuum state of a quantum field when they have a nonzero energy gap. Furthermore, we show that degenerate probes are strongly challenged to become entangled through their interaction through scalar and electromagnetic fields even in full light-contact. We relate these results to previous literature on remote entanglement generation and entanglement harvesting, giving insight into the energy gap's protective role against local noise that prevents the detectors from getting entangled through the interaction with the field.

Quantum information for quantum thermodynamics

Thermodynamics, the branch of physics that analyzes energy exchange between systems, is traditionally studied within the set of transformations of variables such as volume, temperature, or pressure, that are agnostic to the microscopic properties of the systems under scrutiny. The agnosticism of the microscopic structure is so extreme that, in fact, classical equilibrium thermodynamics is grounded in three infinite-type assumptions: (i) interactions between systems are infinitely weak, (ii) as a consequence, in order for a transformation to take place, it must last an infinite amount of time, and (iii) there exist systems, called thermal baths, that are infinitely large and remain unperturbed after any interaction with other systems. While a priori impossible to realize, these assumptions are capable of predicting accurately thermodynamic processes in

the macroscopic regime. In fact, all engines operate based on these principles.

The framework changes drastically when dealing with quantum systems and interactions between them, for two main reasons. One comes from the fact that some effects in quantum mechanics are not possible when working with classical systems, and these effects may aid or hinder the execution of thermodynamic tasks. For instance, it is known that in certain scenarios using entangled systems to extract work from a bath allow for faster extraction, and thus improved power generation. The second reason is purely thermodynamic, as it is not possible to obviate microscopic properties when one deals with single or few particles interacting. Furthermore, the exact description of quantum systems requires a number of parameters that grows exponentially with the number of systems under scrutiny, making impossible to analyze systems even of moderate size. This hurdle can be overcome when considering restricted sets of states and transformations that have efficient characterizations. An example of this, of particular relevance to quantum optics and for implementation of quantum information protocols, is Gaussian quantum mechanics. Interestingly, this framework can efficiently model thermal states of systems of experimental interest and thermodynamic transformations on them.

Motivated by this, we employ the tools from quantum optics and quantum information provided by Gaussian quantum mechanics to analyze thermodynamic processes in the quantum regime. Our aim is to develop a computationally efficient model of a quantum thermal engine, where one can study the effects that the elimination of the standard thermodynamic assumptions have in important figures of merit such as the efficiency or the power output of the engine.

Main results

We study a driven harmonic oscillator operating an Otto cycle between two thermal baths of finite size. By making extensive use of the tools of Gaussian quantum mechanics, we directly simulate the dynamics of the engine as a whole, without the need of making any approximations regarding the strength and time of the interactions, or the size of the systems involved. This allows us to understand the nonequilibrium thermodynamics of the engine not only from the perspective of the working medium, but also as it is seen from the standpoint of the thermal baths. For sufficiently large baths, the engine is capable of running a number of ideal cycles, delivering finite power while operating very close to maximal efficiency. We additionally study the correlations generated in the system, and relate the buildup of correlations between the working medium and the baths, and between the baths themselves, to the degradation of the performance of the engine over the course of many cycles of operation.

Quantum information in machine learning

When seen from the point of view of applications of quantum mechanics in real-world scenarios, one must point out two important periods. First, the theoretical developments occurred during the first third of the XX century led to inventions in the decade of 1950, such as the transistor or the laser, that made use of quantum effects observable at the macroscopic scale. This period was described as the *first quantum revolution*. Nowadays, we are entering the *second quantum revolution*. Enabled by the access and control of individual quantum systems, we are beginning to exploit microscopic quantum effects such as entanglement, superposition and interference for establishing more secure communications, building more accurate sensors, or performing more powerful computations.

In fact, it is a strong belief that encoding information in quantum degrees of freedom will boost exponentially the performance of computers when compared with their classical counterparts. This will have a tremendous impact in science and society as we know it nowadays. For this reason, many efforts have been devoted throughout the last decades to build quantum computers, and as of today, several small prototypes exist, built both by academic institutions and private companies. These are small-scale and noisy computers, that do not allow yet to demonstrate an advantage with respect to classical computers. However, despite some issues remain to be solved, it is widely believed that general-purpose quantum computers will be a reality that will enable tackling the most challenging problems.

One of the areas where quantum computing is expected to bring great improvements is machine learning. Currently, machine learning algorithms—and most prominently among them, deep learning algorithms—process vast quantities of data in order to find patterns with which new information can be classified, predicted, or generated. A typical deep learning model has thousands of parameters that are tuned in order to perform some task, and this represents an important obstacle when analyzing the process that triggers a specific output. Fortunately, there are other machine learning paradigms, like the Bayesian approach, where questions like "which feature triggered this prediction?" are easier to formulate and answer. This does not come for free, as Bayesian machine learning methods are costly in classical computers.

Motivated by this, we explore the question of whether encoding information in quantum systems and processing this information using quantum operations could provide a benefit to the field of machine learning by enabling costly algorithms to be run in an efficient way.

Main results

We develop a new algorithm for Bayesian training of machine learning models in quantum computers, based in Gaussian process regression. The properties of the kernel matrix in the Gaussian process ensure the efficient execution of the core component of the protocol, the quantum linear systems algorithm, providing an at least polynomial speedup over classical algorithms. Furthermore, we demonstrate the execution of the algorithm on contemporary quantum computers and analyze its robustness with respect to realistic noise models.

Quantum information for understanding causality

Causality is a ubiquitous concept in science. Given some observed correlations among some measured variables, which causes could explain these correlations? This question appears in many disciplines, from statistics, physics, or engineering to biology, medicine and social sciences. As such, it has been widely discussed in the last decades. However, no efficient methods for characterizing correlations achievable in arbitrary causal scenarios were known until recently. These improvements, in fact, were inspired by tools commonly employed in the field of quantum information and nonlocality.

The notion of causality requires a redefinition when one inserts quantum systems into consideration. It is well known, since the results of Bell in 1964, that for a same causal pattern quantum systems can be correlated in ways impossible for classical systems. While there has been significant effort to develop an understanding of causes and effects in quantum mechanics, there is an important lack of methods to answer more operational questions. For instance, although methods exist for characterizing the quantum correlations achievable in various causal structures, these are typically of limited applicability, and thus very difficult to generalize. Even for more concrete questions, such as "is it possible to create these correlations in this causal structure?", no generic methods exist. This represents a fundamental gap in knowledge since our current description of nature—especially at the microscopic scale—is through quantum mechanics, and furthermore quantum effects are expected to play a crucial role in many phenomena in areas like thermodynamics, biology, or the study of many-body systems.

This motivates us to attempt closing the gap by looking for methods of general applicability for characterizing the correlations achievable in quantum causal structures. We develop two different projects on this matter. One one hand, we look for a way of including a generic type of constraints present in many causal scenarios, namely the causal independence of systems, in well-known

tools for the study of quantum correlations. On the other, we aim to tackle the characterization of correlations in arbitrary quantum causal structures by the generalization of known methods for the characterization of classical correlations in causal scenarios.

Main results

We present two methods that allow the study of classical and quantum correlations in complex causal structures. One is specifically tailored to scenarios with causally independent parties, such as the scenario underlying entanglement swapping, while the other is of general applicability. We show how the techniques successfully identify correlations not attainable in simple scenarios. Then we employ them to bound the maximal value of Bell-like operators in causal scenarios. Furthermore, we apply them to address study nonlocality in network scenarios. We show how the nonlocal power of measurements can be activated in a network: there exist measuring devices that, despite being unable to generate nonlocal correlations in the standard Bell scenario, provide a classical-quantum separation when used in complex networks.

1.2 Outline of the thesis

This thesis is organized as follows:

- Chapter 2 is dedicated to the study of the quantum information that can be extracted from the quantum vacuum. This chapter is based on the following original results: [PKMM16, PKLMM17].
- In Chapter 3 we employ tools of quantum information science to study the performance of quantum thermal engines. This chapter is based on the following original results: [PKBH18].
- In Chapter 4 we study the impact of quantum computing in the area of machine learning. This chapter is based on the following original results: [ZPKRW19].
- Chapter 5 is devoted to the study of correlations achievable in quantum causal structures.
- Chapter 6 summarizes the results obtained and discusses future research directions. This chapter is based on the following original results: [PKRR$^+$19, WPKG$^+$19], and also contains new, unpublished results[4].

[4] These results arise through collaborations with Antonio Acín, Denis Rosset, and Elie Wolfe, and with Antonio Acín and Daniel Cavalcanti. Manuscripts including these results are expected in the future.

2 | Information in quantum fields: Entanglement harvesting

Vacuum states—also known as ground states—of quantum fields, despite representing states of the field with no particles or excitations, exhibit interesting properties such as entanglement between nonoverlapping spatiotemporal regions. In this chapter we address the question of quantifying the amount of entanglement that can be extracted from the vacuum state of a quantum field. After an introduction to the general framework of entanglement harvesting, we explore the setting of atomic probes harvesting entanglement from the electromagnetic vacuum. We propose a model of interaction between the ground state of a vector field and dipolar probes, and proceed to analyze the impact of the new features available—the relative orientation of the dipole probes, and the exchange of angular momentum between the probes and the field—in the amount of entanglement harvested. We also show that, under the appropriate approximations, the phenomenology of harvesting from scalar fields is recovered. Finally, we show that the existence of a nonzero energy gap in the detectors is key for harvesting entanglement, by formally demonstrating that it is a necessary condition for extracting quantum correlations with spacelike-separated probes.

2.1 Introduction

One of the most fundamental differences between classical and quantum theories of fields is the completely different nature of their respective ground states. While in classical theories the ground state is just a state with zero energy density everywhere, the vacuum state of a quantum field—even a noninteracting one—contains classical and quantum correlations between field observables defined in different regions of spacetime. This is the case even when those regions are spacelike separated [SW85, SW87]. This nonclassical behavior of the vacuum lies deep in the core of quantum theory, and is a key ingredient in phenomena such as, e.g., *quantum energy teleportation* [Hot08, Hot09] (with implications, among other areas, in quantum thermodynamics [FFH14]), *quantum collect calling* [JMMK15] (with implications

in cosmology [BGMBMM15a, BGMBMM15b, BGMBMM16]), or even in the black hole information loss problem [Haw75, Haw76, HPS16, AMPS13, BPŻ13].

The correlations present in the vacuum are physically observable [LR47, Cas48, BCSSF19] and, in principle, usable: both classical and quantum vacuum correlations can be extracted from the field to quantum particle detectors that couple to it locally [Rez03, RRS05]. This allows two parties to entangle particle detectors that each one operates, even if they remain spacelike separated during their whole existence. The extraction of nonclassical correlations from the quantum vacuum to particle detectors was first pointed out by Valentini [Val91], and later on studied by Reznik *et al.* [Rez03, RRS05]. This phenomenon has become known as *entanglement harvesting*.

Entanglement harvesting from scalar fields is now well understood [PKMM15]. In fact, the ability to harvest entanglement from quantum fields has motivated numerous works aiming to harness this phenomenon in a variety of scenarios. These range from the extraction of resources and *entanglement farming* [MMBDK13] or metrology [SMM15], to its study in cosmology, where it has been shown that entanglement harvesting is very sensitive to the geometry of the underlying spacetime [SM09, MMM12] or even its topology [MMST16].

Despite the fact that entanglement harvesting has been proven robust, for instance, against uncertainties in the synchronization and spatial configuration of the particle detectors [MMS16], as most quantum phenomena it is inherently fragile to environmental conditions. To begin with, it has been observed that the entanglement harvested from a field decays with temperature, reaching zero at a finite value [Bro13], particularly when the detectors are in spacelike separation. This can be understood as caused by the known decay of quantum correlations in a quantum field with temperature [MV96, AW08]. In addition, for example, a sudden switching of the detector-field interaction—which induces local excitations of the detectors—is inefficient for harvesting spacelike entanglement since the local noise overshadows the correlations harvested from the field. In contrast, if the interaction is switched on adiabatically, it has been shown that it is possible to harvest entanglement with arbitrarily distant spacelike-separated detectors [PKMM15]. Nevertheless, the variety of situations in which entanglement harvesting has been found relevant has already motivated works that analyze the experimental feasibility of implementing timelike and spacelike entanglement harvesting protocols in both atomic and superconducting systems [OR11, OR12, SPdRMM12].

Most works in entanglement harvesting model the field-probe interaction by the simplified Unruh-DeWitt particle detector model [DHI79]. The Unruh-DeWitt model consists of a linear coupling of a spherically symmetric (often pointlike) two-level quantum system and a massless, scalar quantum field, some-

times with slight variations such as a spatial smearing function that endows the two-level system with a finite size [PKMM15]. Although in principle simple, this model was shown to capture the fundamental features of the light-matter interaction in scenarios where exchange of angular momentum does not play a role [MMMdR13, AKMM14]. The fact that this simple toy model of a particle detector can reproduce the light-matter interaction to a great extent is one of the reasons behind the strong interest that this model (in entanglement harvesting in particular) has attracted in more applied scenarios.

Nevertheless, and despite its great success, the Unruh-DeWitt model fails to capture the complete interaction between fully-featured atoms and the electromagnetic field vacuum. The electromagnetic field is not a scalar but a vector field, which carries angular momentum. Moreover, realistic atomic orbitals are, in general, not isotropic. These two features cannot be captured in the scalar Unruh-DeWitt model or any of its variants. This means that any study based on the Unruh-DeWitt model will not be able to see the inherent anisotropies and orientation dependence that entanglement harvesting has, and will not predict any effects related to the fact that the atoms can exchange angular momentum with the field (which, as we will see, may help or hinder their ability to harvest entanglement).

2.2 Preliminaries

Before describing the techniques we employ for studying entanglement harvesting, we first introduce the relevant concepts and notation that will be used throughout.

2.2.1 Measuring entanglement in bipartite systems

The main problem addressed in this chapter is that of the quantification of the entanglement that can be extracted from the vacuum state of a quantum field to two localized probes. The determination of a quantum state of a multipartite system being separable or entangled is in general an NP-hard problem [Gha10], and widely-used quantities such as the entanglement entropy $S(\rho) = -\mathrm{Tr}\left[\rho \log \rho\right]$ or distillable entanglement [BDSW96] have issues in certain situations. For instance, entanglement entropy can be nonzero for separable mixed states. However, for systems of two qubits, the negativity [VW02] is a measure of entanglement with many desirable properties: it vanishes for separable states, it is valid for mixed states, it does not increase under local operations and classical communication (i.e, it is an entanglement monotone), and it is—in its logarithmic version—additive.

The concept of negativity relies on the Peres-Horodecki criterion [Per96, H$^{\otimes 3}$96] for separability of quantum systems. Both negativity and the Peres-Horodecki criterion are based in the concept of positive partial transpose. For a general bipartite density matrix

$$\rho_{\mathrm{AB}} = \sum_{ijkl} \rho_{\mathrm{AB}}^{ijkl} \left|i\right\rangle \left\langle j\right| \otimes \left|k\right\rangle \left\langle l\right| \tag{2.1}$$

the partial transpose is defined as the transposition only of the system of one of the parties,

$$\rho_{\mathrm{AB}}^{\mathrm{T_B}} = \sum_{ijkl} \rho_{\mathrm{AB}}^{ijkl} \left|i\right\rangle \left\langle j\right| \otimes \left(\left|k\right\rangle \left\langle l\right|\right)^{\mathrm{T}} = \sum_{ijkl} \rho_{\mathrm{AB}}^{ijkl} \left|i\right\rangle \left\langle j\right| \otimes \left|l\right\rangle \left\langle k\right|. \tag{2.2}$$

Note that there exists a simple relation between the partial transpositions with respect to each of the subsystems:

$$\rho_{\mathrm{AB}}^{\mathrm{T_B}} = \sum_{ijkl} \rho_{\mathrm{AB}}^{ijkl} \left|i\right\rangle \left\langle j\right| \otimes \left(\left|k\right\rangle \left\langle l\right|\right)^{\mathrm{T}} = \sum_{ijkl} \rho_{\mathrm{AB}}^{ijkl} \left[\left(\left|j\right\rangle \left\langle i\right|\right)^{\mathrm{T}} \otimes \left|k\right\rangle \left\langle l\right|\right]^{\mathrm{T}} = \left(\rho_{\mathrm{AB}}^{\mathrm{T_A}}\right)^{\mathrm{T}}. \tag{2.3}$$

Peres and Horodecki showed that a necessary condition for density matrices representing bipartite quantum systems to be separable is that the eigenvalues of the partial transpose of the density matrix of the state must be all nonnegative. Therefore, if the partial transpose of the density matrix representing a state of a quantum bipartite system has at least one negative eigenvalue, then such state is guaranteed to be entangled.

This condition is also sufficient for bipartite qubit-qubit and qubit-qutrit systems. Therefore for the cases covered in this study we can easily relate the existence of negative eigenvalues with the state being entangled. In the desire of having an intuitive way of quantifying entanglement, Vidal and Werner [VW02] defined negativity as

$$\mathcal{N} \equiv \frac{\parallel \rho^{\mathrm{T_A}} \parallel_1 - 1}{2}, \tag{2.4}$$

where $\|O\|_1 = \mathrm{Tr}\sqrt{O^\dagger O}$ represents the trace norm of the operator O. This measure can be simply written as the absolute value of the sum of the negative eigenvalues of $\rho^{\mathrm{T_A}}$—or, since the eigenvalues of transposed matrices are the same, the sum of the negative eigenvalues of $\rho^{\mathrm{T_B}}$—and measures the degree to which $\rho^{\mathrm{T_A}}$ fails to be separable. In fact, it vanishes for all qubit-qubit and qubit-qutrit unentangled states, so entangled and separable states are clearly distinguished with negativity in the case of a two-qubit system like that used in this chapter.

2.2.2 Models of light-matter interaction

The first step when studying entanglement harvesting is to select a realistic model of light-matter interaction. We first review two scalar models that have been used in the past to approximate the coupling of atomic electrons with the electromagnetic field. Then we consider and thoroughly discuss the validity of an interaction model where the electron of a hydrogen atom couples as a dipole to the full electromagnetic field. This is the model that we will employ to derive the main results in this chapter.

Scalar coupling: Unruh-DeWitt Hamiltonian

The interaction between a nonrelativistic electron of momentum $\boldsymbol{p}$ and an electromagnetic field defined by a vector potential $\boldsymbol{A}(\boldsymbol{x},t)$ and a scalar potential $V(\boldsymbol{x},t)$ is given by the standard classical minimal coupling Hamiltonian

$$H = \frac{\boldsymbol{p}^2}{2m} - \frac{e}{m}\boldsymbol{p}\cdot\boldsymbol{A}(\boldsymbol{x},t) + \frac{e^2}{2m}\left[\boldsymbol{A}(\boldsymbol{x},t)\right]^2 + V(\boldsymbol{x},t). \tag{2.5}$$

This Hamiltonian is behind the so-called light-matter interaction when the electron is bound to an atom. Instead of using the full Hamiltonian (2.5), it is commonplace in the literature on entanglement harvesting to simplify the interaction and replace the electromagnetic coupling by a linear coupling between the monopole moment $\hat{\mu}$ of a pointlike two-level system (often referred to as *detector*) and a quantum scalar field $\hat{\phi}$. This coupling mimics the $\boldsymbol{p}\cdot\boldsymbol{A}$ term in Eq. (2.5). When this model is used to capture the features of the light-matter interaction, it is very common to argue that the scalar potential term in Eq. (2.5) can be ignored (typically by working in the Coulomb gauge) and that the quadratic term $\propto \boldsymbol{A}^2$ can be neglected since it is of higher order in e (for an exception to this see, e.g., [AKMM14]). The monopole moment in the interaction picture is given by

$$\hat{\mu}(t) = \hat{\sigma}^+ e^{\mathrm{i}\Omega t} + \hat{\sigma}^- e^{-\mathrm{i}\Omega t}, \tag{2.6}$$

where Ω is the gap between the two detector states. This monopole moment of the detector is then coupled to a scalar field $\hat{\phi}$ at the position where the detector is, denoted by $\boldsymbol{x}_d$. The specific form of the interaction Hamiltonian is

$$H_{\text{UDW}} = e\,\mathcal{X}(t)\hat{\mu}(t)\,\hat{\phi}(\boldsymbol{x}_d,t), \tag{2.7}$$

where e is the coupling constant and $\mathcal{X}(t)$ is a switching function that controls the time dependence of the interaction strength. This model of interaction is

known as the Unruh-DeWitt particle detector model [DHI79], which has been extensively used in fundamental studies in quantum field theory [BD84].

Under the assumption of interactions without exchange of angular momentum, the Unruh-DeWitt Hamiltonian captures the main features of the light-matter interaction [MMMdR13, AKMM14]. Oftentimes, the model is upgraded with a detector spatial smearing. The smearing of particle detectors may respond to the need to regularize divergences of the pointlike model [Sch04] or, as for example in quantum optics, to improve on the accuracy of the models of light-matter interaction considering that the atoms are not really pointlike objects, and instead they are localized in the full extension of their atomic wavefunctions [MMS16, AKMM14]. Furthermore, as discussed in [Sch04] spatial smearings are sometimes introduced implicitly in some form of soft UV regularization (see, e.g., [Tak86]).

To take into account corrections coming from the finite size of atoms, we can introduce an *ad hoc* spatial profile or *smearing function*, typically strongly supported on a finite spatial region, that controls how much each point of the detector in that region interacts with the field. This leads to the interaction Hamiltonian

$$H_{\text{UDW}} = e\mathcal{X}(t) \int \mathrm{d}^3\boldsymbol{x}\, S(\boldsymbol{x}-\boldsymbol{x}_d)\hat{\mu}(t)\, \hat{\phi}(\boldsymbol{x},t). \tag{2.8}$$

One could argue that to include the atomic orbital wavefunction geometry, it is natural to think that the spatial support of the atom could be associated with the spatial probability profile of the atomic wavefunctions [AKMM14]. In previous works in entanglement harvesting, different spatial smearings of strong support on a compact region have been studied [PKMM15]. Nevertheless, this is a feature that has to be added *ad hoc* to the spatial smearing profile in Eq. (2.7), since the atomic wavefunction association with the smearing function does not naturally arise in this simplified model.

Derivative coupling

In the same fashion that the minimal coupling Hamiltonian $\boldsymbol{p}\cdot\boldsymbol{A}$ is simplified into the Unruh-DeWitt Hamiltonian in Eq. (2.7), the dipole coupling Hamiltonian $\boldsymbol{d}\cdot\boldsymbol{E}$ could be again simplified to a scalar coupling when the interactions involve no exchange of angular momentum. Intuitively, given that in the Coulomb gauge the electric field is defined from the electromagnetic vector potential as $\boldsymbol{E} = -\partial_t \boldsymbol{A}$, an interaction Hamiltonian that couples the atomic monopole moment to the time derivative of a scalar field should capture some of the features

of the dipole coupling. Concretely, the following Hamiltonian is used:

$$H_{\mathrm{UDW_d}} = e\mathcal{X}(t) \int \mathrm{d}^3\boldsymbol{x}\, S(\boldsymbol{x}-\boldsymbol{x}_d)\hat{\mu}(t)\, \partial_t\hat{\phi}(\boldsymbol{x},t). \tag{2.9}$$

This model has been particularly useful in (1+1)-dimensional analyses, where the use of a derivative coupling alleviates IR divergences in the behavior of the Unruh-DeWitt model [Sch04, JAL14, MML14]. We see in Eq. (2.9) that, as in the case of Unruh-DeWitt detectors, a spatial profile can also be introduced *ad hoc* in this case to account for the finite size of the atomic probes [TBMM16].

Dipole coupling of an atom to the electromagnetic field

Let us now consider a model for the complete interaction of an atom with a vector electromagnetic field. We begin with the local *dipole coupling* between an electric dipole and an electric field,

$$\hat{\boldsymbol{d}} \cdot \hat{\boldsymbol{E}} = e\hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{E}}, \tag{2.10}$$

e being the charge of the dipole, $\hat{\boldsymbol{x}}$ its position operator, and $\hat{\boldsymbol{E}}$ the electric field operator. This coupling is extensively used in quantum optics to describe the light-matter interaction [SZ97]. It is well known that the leading-order contribution to atomic transitions is of a dipole nature and is governed by a term of the form (2.10). The intensity of higher multipole transitions is strongly suppressed and only becomes relevant for transitions forbidden by the dipole selection rules (see for instance [BJ03]).

Indeed, the dipole coupling is only an approximation for the full electromagnetic interaction of an atomic electron with the electromagnetic field. However, it is discussed in [LSS87, SZ97] that for realistically small atoms, an approximate gauge transformation yields the dipole coupling out of the full atomic-field coupling [an atomic electron minimally coupled to the electromagnetic field vector potential of Eq. (2.5)]. This approximation may break when the initial state of the field and the atoms is not excited and only for interaction times that are comparable or smaller than the length scale of the atoms. However, for interaction times much larger than the light-crossing time of the atomic radius the dipole coupling should yield a good approximation even for ground state dynamics. The coupling (2.10) is extensively used in atomic physics and quantum optics to successfully reproduce experiments [SVA$^+$07, Kra10] and in theoretical proposals [Tak86, SZ97, LS09, IMB$^+$15].

The dipole coupling is also convenient since the atom couples explicitly to a gauge-invariant field observable, and because, when the approximation holds,

the gauge choice made for the field degrees of freedom corresponds to the choice made in the conventional textbook solutions of the Schrödinger equation for an electron trapped in a Coulomb potential. That is, $\boldsymbol{A}=0$ in the absence of currents (for further discussion see [LSS87, SZ97] and the multipolar gauge in e.g. [Jac02]). Moreover, the fact that the commutator of $\boldsymbol{E}$ satisfies microcausality (i.e., it vanishes for spacelike separated events) means that the results of [MM15] for a scalar field can be quickly reproduced here for the electromagnetic interaction and thus the interaction between two atoms through the field as given in Eq. (2.10) is fully causal. This kind of coupling has also been used in other contexts outside quantum optics, as for instance to analyze the Unruh effect with an electromagnetic field [Tak86], or in quantum friction [IMB+15].

Let us, for simplicity and for the sake of comparison with the scalar models presented above, assume that only two levels of the atomic structure are relevant in our setup. In that case, the dipole operator enables transitions between the atomic ground state and one single relevant excited state. When we restrict it to only two levels, the operator $\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{E}}$ in the interaction picture reads

$$\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{E}} = \langle e|\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{E}}|g\rangle\, e^{\mathrm{i}\Omega t}\,|e\rangle\langle g| + \langle g|\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{E}}|e\rangle\, e^{-\mathrm{i}\Omega t}\,|g\rangle\langle e|\,.$$

Inserting resolutions of the identity in the position eigenbasis and noting that $\psi_g(\boldsymbol{x})=\langle\boldsymbol{x}|g\rangle$ and $\psi_e(\boldsymbol{x})=\langle\boldsymbol{x}|e\rangle$ are the position representation of the ground and excited level wavefunctions, respectively, the operator (2.10) can be recast as

$$\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{E}}(\boldsymbol{x},t) = \int \mathrm{d}^3\boldsymbol{x}\left[\boldsymbol{S}(\boldsymbol{x})\cdot\hat{\boldsymbol{E}}(\boldsymbol{x},t)e^{\mathrm{i}\Omega t}\,|e\rangle\langle g| + \boldsymbol{S}^*(\boldsymbol{x})\cdot\hat{\boldsymbol{E}}(\boldsymbol{x},t)e^{-\mathrm{i}\Omega t}\,|g\rangle\langle e|\right], \tag{2.11}$$

where $\hat{\boldsymbol{E}}(\boldsymbol{x},t)$ is an operator which acts on the field Hilbert space, and the spatial *smearing vector* $\boldsymbol{S}(\boldsymbol{x})$ is defined as

$$\boldsymbol{S}(\boldsymbol{x}) = \psi_e^*(\boldsymbol{x})\boldsymbol{x}\,\psi_g(\boldsymbol{x}). \tag{2.12}$$

Note that in this case, the specific form of the spatial smearing arises *naturally* from the coupling, and does not have to be inserted artificially as in the case of the Unruh-DeWitt and derivative coupling Hamiltonians (2.8) and (2.9).

By direct comparison with Eq. (2.10), the position-space representation of the dipole moment in the interaction picture can be written as

$$\hat{\boldsymbol{d}}(\boldsymbol{x},t) = e\left[\boldsymbol{S}(\boldsymbol{x})e^{\mathrm{i}\Omega t}\hat{\sigma}^+ + \boldsymbol{S}^*(\boldsymbol{x})e^{-\mathrm{i}\Omega t}\hat{\sigma}^-\right], \tag{2.13}$$

where we have adopted the usual notation for the SU(2) ladder operators $\hat{\sigma}^+=|e\rangle\langle g|$ and $\hat{\sigma}^-=|g\rangle\langle e|$. With this expression for the dipole moment, the

Hamiltonian for the dipole interaction reads

$$H_{\text{EM}} = \mathcal{X}(t) \int \mathrm{d}^3\boldsymbol{x}\, \hat{\boldsymbol{d}}(\boldsymbol{x} - \boldsymbol{x}_d, t) \cdot \hat{\boldsymbol{E}}(\boldsymbol{x}, t). \tag{2.14}$$

In the following we use Eq. (2.14) as the model of light-matter interaction with which we will analyze entanglement harvesting. In doing so, we will be able to qualitatively and quantitatively compare the results with previous models that neglected the vector nature of the field.

2.2.3 System dynamics

We proceed now to study the evolution of two probes—denoted as A and B—that interact locally with the vacuum state of a field. The full interaction Hamiltonian is

$$H_{\text{EM}} = \sum_{\nu} \mathcal{X}_\nu(t) \int \mathrm{d}^3\boldsymbol{x}\, \hat{\boldsymbol{d}}_\nu(\boldsymbol{x} - \boldsymbol{x}_\nu, t) \cdot \hat{\boldsymbol{E}}(\boldsymbol{x}, t) \tag{2.15}$$

in the interaction picture, where $\nu = \mathrm{A}, \mathrm{B}$ labels each of the two atoms. Each atom ν is located in position $\boldsymbol{x}_\nu$. We recall that the dipole moment operator is given by Eq. (2.13), that is

$$\hat{\boldsymbol{d}}_\nu(\boldsymbol{x}, t) = e_\nu \left[\boldsymbol{S}_\nu(\boldsymbol{x}) e^{\mathrm{i}\Omega_\nu t} \hat{\sigma}_\nu^+ + \boldsymbol{S}_\nu^*(\boldsymbol{x}) e^{-\mathrm{i}\Omega_\nu t} \hat{\sigma}_\nu^- \right], \tag{2.16}$$

and

$$\boldsymbol{S}_\nu(\boldsymbol{x}) = \psi^*_{e\,\nu}(\boldsymbol{x}) \boldsymbol{x}\, \psi_{g_\nu}(\boldsymbol{x}). \tag{2.17}$$

In addition to the study of the electromagnetic coupling, we will compare the analysis with the other scalar approximations used in past studies of entanglement harvesting, and discussed in the previous section. Again in the interaction picture, the scalar-field interaction Hamiltonians for two detectors coupled to the field are

$$H_{\text{UDW}} = \sum_{\nu} e_\nu \mathcal{X}_\nu(t) \int \mathrm{d}^3\boldsymbol{x}\, S_\nu(\boldsymbol{x} - \boldsymbol{x}_\nu) \hat{\mu}_\nu(t) \hat{\phi}(\boldsymbol{x}, t) \tag{2.18}$$

for the Unruh-DeWitt coupling and

$$H_{\text{UDW}_\text{d}} = \sum_{\nu} e_\nu \mathcal{X}_\nu(t) \int \mathrm{d}^3\boldsymbol{x} S_\nu(\boldsymbol{x} - \boldsymbol{x}_\nu) \hat{\mu}_\nu(t) \partial_t \hat{\phi}(\boldsymbol{x}, t) \tag{2.19}$$

for the derivative coupling.

In the three cases in Eqs. (2.15), (2.18) and (2.19) we can express the fields in terms of plane-wave mode expansions in the standard way [BD84, SZ97], i.e.,

$$\hat{\phi}(\boldsymbol{x},t) = \int \frac{\mathrm{d}^3\boldsymbol{k}}{\sqrt{(2\pi)^3 2|\boldsymbol{k}|}} \left[\hat{a}_{\boldsymbol{k}} e^{\mathrm{i}k\cdot x} + \hat{a}^\dagger_{\boldsymbol{k}} e^{-\mathrm{i}k\cdot x}\right], \tag{2.20}$$

$$\partial_t\hat{\phi}(\boldsymbol{x},t) = \int \frac{\mathrm{d}^3\boldsymbol{k}}{\sqrt{(2\pi)^3}} \sqrt{\frac{|\boldsymbol{k}|}{2}} \left[-\mathrm{i}\hat{a}_{\boldsymbol{k}} e^{\mathrm{i}k\cdot x} + \mathrm{i}\hat{a}^\dagger_{\boldsymbol{k}} e^{-\mathrm{i}k\cdot x}\right], \tag{2.21}$$

$$\hat{\boldsymbol{E}}(\boldsymbol{x},t) = \sum_s \int \frac{\mathrm{d}^3\boldsymbol{k}}{\sqrt{(2\pi)^3}} \sqrt{\frac{|\boldsymbol{k}|}{2}} \left[-\mathrm{i}\hat{a}_{\boldsymbol{k},s}\boldsymbol{\epsilon}(\boldsymbol{k},s) e^{\mathrm{i}k\cdot x} + \mathrm{i}\hat{a}^\dagger_{\boldsymbol{k},s}\boldsymbol{\epsilon}^*(\boldsymbol{k},s) e^{-\mathrm{i}k\cdot x}\right], \tag{2.22}$$

where $k\cdot x \coloneqq \boldsymbol{k}\cdot\boldsymbol{x} - |\boldsymbol{k}|t$. Note that each $\hat{a}$ (resp. $\hat{a}^\dagger$) represents an annihilation (resp. creation) operator for a field mode of momentum $\boldsymbol{k}$ and, in the electromagnetic case, polarization s. These operators satisfy the following canonical commutation relations:

$$[\hat{a}_{\boldsymbol{k}}, \hat{a}^\dagger_{\boldsymbol{k}'}] = \delta^{(3)}(\boldsymbol{k}-\boldsymbol{k}'),$$
$$[\hat{a}_{\boldsymbol{k},s}, \hat{a}^\dagger_{\boldsymbol{k}',s'}] = \delta^{(3)}(\boldsymbol{k}-\boldsymbol{k}')\delta_{s,s'}.$$

In the case of the electromagnetic field, its vector nature is encoded in the set $\{\boldsymbol{\epsilon}(\boldsymbol{k},s)\}_{s=1}^3$ of orthonormal (in general complex) *polarization vectors*. Maxwell's equations force the electric field to satisfy $\nabla\cdot\boldsymbol{E} = 0$, which yields the *transversality condition* $\boldsymbol{k}\cdot\boldsymbol{\epsilon}(\boldsymbol{k},s) = 0$. This constraint restricts the sum over polarizations in Eq. (2.22) to $s = 1, 2$, representing the two physical polarizations of the electromagnetic field, which are typically denoted as $\{\perp_1, \perp_2\}$ and satisfy the *completeness relation*

$$\sum_{s=\perp_1,\perp_2} \boldsymbol{\epsilon}(\boldsymbol{k},s)\otimes\boldsymbol{\epsilon}(\boldsymbol{k},s) = \mathbb{1} - \frac{\boldsymbol{k}\otimes\boldsymbol{k}}{|\boldsymbol{k}|^2}. \tag{2.23}$$

For completeness, an explicit derivation of this (otherwise well-known) expression can be seen in Appendix A.1.

The time evolution generated by each respective interaction Hamiltonian, (2.15), (2.18) and (2.19), can be obtained by a perturbative Dyson expansion of the time-evolution operator

$$U = \underbrace{\mathbb{1}}_{U^{(0)}} \underbrace{-\mathrm{i}\int_{-\infty}^{\infty}\mathrm{d}t\, H(t)}_{U^{(1)}} \underbrace{-\int_{-\infty}^{\infty}\mathrm{d}t\int_{-\infty}^{t}\mathrm{d}t'\, H(t)H(t')}_{U^{(2)}} + \dots$$

Therefore, the time evolution of an initial state $\hat{\rho}_0$, given by $\hat{\rho} = U\hat{\rho}_0 U^\dagger$, can be written as a perturbative expansion in the overall coupling strength $e^{i+j} = e_{\text{A}}^i e_{\text{B}}^j$,

$$\hat{\rho} = \hat{\rho}_0 + \hat{\rho}^{(1,0)} + \hat{\rho}^{(0,1)} + \hat{\rho}^{(2,0)} + \hat{\rho}^{(0,2)} + \hat{\rho}^{(1,1)} + \mathcal{O}(e^3).$$

Here the notation $\hat{\rho}^{(i,j)}$ represents the correction to the initial state $\hat{\rho}_0$ obtained by acting on it with $U^{(i)}$ from the left and with ${U^{(j)}}^\dagger$ from the right, namely

$$\hat{\rho}^{(i,j)} = U^{(i)}\hat{\rho}_0 {U^{(j)}}^\dagger. \tag{2.24}$$

We are interested in the harvesting of entanglement from the vacuum state of the field. Therefore, we consider that the initial state of the atoms-field system is

$$\hat{\rho}_0 = |0\rangle\langle 0| \otimes \hat{\rho}_{\text{AB},0}, \tag{2.25}$$

where $|0\rangle$ is the vacuum state of the field and $\hat{\rho}_{\text{AB},0}$ is the joint initial state of the atoms. After the evolution, one is only interested in the partial state of the probes, which is given by

$$\hat{\rho}_{\text{AB}} = \text{Tr}_{\hat{\phi}}(\hat{\rho}), \tag{2.26}$$

where $\text{Tr}_{\hat{\phi}}$ represents the partial trace over the field degrees of freedom.

This means that the nondiagonal terms in the field produced by the time evolution will not be relevant to the phenomenon. In particular, as long as the initial state of the field is diagonal in the Fock basis (as it is for the case of the vacuum, or any incoherent superposition of Fock states such as a thermal state), any contribution $\hat{\rho}^{(i,j)}$ for which the parities of i and j are different (i.e., any correction with overall coupling constant e^{i+j} with $i+j$ odd) will not contribute to the final state of the detectors (2.26).

Furthermore, we assume that both detectors are initially on their respective ground states as in previous studies of entanglement harvesting from scalar fields [PKMM15]. In this fashion, the initial joint state has even parity and reads

$$\hat{\rho}_{\text{AB},0} = |g_{\text{A}}\rangle\langle g_{\text{A}}| \otimes |g_{\text{B}}\rangle\langle g_{\text{B}}|. \tag{2.27}$$

Although they have a completely different spatial structure, the monopole moment operator (2.6) and the dipole moment (2.13) have the same internal Hilbert space structure, namely, they are Hermitian linear combinations of the SU(2) ladder operators. This is due to the fact that, for now, we are focusing our study on a particular transition involving only two atomic levels. This in turn means that the time-evolved density matrix of the detectors in the three cases discussed will have the same structure (although with notably different coefficients) when written in the basis

$$\{|g_{\text{A}}\rangle \otimes |g_{\text{B}}\rangle, |e_{\text{A}}\rangle \otimes |g_{\text{B}}\rangle, |g_{\text{A}}\rangle \otimes |e_{\text{B}}\rangle, |e_{\text{A}}\rangle \otimes |e_{\text{B}}\rangle\}.$$

Namely, the time-evolved density matrix is

$$\hat{\rho}_{\mathrm{AB}} = \begin{pmatrix} 1-\mathcal{L}_{\mathrm{AA}}-\mathcal{L}_{\mathrm{BB}} & 0 & 0 & \mathcal{M}^* \\ 0 & \mathcal{L}_{\mathrm{AA}} & \mathcal{L}_{\mathrm{AB}} & 0 \\ 0 & \mathcal{L}_{\mathrm{BA}} & \mathcal{L}_{\mathrm{BB}} & 0 \\ \mathcal{M} & 0 & 0 & 0 \end{pmatrix} + \mathcal{O}(e^4). \tag{2.28}$$

At a first sight, one can think that this matrix fails to be positive given the conditions in [MMST16]. We show in Appendix A.2 that this matrix is indeed positive at leading order, $\mathcal{O}(e^2)$, in perturbation theory.

The functions $\mathcal{L}_{\mu\nu}$ and $\mathcal{M}$ in the electromagnetic case can be evaluated as

$$\begin{aligned}\mathcal{L}^{\mathrm{EM}}_{\mu\nu} =& e_\mu e_\nu \int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{\infty}\mathrm{d}t_2\int\mathrm{d}^3\boldsymbol{x}_1\int\mathrm{d}^3\boldsymbol{x}_2 \\ &\times e^{\mathrm{i}(\Omega_\mu t_1-\Omega_\nu t_2)}\mathcal{X}_\mu(t_1)\mathcal{X}_\nu(t_2)\boldsymbol{S}^{*\mathbf{t}}_\nu(\boldsymbol{x}_2-\boldsymbol{x}_\nu)\mathbf{W}(\boldsymbol{x}_2,\boldsymbol{x}_1,t_2,t_1)\boldsymbol{S}_\mu(\boldsymbol{x}_1-\boldsymbol{x}_\mu),\end{aligned} \tag{2.29}$$

$$\begin{aligned}\mathcal{M}^{\mathrm{EM}} =& -e_{\mathrm{A}}e_{\mathrm{B}}\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\int\mathrm{d}^3\boldsymbol{x}_1\int\mathrm{d}^3\boldsymbol{x}_2 \\ &\times\Big[e^{\mathrm{i}(\Omega_{\mathrm{A}}t_1+\Omega_{\mathrm{B}}t_2)}\mathcal{X}_{\mathrm{A}}(t_1)\mathcal{X}_{\mathrm{B}}(t_2)\boldsymbol{S}^{\mathbf{t}}_{\mathrm{A}}(\boldsymbol{x}_1-\boldsymbol{x}_{\mathrm{A}})\mathbf{W}(\boldsymbol{x}_1,\boldsymbol{x}_2,t_1,t_2)\boldsymbol{S}_{\mathrm{B}}(\boldsymbol{x}_2-\boldsymbol{x}_{\mathrm{B}}) \\ &+e^{\mathrm{i}(\Omega_{\mathrm{B}}t_1+\Omega_{\mathrm{A}}t_2)}\mathcal{X}_{\mathrm{B}}(t_1)\mathcal{X}_{\mathrm{A}}(t_2)\boldsymbol{S}^{\mathbf{t}}_{\mathrm{B}}(\boldsymbol{x}_1-\boldsymbol{x}_{\mathrm{B}})\mathbf{W}(\boldsymbol{x}_1,\boldsymbol{x}_2,t_1,t_2)\boldsymbol{S}_{\mathrm{A}}(\boldsymbol{x}_2-\boldsymbol{x}_{\mathrm{A}})\Big],\end{aligned} \tag{2.30}$$

where $\boldsymbol{S}^{\mathbf{t}}_\nu$ and $\boldsymbol{S}^{*\mathbf{t}}_\nu$ are respectively the transpose and Hermitian conjugate of the vector $\boldsymbol{S}_\nu$, the spatial smearing vectors are given by Eq. (2.12), and the matrix $\mathbf{W}$ is the Wightman function 2-tensor of the field, whose components are given by

$$[\mathbf{W}]_{ij} = W^{\mathrm{EM}}_{ij}(\boldsymbol{x},\boldsymbol{x}',t,t') = \langle 0|E_i(\boldsymbol{x},t)E_j(\boldsymbol{x}',t')|0\rangle\,. \tag{2.31}$$

In the scalar cases, Eqs. (2.29) and (2.30) have a similar form, but the corresponding scalar analogues of the vector quantities have to be used. This is, the smearing vectors $\boldsymbol{S}(\boldsymbol{x})$ are replaced by the scalar smearing functions $S(\boldsymbol{x})$, and instead of the Wightman tensor one has to use the scalar Wightman function of the corresponding field,

$$W^{\mathrm{UDW}}(\boldsymbol{x},\boldsymbol{x}',t,t') = \langle 0|\hat{\phi}(\boldsymbol{x},t)\hat{\phi}(\boldsymbol{x}',t')|0\rangle\,, \tag{2.32}$$

$$\begin{aligned}W^{\mathrm{UDW_d}}(\boldsymbol{x},\boldsymbol{x}',t,t') &= \langle 0|\partial_t\hat{\phi}(\boldsymbol{x},t)\partial_{t'}\hat{\phi}(\boldsymbol{x}',t')|0\rangle \\ &= \partial_t\partial_{t'}W^{\mathrm{UDW}}(\boldsymbol{x},\boldsymbol{x}',t,t').\end{aligned} \tag{2.33}$$

2.2.4 Quantifying the harvested entanglement

As stated earlier, we are interested in analyzing the entanglement acquired between the two atoms after the interaction with the field vacuum. To quantify the entanglement between them we will use the negativity. For a state of the form of Eq. (2.28), the negativity is given by $\mathcal{N} = \max\left(0, \mathcal{N}^{(2)}\right) + \mathcal{O}(e^4)$, where

$$\mathcal{N}^{(2)} = -\frac{1}{2}\left(\mathcal{L}_{\mathrm{AA}} + \mathcal{L}_{\mathrm{BB}} - \sqrt{\left(\mathcal{L}_{\mathrm{AA}} - \mathcal{L}_{\mathrm{BB}}\right)^2 + 4|\mathcal{M}|^2}\right). \tag{2.34}$$

A naïve inspection of the partial transpose of Eq. (2.28) as it stands produces the seemingly always-negative eigenvalue $E_2 = -|\mathcal{L}_{\mathrm{AB}}|^2$, potentially leading to having the probes always entangled, regardless of the details in the setup. However, as discussed in [PKMM15], this term is $\mathcal{O}(e^4)$; thus the second-order expansion in Eq. (2.34) is not enough to study this subdominant eigenvalue since the full fourth-order density matrix must be computed. It was shown in [PKMM15] that (for the Unruh-DeWitt model) no relevant changes to the second-order result were obtained when subleading effects were considered.

We will consider the simple case in which both atoms are identical, which implies that $\Omega_{\mathrm{A}} = \Omega_{\mathrm{B}} \equiv \Omega$, $e_{\mathrm{A}} = e_{\mathrm{B}} \equiv e$ and $\boldsymbol{S}_{\mathrm{A}} = \boldsymbol{S}_{\mathrm{B}}$. In this case, the $\mathcal{L}$ terms are equal, and Eq. (2.34) simplifies to

$$\mathcal{N}^{(2)} = |\mathcal{M}| - \mathcal{L}_{\mu\mu}, \tag{2.35}$$

where $\mathcal{L}_{\mu\mu} \equiv \mathcal{L}_{\mathrm{AA}} = \mathcal{L}_{\mathrm{BB}}$.

This expression allows for a very intuitive interpretation. When $\mu = \nu$, Eq. (2.29) is a local term: it only depends on the properties of just one of the atoms. On the other hand, Eq. (2.30) depends on the properties of both atoms and therefore is a nonlocal term. This means that in Eq. (2.35) there is a direct competition between nonlocal, entangling excitations, and local noise, leading to the intuitive notion that in order to have entanglement between the atoms, the nonlocal term must overcome the local, single-atom "noisy" excitations [RRS05, PKMM15, MMS16]. It is worth noticing that, as shown in [MMST16], for this case of two identical atoms, the second-order correction to the negativity is related to another conventional entanglement measure, the concurrence [Woo98], by the simple relation $\mathcal{C}^{(2)} = 2\mathcal{N}^{(2)}$.

2.3 Entanglement harvesting with hydrogenlike atoms

After having an overview on the tools employed for analyzing entanglement harvesting, let us use them to analyze the ability of atomic probes to harvest

entanglement from the vacuum state of the electromagnetic field. We will model the interactions between individual atomic probes and the quantum electromagnetic vacuum through the dipole coupling Hamiltonian of Eq. (2.15). Making use of it, we present a full study of how the harvesting power of transitions of fully-featured hydrogenoid atoms depends on the anisotropic character of the atoms' excited states, and how the exchange of angular momentum impacts the harvesting of entanglement.

2.3.1 Characterization of the atomic model

In the case of the electromagnetic coupling one needs to specify the wavefunctions of the ground and excited states of the probes as per Eq. (2.17). We will consider two identical, static, hydrogenlike atoms located at positions $\boldsymbol{x}_{\mathrm{A}}$ and $\boldsymbol{x}_{\mathrm{B}}$, which interact with an electromagnetic field in its vacuum state for a finite time scale T implemented through strongly-supported smooth Gaussian switching functions[1]

$$\mathcal{X}_\nu(t) = e^{-\frac{(t-t_\nu)^2}{T^2}}, \tag{2.36}$$

where t_ν is the center of the Gaussian of atom ν. This choice of switching function arises naturally when one thinks, for instance, of atoms flying transversely through a long optical cavity: the ground transversal mode has precisely a Gaussian profile [Sve10]. Also, in the case of harvesting from scalar fields, smooth switchings in the form of Eq. (2.36) were shown to allow for increased harvesting over more abrupt switchings [PKMM15].

We will consider only two of the atomic levels in the probes. We choose the ground state to be the hydrogenoid-1s state [GP92]

$$\psi_g^{\mathrm{EM}}(\boldsymbol{x}) = \psi_{100}(\boldsymbol{x}) = R_{10}(|\boldsymbol{x}|)Y_{00}(\theta,\phi) = \frac{1}{\sqrt{\pi a_0^3}} e^{-\frac{|\boldsymbol{x}|}{a_0}}, \tag{2.37}$$

where a_0 is the generalized Bohr radius.

In the electromagnetic case, under the realistic consideration of hydrogenlike atoms, the smearing functions (smearing vectors) are no longer rotationally invariant. Indeed, the excited state has to be at least a $2p$ level—which is no

[1] Although these switching functions are not strictly compactly supported, whenever we study harvesting in "spacelike" regimes (at more than 7–8 standard deviations away from maximal light contact) we also double-test the results by performing numerical studies where we substitute the Gaussian switchings with a compactified version for which we enforce that $\mathcal{X}(t) = 0$ when $t > 8T/\sqrt{2}$ (i.e., 8-sigma away from the Gaussian peak) and we make sure that the difference is below the numerical double-precision threshold for zero. This was done to guarantee that the results are not an artifact of the switching function Gaussian tails.

longer spherically symmetric—due to the angular momentum selection rules of the dipole interaction. This shows already one of the main differences between previous scalar models and the electromagnetic interaction: new features will appear as we consider transitions with exchange of angular momentum that could not possibly be captured by any scalar model. Furthermore, the lack of rotational invariance of the smearing vectors $\boldsymbol{S}_\nu$ translates into a dependence of entanglement harvesting on the relative orientation of the two atoms.

For illustration, we will consider that the excited level is the nonisotropic, $2p_z$ excited state[2]:

$$\psi_e^{\mathrm{EM}}(\boldsymbol{x}) = R_{21}(|\boldsymbol{x}|)Y_{10}(\theta,\phi) = \frac{1}{\sqrt{32\pi a_0^5}} e^{-\frac{|\boldsymbol{x}|}{2a_0}} |\boldsymbol{x}| \cos\theta. \tag{2.38}$$

Equations (2.37) and (2.38) completely determine the spatial smearing vector (2.12) for each atom:

$$\boldsymbol{S}_{\mathrm{A}}(\boldsymbol{x}) = \frac{\cos\theta}{4\pi a_0^4\sqrt{2}} e^{-\frac{3|\boldsymbol{x}|}{2a_0}} |\boldsymbol{x}|^2 \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}, \tag{2.39}$$

$$\boldsymbol{S}_{\mathrm{B}}(\boldsymbol{x}) = \frac{\cos\theta\cos\vartheta - \sin\theta\sin\vartheta\cos(\psi+\phi)}{4\pi a_0^4\sqrt{2}} e^{-\frac{3|\boldsymbol{x}|}{2a_0}} |\boldsymbol{x}|^2 \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}, \tag{2.40}$$

where the triplet (ψ,ϑ,φ) denotes the three Euler angles that define the relative orientation of the reference frame of atom B with respect to atom A's reference frame (see Fig. 2.1). The different angular dependences of the smearing vectors $\boldsymbol{S}_{\mathrm{A}}$ and $\boldsymbol{S}_{\mathrm{B}}$ are due to the fact that we have expressed the reference frame of atom B in terms of the reference frame of atom A. In so doing, the spherical harmonics of atom B (which take a canonical form in its own reference frame) transform to A's frame linearly,

$$Y_{lm}^{\mathrm{B}}(\theta,\phi) = \sum_{\mu=-l}^{l} Y_{l\mu}^{\mathrm{A}}(\theta,\phi)\mathcal{D}^l_{\mu,m}(\psi,\vartheta,\varphi), \tag{2.41}$$

where the different $\mathcal{D}^l_{m_1,m_2}(\psi,\vartheta,\varphi)$ are the Wigner D-functions that characterize the rotation of the angular momentum operators under changes of reference frame. For a more detailed description and the explicit definition of the Wigner D-functions, see Appendix A.3.

[2] Identical results would be obtained for the $2p_x$ and $2p_y$ levels.

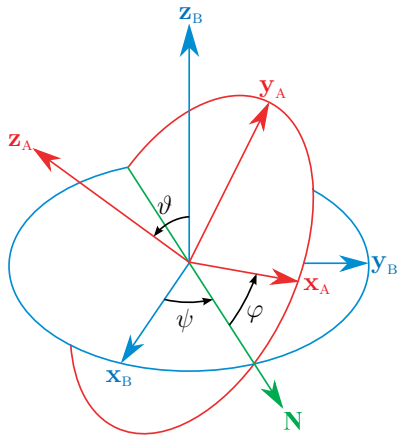

Figure 2.1: Euler angles characterizing the relative orientation of atom B (blue) with respect to atom A's reference frame (red).

2.3.2 Local noise and correlation term

Under the assumptions above, the calculations are (very) lengthy but straightforward. In Appendix A.3 we show in very great detail how to compute the local term (2.29) and the correlation term (2.30) for general transitions between any two arbitrary states of the two hydrogenoid atoms. In particular, for a transition between the $1s$ and the $2p_z$ orbitals, as we also show in the appendix, the two terms involved in Eq. (2.35) take the following form:

$$\mathcal{L}^{\mathrm{EM}}_{\mu\mu} = e^2 \frac{49152}{\pi} a_0^2 T^2 \int_0^\infty \mathrm{d}|\boldsymbol{k}||\boldsymbol{k}|^3 e^{-\frac{1}{2}T^2(\Omega+|\boldsymbol{k}|)^2} \frac{1}{\left(4a_0^2|\boldsymbol{k}|^2+9\right)^6}, \tag{2.42}$$

$$\begin{aligned}|\mathcal{M}|^{\mathrm{EM}} = e^2 \frac{24576\,|\cos\vartheta|}{\pi} a_0^2 T^2 \Bigg| \int_0^\infty \mathrm{d}|\boldsymbol{k}||\boldsymbol{k}|^3 e^{-\frac{1}{2}T^2\left(\Omega^2+|\boldsymbol{k}|^2\right)} \\ \times \frac{j_0(|\boldsymbol{k}||\boldsymbol{x}_\mathrm{A}-\boldsymbol{x}_\mathrm{B}|)+j_2(|\boldsymbol{k}||\boldsymbol{x}_\mathrm{A}-\boldsymbol{x}_\mathrm{B}|)}{\left(4a_0^2|\boldsymbol{k}|^2+9\right)^6} \left[E(|\boldsymbol{k}|,t_{\mathrm{BA}})+E(|\boldsymbol{k}|,-t_{\mathrm{BA}})\right]\Bigg|,\end{aligned} \tag{2.43}$$

where

$$E(|\boldsymbol{k}|,t_{\mathrm{BA}}) = e^{\mathrm{i}|\boldsymbol{k}|t_{\mathrm{BA}}}\mathrm{erfc}\left(\frac{\mathrm{i}T^2|\boldsymbol{k}|+t_{\mathrm{BA}}}{\sqrt{2}T}\right), \tag{2.44}$$

$j_l(x)$ are the spherical Bessel functions of the first kind, $\mathrm{erfc}(x)=1-\mathrm{erf}(x)$ is the complementary error function, $t_{\mathrm{BA}}=t_\mathrm{B}-t_\mathrm{A}$ represents the time delay between

the switchings, and we recall that the Euler angle ϑ is the relative angle between the axes of symmetry of the $2p_z$ orbitals of atoms A and B.

We see explicitly in Eq. (2.42) that the local noise term $\mathcal{L}_{\mu\mu}$ depends only on the properties of only one of the atoms. As expected, the dependences on the relative distance $|\boldsymbol{x}_{\mathrm{A}} - \boldsymbol{x}_{\mathrm{B}}|$ and relative angle between the $2p_z$ orbitals' axes of symmetry ϑ appear only in the nonlocal correlations term (2.43).

2.3.3 Orientation dependence of entanglement harvesting

Given that, as discussed above, the orbitals of the excited states are not isotropic, the relative orientation of the atoms is a new, unexplored degree of freedom that has a nontrivial influence on the ability of the atomic probes to harvest entanglement from the field. In Fig. 2.2 we show the amount of entanglement that the two atoms harvest when placed in light contact ($d = t_{\mathrm{BA}}$) as the relative angle ϑ between them varies. As it can be seen from Eqs. (2.34) and (2.43), the same would apply for atoms in spacelike separation.

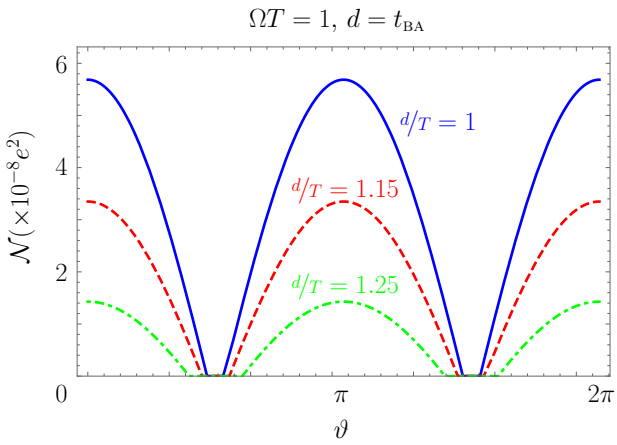

Figure 2.2: Negativity (at leading order) when the two atoms are in full light-contact, $d = t_{\mathrm{BA}}$, for two hydrogenoid atoms with energy gap $\Omega T = 1$, atomic radius $a_0\Omega = 0.001$ and spatial separations (solid blue) $d/T = 1$, (dashed red) $d/T = 1.15$ and (dashed-dotted green) $d/T = 1.25$, as a function of the relative orientation between the atoms.

In particular, the specific dependence of the nonlocal term (2.43) on ϑ implies that atoms oriented along perpendicular axes will not be able to harvest any entanglement from the field, regardless of the amount of time they are interacting with it. Nevertheless, this claim needs to be qualified: what the ϑ dependence of Eq. (2.43) actually shows is that the atoms cannot harvest entanglement using the transition $1s \to 2p_z$ when the two $2p_z$ orbitals are perpendicular. Note how-

ever that, in principle, entanglement harvesting would be possible through the participation of the two atoms' $2p_x$ and $2p_y$ orbitals, which in general will not be perpendicular even when the two atoms' $2p_z$ orbitals are. Although, rigorously speaking, to assess the entanglement harvested by all the possible transitions to the $2p$ orbitals requires a higher-dimensional study, the symmetry of the problem and the perturbative nature of the interaction (which makes multipartite entanglement appear only at higher order in perturbation theory, which is beyond the scope of our study) allows us to already conclude that there exists an optimal orientation of the two atoms that maximizes entanglement harvesting: that in which the sum of the absolute values of the projections of (the director vectors of) all the axes of B's frame on axes of A's frame is maximal. This optimal case is obtained when the $2p$ atomic orbitals of B are oriented such that their axes of symmetry maximize their angular separation with respect to the $2p$ orbitals of A. Namely, the Euler angles for all 96 configurations[3] where the entanglement harvested is maximal are

$$(\psi,\,\vartheta,\,\varphi) = \begin{cases} \left(\frac{\pi}{4} + n\frac{\pi}{2},\, \vartheta_1,\, \frac{\pi}{4} + m\frac{\pi}{2}\right) \\ \left(\frac{\pi}{4} + n\frac{\pi}{2},\, \pi - \vartheta_1,\, \frac{\pi}{4} + m\frac{\pi}{2}\right) \\ \left(\psi_1 + n\frac{\pi}{2},\, \vartheta_2,\, \psi_1 - l\frac{\pi}{2}\right) \\ \left(\psi_2 + n\frac{\pi}{2},\, \vartheta_2,\, \psi_2 - l\frac{\pi}{2}\right) \end{cases} , \tag{2.45}$$

for $n,\, m = 0 \ldots 3$, $l = 1 \ldots 8$, $\vartheta_1 = \arccos(1/3)$, $\vartheta_2 = \arccos(2/3)$, $\psi_1 = \arctan(1/2)$ and $\psi_2 = \arctan(2)$. Conversely, entanglement harvesting would be minimum when every axis of one atom's reference frame is parallel to an axis of the other atom's frame. In such configuration each axis of B's frame has zero projection onto two of the axes of A's frame.

In Fig. 2.3 we show examples of configurations which minimize and maximize the entanglement harvested.

2.3.4 Quantitative analysis

After exploring the genuine characteristics of entanglement harvesting from electromagnetic fields with nonisotropic probes, we proceed to study the remaining features of entanglement harvesting from an electromagnetic field to two identical, hydrogenlike atoms.

[3] These 96 total configurations correspond to the 8 possibilities of choice of A's quadrant where to place B's z axis, for each of these, the 4 possibilities of choice of quadrant where to orient B's x axis, and finally, the 3 choices of A's axis with which to maximize the overlap of B's z axis.

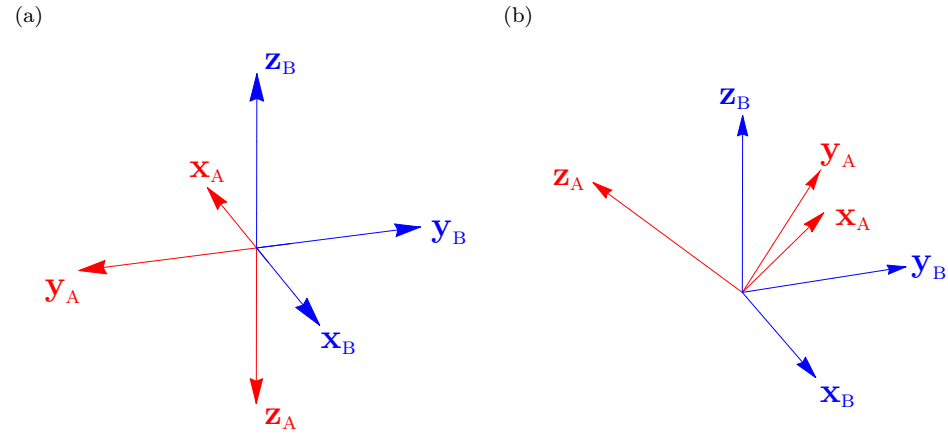

Figure 2.3: Examples of relative orientation of the atoms that (a) minimize and (b) maximize the entanglement harvested from the field when all the $1s \to 2p$ transitions are considered. The Euler angles for the configuration of maximal harvesting shown are $(\pi/4, \arccos(1/3), \pi/4)$.

We have the aim of having a model as realistic as possible, and for such we set realistic values for the parameters needed. The $1s \to 2p$ transition in the hydrogen atom has a wavelength of $\Lambda \sim 100$ nm, and the Bohr radius is $a_0 \sim 0.1$ nm. This gives (in natural units) $a_0\Omega = a_0\Lambda^{-1} = 0.001$. We will use these values throughout this section when computing the negativity from Eqs. (2.34), (2.42) and (2.43).

We first analyze how often it is possible to harvest entanglement from the field to the two atoms. For comparison, we recall [PKMM15], where it was shown that in the case of harvesting from scalar fields it is possible to harvest entanglement with arbitrarily distant detectors using a smooth switching and scaling up the energy gap consistently.

In Fig. 2.4 we show in dark red the values of energy gap ΩT of each atom, and the spatial distance d/T between them, for which entanglement can be harvested from the electromagnetic vacuum, regardless of its magnitude. In the figure, the delay between the switchings is set to $t_{BA}/T = 10$. Fig. 2.4 shows that harvesting of entanglement is possible even from regions that are arbitrarily spacelike-separated, provided that the atoms have an energy gap that is large enough: given an arbitrarily large distance between the atoms, one can consider a transition energy large enough so that the atoms can harvest entanglement. This is similar to the case discussed in [PKMM15] for the scalar-field Unruh-DeWitt entanglement harvesting: higher energy gaps reduce the impact of the local noise, making it possible to harvest entanglement for arbitrarily separated

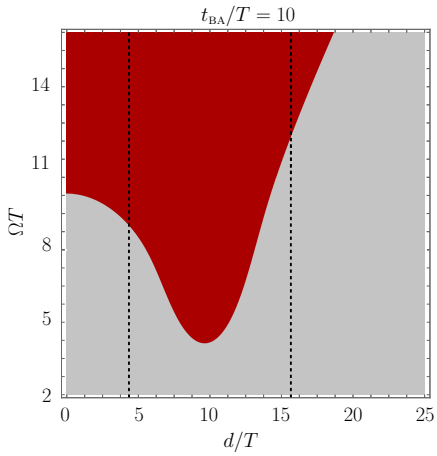

Figure 2.4: Binary plot showing in (dark) red the values of d/T and ΩT for which entanglement harvesting is possible, and in (light) grey the values for which there is zero entanglement harvesting. We consider $a_0\Omega = 0.001$ (hydrogen proportions) and that the atomic interactions with the field are delayed by a time delay $t_{BA}/T = 10$. The dashed black lines represent the boundaries of the regions of the plot where atom B is in the lightcone of atom A. We have located these boundaries at 8 standard deviations of the switching function (2.36), that is, $d = t_{BA} \pm 8T/\sqrt{2}$. Notice that we can harvest entanglement from arbitrarily far-away distances by increasing the transition energy gap.

atoms. To increase the gap of realistic atomic species one could consider higher energetic transitions $1s \to n\,p$, for which the product $a_0\Omega = 0.001$ does not scale with n.

However, increasing the gap to compensate for an increase of spatial separation has a downside: higher energy gaps also suppress the nonlocal correlation term $\mathcal{M}$ (albeit more slowly than $\mathcal{L}_{\mu\mu}$), thereby decreasing the total amount of entanglement harvested, leading to the same "damping-and-leakage" effect described for the Unruh-DeWitt case [PKMM15]. Despite this, one must note that it is possible to find configurations that harvest enough entanglement to violate Bell inequalities [Rez03], and even if the configuration does not allow for this, it is possible to use distillation protocols to obtain EPR pairs [Rez00].

In Fig. 2.5 we show the amount of entanglement that can be harvested from the field for a relatively large value of $\Omega T = 12$ as a function of the distance between the atomic centers of mass d/T and the time delay of their (recall local) interactions with the field t_{BA}/T. As expected given the results on harvesting from scalar fields [PKMM15], the entanglement is maximum inside the lightcone, where direct communication between the atoms can be achieved via exchange of real photons. Moving away from the lightcone the entanglement decreases

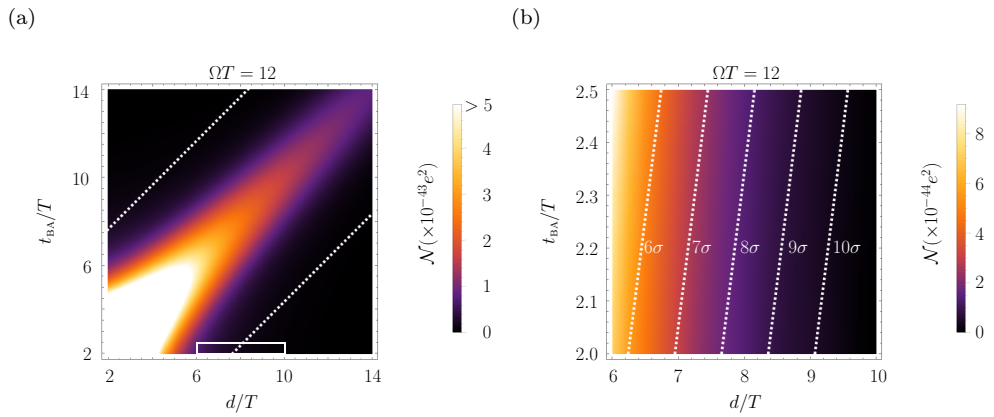

Figure 2.5: (a) Negativity to leading order as a function of spatial distance d and time delay between the atoms' interactions t_{BA} of two hydrogenoid atoms satisfying $a_0\Omega = 0.001$ for a fixed energy gap of $\Omega T = 12$. The white, dashed lines represent the boundaries of the lightcone, located at $d = t_{\mathrm{BA}} \pm 8T/\sqrt{2}$. (b) Zoom-in on the area marked with a white rectangle in (a), to demonstrate spacelike entanglement harvesting. The white dashed lines represent the distance to the point of light contact $d = t_{\mathrm{BA}}$ in number of standard deviations $\sigma = T/\sqrt{2}$. Both for the Gaussian switching and the compactly supported cropped version of the Gaussian switching at 8σ we observe spacelike entanglement harvesting.

with spatiotemporal distance, eventually leaking into the region of spacelike separation (shown in Fig. 2.5b). Entanglement harvesting also decays very quickly with the spatial separation of the atoms, as it can be seen in Figs. 2.2 and 2.5a.

One may perhaps be tempted to ascribe the harvested entanglement in the region deemed spacelike separated in Fig. 2.5b to the overlap of the atomic wavefunctions. After all, atomic orbitals are not compactly supported, and as such, the atoms are "never" completely spacelike separated. In Fig. 2.5b, we see that there is entanglement harvesting when the atoms are separated by 9σ [recall that $\sigma = T/\sqrt{2}$, with T being the interaction time scale as given in (2.36)]. At these points, the spatial separation of the atoms is $d \approx 10^4 a_0$ (where $a_0 = 0.001\Omega^{-1}$). With these numbers, the atomic wavefunction of atom A in the region of the center of mass of atom B is suppressed by an exponential factor e^{-10^4}, and indeed the overlap of the two wavefunctions yields $\int \mathrm{d}|\boldsymbol{x}||\boldsymbol{x}|^2 \psi^{\mathrm{A}}(|\boldsymbol{x}|)\psi^{\mathrm{B}}(|\boldsymbol{x}|) \approx 10^{-4343}$. Although the entanglement harvested $\mathcal{N}$ at this distance is very small, it is still several thousands of orders of magnitude too large to be explained by the overlap of the atomic wavefunctions. These results were obtained both for a Gaussian switching function and for a compactly-supported Gaussian-like switching function that is made zero at $t - t_\nu = \pm\, 8\sigma = \pm\, 8T/\sqrt{2}$ (that is, at 8 standard

deviations of the center of the Gaussian), so there is no appreciable light-like contact due to the switching either. Due to the absence of direct contact between the atoms when they are spacelike separated, the entanglement is really harvested from that existing previously in the electromagnetic field. The reason why this entanglement is small is because entanglement harvesting will only be relevant in atomic physics when the atoms are separated by several times the atomic sizes, as was already discussed in [MMS16].

2.3.5 Comparison with scalar models

Finally, we compare the electromagnetic coupling results with entanglement harvesting from scalar fields using the Unruh-DeWitt and derivative coupling models. Recall that in these scalar cases the smearing function needs to be introduced *ad hoc* in the model. In order to compare the electromagnetic coupling with the two scalar cases, Eqs. (2.18) and (2.19), we need to be mindful of the scalar nature of the coupling that prevents transfer of angular momentum through the field. As a consequence, in the cases of harvesting from scalar fields, the transitions $1s \to 2p$ are not allowed—in the same way that a $1s \to 2s$ transition is not allowed in the dipole electromagnetic case. Thus, in order to keep the comparison fair, we consider the most similar transition allowed by the scalar selection rules, which in this case would be between the two spherically symmetric levels of the hydrogenoid atom with lowest energy, namely the $1s$ and $2s$ states, whose position-representation wavefunctions are given by

$$\psi_g(\boldsymbol{x}) = \frac{1}{\sqrt{\pi a_0^3}} e^{-\frac{|\boldsymbol{x}|}{a_0}}, \tag{2.46}$$

$$\psi_e(\boldsymbol{x}) = \frac{1}{4\sqrt{2\pi a_0^3}} e^{-\frac{|\boldsymbol{x}|}{2a_0}} \left(2 - \frac{|\boldsymbol{x}|}{a_0}\right). \tag{2.47}$$

Although obvious, one way of seeing directly from these expressions that a transition between these two levels is forbidden in the electromagnetic dipole coupling is to realize that these two levels have equal (even) parity, so the spatial smearing vector Eq. (2.12) (which has odd overall parity) vanishes when integrated over all space.

As the smearing function in the scalar case we propose to use the scalar analogue to the smearing vector (2.12), which associates the smearing with the two-level wavefunctions. Such an appropriate definition for the scalar cases with overall even parity is

$$S(\boldsymbol{x}) = \psi_e(\boldsymbol{x})\psi_g(\boldsymbol{x}) = \frac{1}{4\pi a_0^3\sqrt{2}} e^{-\frac{3|\boldsymbol{x}|}{2a_0}} \left(2 - \frac{|\boldsymbol{x}|}{a_0}\right). \tag{2.48}$$

Considering this analysis, the two terms involved in Eq. (2.35) take the following form:

$$\mathcal{L}_{\mu\mu}^{\text{UDW}} = e^2 \frac{32768}{\pi} a_0^4 T^2 \int_0^\infty \mathrm{d}|\boldsymbol{k}| \frac{|\boldsymbol{k}|^5 e^{-\frac{1}{2}T^2(\Omega+|\boldsymbol{k}|)^2}}{\left(4a_0^2|\boldsymbol{k}|^2+9\right)^6}, \tag{2.49}$$

$$\begin{aligned}|\mathcal{M}|^{\text{UDW}} = e^2 \frac{16384}{\pi|\boldsymbol{x}_{\text{A}}-\boldsymbol{x}_{\text{B}}|} a_0^4 T^2 \Bigg| \int_0^\infty \mathrm{d}|\boldsymbol{k}|\, |\boldsymbol{k}|^4 e^{-\frac{1}{2}T^2\left(\Omega^2+|\boldsymbol{k}|^2\right)} \\ \times \frac{\sin(|\boldsymbol{k}||\boldsymbol{x}_{\text{A}}-\boldsymbol{x}_{\text{B}}|)}{\left(4a_0^2|\boldsymbol{k}|^2+9\right)^6} \left[E(|\boldsymbol{k}|, t_{\text{BA}}) + E(|\boldsymbol{k}|, -t_{\text{BA}})\right]\Bigg|,\end{aligned} \tag{2.50}$$

$$\mathcal{L}_{\mu\mu}^{\text{UDWd}} = e^2 \frac{32768}{\pi} a_0^4 T^2 \int_0^\infty \mathrm{d}|\boldsymbol{k}| \frac{|\boldsymbol{k}|^7 e^{-\frac{1}{2}T^2(\Omega+|\boldsymbol{k}|)^2}}{\left(4a_0^2|\boldsymbol{k}|^2+9\right)^6}, \tag{2.51}$$

$$\begin{aligned}|\mathcal{M}|^{\text{UDWd}} = e^2 \frac{16384}{\pi|\boldsymbol{x}_{\text{A}}-\boldsymbol{x}_{\text{B}}|} a_0^4 T^2 \Bigg| \int_0^\infty \mathrm{d}|\boldsymbol{k}|\, |\boldsymbol{k}|^6 e^{-\frac{1}{2}T^2\left(\Omega^2+|\boldsymbol{k}|^2\right)} \\ \times \frac{\sin(|\boldsymbol{k}||\boldsymbol{x}_{\text{A}}-\boldsymbol{x}_{\text{B}}|)}{\left(4a_0^2|\boldsymbol{k}|^2+9\right)^6} \left[E(|\boldsymbol{k}|, t_{\text{BA}}) + E(|\boldsymbol{k}|, -t_{\text{BA}})\right]\Bigg|.\end{aligned} \tag{2.52}$$

We see that these two pairs of expressions are very similar. Comparing on one hand Eqs. (2.49) and (2.51), and on the other hand Eqs. (2.50) and (2.52), we see that the expressions in each pair only differ in an extra factor of $|\boldsymbol{k}|^2$ on the integrals in momentum for the derivative coupling case. This could be easy to anticipate given the derivative nature of the coupling. Other than that, the expressions of the matrix elements of $\hat{\rho}_{\text{AB}}$ are identical in the Unruh-DeWitt model and in the derivative coupling model.

It is much more interesting to compare Eqs. (2.49), (2.51), (2.50) and (2.52) with their electromagnetic counterparts, Eqs. (2.42) and (2.43). There are not many significant differences in the local terms $\mathcal{L}_{\mu\mu}$: the only changes stem from the different dimensions of the couplings.

The case of the nonlocal term $\mathcal{M}$ is much more interesting, since it is here where the anisotropy of the excited levels and the vector nature of the electromagnetic field become apparent. The anisotropy of the excited levels can be seen through the factor $\cos\vartheta$ in Eq. (2.43), where, as discussed above, ϑ is the relative angle between the $2p_z$ orbitals of the two atoms [see Fig. 2.1, Eq. (2.41) and Appendix A.3].

In addition to the anisotropy of entanglement harvesting, the vector nature of the interaction manifests itself in the appearance of the spherical Bessel functions $j_0(|\boldsymbol{k}|d)$ and $j_2(|\boldsymbol{k}|d)$ in Eq. (2.43), substituting the factor $\sin(|\boldsymbol{k}|d)/(|\boldsymbol{k}|d)$

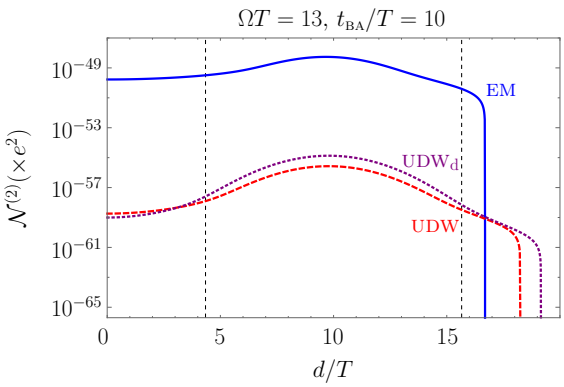

Figure 2.6: Negativity as a function of the spatial separation between the detectors in the different models studied: (blue, solid) hydrogen atom dipolarly coupled to the electromagnetic field, (red, dashed) monopole, spherically symmetric detector linearly coupled to a scalar field and (purple, dotted) scalar derivative coupling. In the electromagnetic coupling case, the atomic $2p$ orbitals are parallel. For all models, the energy gap is chosen to be $\Omega T = 13$, the detectors' size is given by $a_0\Omega = 0.001$ and the time delay between the switchings is $t_{BA}/T = 10$. The black dashed lines represent the boundaries of the lightcone, located at $d = t_{BA} \pm 8T/\sqrt{2}$. Electromagnetic harvesting is stronger, but has a shorter reach, than the scalar cases. Note that the sharp drop in all cases is due to the logarithmic scale.

(with $d = |\boldsymbol{x}_A - \boldsymbol{x}_B|$) that appears in both scalar cases. Notice that $\mathcal{M}$ is the $|g_A\rangle\langle e_A| \otimes |g_B\rangle\langle e_B|$ matrix element of the density matrix $\hat{\rho}_{AB}$, so the subindices $l = 0, 2$ of the spherical Bessel functions in the electromagnetic case correspond exactly with the even values of possible total angular momentum that one can obtain by combining the angular momentum of the dipole coupling term $\boldsymbol{d} \cdot \boldsymbol{E}$, $l_d = 1$, with the excited and ground atomic states' angular momenta $l_g = 0$, $l_e = 1$.

Furthermore, another way to see that the appearance of higher-order spherical Bessel functions is precisely the difference introduced by the exchange of angular momentum between the two atoms is the following: the j_0 summand in the integral (2.43) yields the same qualitative features of the scalar models, since $j_0(|\boldsymbol{k}|d) = \sin(|\boldsymbol{k}|d)/(|\boldsymbol{k}|d)$. We can wave our hands and roughly say that if we were to restrict the two atomic levels to be spherically symmetric ($l_g = l_e = 0$) and thus naïvely replace the coupling term by the scalar case (with a zero angular momentum $l_s = 0$) the only possible value of total angular momentum would be, of course, $l_e \oplus l_g \oplus l_s = 0$. Under this light, it is not surprising that we recover the same behavior as in the scalar cases just by simply neglecting the contribution of the higher-order Bessel functions in the electromagnetic case.

Indeed, in a rough qualitative analysis, comparing the results reported

in [PKMM15] with Figs. 2.4 and 2.5 we can conclude that the scalar models capture the essence of the phenomenon of harvesting if we do not pay attention to the effects introduced due to the exchange of angular momentum. However, under closer scrutiny, both the amount of entanglement harvested and the distance at which the atoms can still harvest entanglement are significantly different in the different models. As an illustration, in Fig. 2.6 we show the values of the negativity in the three models as a function of spatial separation between the detectors, for the same values of the energy gap $\Omega T = 13$ and time delay between the switchings $t_{\mathrm{BA}}/T = 10$.

Fig. 2.6 shows (i) that in the dipole model the amount of entanglement harvested from the field when the atoms' symmetry axes are parallel is much higher than in the scalar models, and (ii) in the scalar models it is possible to harvest entanglement when the atoms are further away than in the dipole, electromagnetic model. Regarding the two scalar models, there exist no major differences between them, which again points to the fact that the main differences appear when we incorporate the exchange of angular momentum to the model rather than when we change from an amplitude coupling to a derivative coupling.

2.4 Entanglement harvesting with gapless probes

In the following we focus on the special case when the probes used for harvesting entanglement have their two energy levels degenerate. This implies that there is no energy gap "shielding" from the local excitations created by the switching function, but neither the nonlocal excitations that entangle the detectors are suppressed. Therefore, a priori it is unclear whether gapless probes are better or worse at harvesting entanglement than gapped systems.

We will primarily focus on harvesting from scalar fields, although we will also deal with the example of harvesting from the electromagnetic field with zero-gap atomic transitions in Section 2.4.3. In the case of harvesting from scalar fields using the linear coupling of Eq. (2.18), Eqs. (2.29) and (2.30) simplify to

$$\begin{aligned}
\mathcal{L} &= e^2 \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2\, \mathcal{X}(t_1)\mathcal{X}(t_2) e^{\mathrm{i}\Omega(t_1-t_2)} \\
&\quad \times \int \mathrm{d}^n \boldsymbol{x}_1 \int \mathrm{d}^n \boldsymbol{x}_2\, S(\boldsymbol{x}_1) S(\boldsymbol{x}_2) W_n(t_2, \boldsymbol{x}_2, t_1, \boldsymbol{x}_1) \\
&= e^2 \int \mathrm{d}^n \boldsymbol{k} \frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|} \int_{-\infty}^{\infty} \mathrm{d}t_1\, \mathcal{X}(t_1) e^{\mathrm{i}(|\boldsymbol{k}|+\Omega)t_1} \int_{-\infty}^{\infty} \mathrm{d}t_2\, \mathcal{X}(t_2) e^{-\mathrm{i}(|\boldsymbol{k}|+\Omega)t_2} \\
&= e^2 \int \mathrm{d}^n \boldsymbol{k} \frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|} \left| \int_{-\infty}^{\infty} \mathrm{d}t\, \mathcal{X}(t) e^{\mathrm{i}(|\boldsymbol{k}|+\Omega)t} \right|^2 ,
\end{aligned} \tag{2.53}$$

$$\begin{aligned}\mathcal{M} =& -e^2 \int_{-\infty}^{\infty}\mathrm{d}t_1 \int_{-\infty}^{t_1}\mathrm{d}t_2 \int \mathrm{d}^n\boldsymbol{x}_1 \int \mathrm{d}^n\boldsymbol{x}_2 S(\boldsymbol{x}_1-\boldsymbol{x}_{\mathrm{A}})S(\boldsymbol{x}_2-\boldsymbol{x}_{\mathrm{B}})W_n(t_1,\boldsymbol{x}_1,t_2,\boldsymbol{x}_2)\\ &\times e^{\mathrm{i}\Omega(t_1+t_2)}\left[\mathcal{X}\left(t_1-t_{\mathrm{A}}\right)\mathcal{X}\left(t_2-t_{\mathrm{B}}\right)+\mathcal{X}\left(t_1-t_{\mathrm{B}}\right)\mathcal{X}\left(t_2-t_{\mathrm{A}}\right)\right]\\ =& -e^2 \int \mathrm{d}^n\boldsymbol{k}\,\frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|}e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}})}\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\, e^{-\mathrm{i}(|\boldsymbol{k}|-\Omega)t_1}e^{\mathrm{i}(|\boldsymbol{k}|+\Omega)t_2}\\ &\times\left[\mathcal{X}\left(t_1-t_{\mathrm{A}}\right)\mathcal{X}\left(t_2-t_{\mathrm{B}}\right)+\mathcal{X}\left(t_1-t_{\mathrm{B}}\right)\mathcal{X}\left(t_2-t_{\mathrm{A}}\right)\right],\end{aligned} \tag{2.54}$$

We want to consider gapless detectors, with $\Omega=0$, so that the monopole moment (2.6) becomes time-independent. We assume that both detectors are switched on and off with the same profile $\mathcal{X}$, which will have compact support, writing

$$\mathcal{X}\left(t\right)=\begin{cases}\chi(t) & \text{for } 0\leq t\leq T\\ 0 & \text{otherwise}\end{cases}, \tag{2.55}$$

where $T>0$ is the duration of the interaction of each detector with the field. We emphasize that the times t_ν, at which the interaction of each detector with the field begin, remain arbitrary.

2.4.1 Nonoverlapping switchings

When the domains of the switching functions do not overlap, the time integrals in the nonlocal term (2.54) simplify greatly. There are two summands in this term, which require separate study.

In the first summand the integrand is nonzero for $t_1\in[t_{\mathrm{A}},\,t_{\mathrm{A}}+T]$ and $t_2\in[t_{\mathrm{B}},\,t_1\leq\min(t_{\mathrm{A}},t_{\mathrm{B}})+T]$. Without loss of generality, let us assume that detector B is switched on after detector A has been switched off (i.e., $t_{\mathrm{B}}>t_{\mathrm{A}}+T$). In this case, because of the nested nature of the integrals, the region of integration over t_2 is limited by the support of $\mathcal{X}(t_1-t_{\mathrm{A}})$. Since detector B is switched on after detector A is switched off, the region of integration over t_2 lies out of the support of $\mathcal{X}(t_1-t_{\mathrm{A}})$, and therefore the integral evaluates to 0 regardless of the specific shape of $\chi(t_2)$.

In the second summand, in contrast, the integrand is supported in $t_1\in[t_{\mathrm{B}},\,t_{\mathrm{B}}+T]$ and $t_2\in[t_{\mathrm{A}},\,t_1\leq\min(t_{\mathrm{A}},t_{\mathrm{B}})+T]$. Now, in the case that detector B is switched on after detector A has been switched off, the effective region of integration over t_2 after taking into account the supports of $\mathcal{X}(t_1-t_{\mathrm{B}})$ and

$\mathcal{X}(t_2 - t_{\text{A}})$ is $[t_{\text{A}}, t_{\text{A}} + T]$. This means that we can decouple the two integrals,

$$\begin{aligned}\int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{t_1} \mathrm{d}t_2\, e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)} \mathcal{X}\left(t_1 - t_{\text{B}}\right) \mathcal{X}\left(t_2 - t_{\text{A}}\right) \\ = \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2\, e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)} \mathcal{X}\left(t_1 - t_{\text{B}}\right) \mathcal{X}\left(t_2 - t_{\text{A}}\right),\end{aligned} \tag{2.56}$$

where the equality follows because all the values of t_2 in the support of $\mathcal{X}\left(t_2 - t_{\text{A}}\right)$ are strictly smaller than the smallest value of t_1 in the support of $\mathcal{X}\left(t_1 - t_{\text{B}}\right)$.

Now, using the fact that the modulus of an integral is upper bounded by the integral of the modulus of the integrand, i.e.,

$$\left| \int \mathrm{d}x\, f(x) \right| \leq \int \mathrm{d}x\, |f(x)|\,, \tag{2.57}$$

we see that

$$\begin{aligned}|\mathcal{M}| &= e^2 \left| \int \mathrm{d}^n\boldsymbol{k}\, \frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|} e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_{\text{A}} - \boldsymbol{x}_{\text{B}})} \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2\, e^{\mathrm{i}|\boldsymbol{k}|(t_1-t_2)} \mathcal{X}(t_1 - t_{\text{B}}) \mathcal{X}(t_2 - t_{\text{A}}) \right| \\ &\leq e^2 \int \mathrm{d}^n\boldsymbol{k} \frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|} \left| \int_{-\infty}^{\infty} \mathrm{d}t \int_{-\infty}^{\infty} \mathrm{d}t'\, e^{\mathrm{i}|\boldsymbol{k}|(t-t')} \mathcal{X}(t) \mathcal{X}(t') \right| \\ &= \mathcal{L},\end{aligned} \tag{2.58}$$

where the changes of variables $t = t_1 - t_{\text{B}}$ and $t' = t_2 - t_{\text{A}}$ have been performed.

This yields the following conclusion: when there is no overlap between the time intervals of individual interactions with the field, the nonlocal, entangling term is always upper bounded by the local one and therefore $\mathcal{N}^{(2)} = \max\left(0, |\mathcal{M}| - \mathcal{L}\right) = 0$ for any—compactly supported or not—smearing function of the (recall gapless) detectors and any compactly supported, nonoverlapping switchings.

This means that gapless inertial comoving detectors with the same switching functions are unable to harvest any entanglement regardless of their relative positioning (spacelike, timelike or lightlike) from even arbitrarily close regions if they are switched on at different times with no overlap between the time intervals in which each individual detector interacts with the field.

It is important to stress that this is the case even for gapless detectors which are in regions that can be connected by light. This is true even if the smearing functions overlap (which means having effectively zero distance between the detectors).

Although this proof assumed that the switchings were the same for both detectors, numerical evidence for a generality of compactly supported switching

functions suggests that the detectors are unable to harvest entanglement also in the case of switchings of different duration $T_{\mathrm{A}} \neq T_{\mathrm{B}}$. This is true for detectors in timelike, spacelike or even lightlike separation.

Finally, notice that this proof carries over to the case of 1+1 dimensions if we add an infrared cutoff. Even with an infrared cutoff, the identity (2.57) still holds in the same way as in Eq. (2.58), so the inability of gapless detectors to harvest entanglement applies also to this case.

2.4.2 Overlapping switchings

We now explore the case when the time intervals of interaction overlap, either partially or totally. For this scenario, numerical evidence shows that entanglement harvesting is possible in general for timelike and lightlike separations, so we focus on the harvesting of entanglement from spacelike-separated regions and ask the following question: can two gapless detectors harvest entanglement from the field vacuum while they remain spacelike separated?

To talk properly about spacelike separation, we consider detectors with arbitrary, compactly supported smearings. Concretely, detectors A and B have finite characteristic lengths R_{A} and R_{B}, respectively. In analogy with Eq. (2.55), the smearing functions of the detectors are given by

$$S_\nu(\boldsymbol{x}) = \begin{cases} s_\nu(\boldsymbol{x}) & \text{for } |\boldsymbol{x}| \leq \frac{1}{2} R_\nu \\ 0 & \text{otherwise} \end{cases} . \tag{2.59}$$

For the following proof, we furthermore assume that the shapes of the detectors have spherical symmetry, which amounts to saying that their Fourier transform, as given by

$$\tilde{S}(\boldsymbol{k}) = \frac{1}{\sqrt{(2\pi)^n}} \int \mathrm{d}^n \boldsymbol{x}\, S(\boldsymbol{x}) e^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{2.60}$$

only depends on the norm of the Fourier variable $\boldsymbol{k}$. Explicitly, writing Eqs. (2.53) and (2.54) in spherical coordinates after setting $\Omega = 0$ leads to the expressions

$$\mathcal{L} = e^2 \int_0^\infty \mathrm{d}|\boldsymbol{k}| \int \mathrm{d}\Omega_{n-1} |\boldsymbol{k}|^{n-2} \frac{|\tilde{S}(|\boldsymbol{k}|)|^2}{2} \left| \int_{-\infty}^\infty \mathrm{d}t\, \mathcal{X}(t) e^{\mathrm{i}|\boldsymbol{k}|t} \right|^2 , \tag{2.61}$$

$$\begin{aligned} \mathcal{M} = -\, e^2 \int_0^\infty \mathrm{d}|\boldsymbol{k}| \int \mathrm{d}\Omega_{n-1} |\boldsymbol{k}|^{n-2} \frac{|\tilde{S}(|\boldsymbol{k}|)|^2}{2} e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_{\mathrm{A}} - \boldsymbol{x}_{\mathrm{B}})} \\ \int_{-\infty}^\infty \mathrm{d}t_1 \int_{-\infty}^{t_1} \mathrm{d}t_2\, e^{-\mathrm{i}|\boldsymbol{k}|(t_1 - t_2)} \left[\mathcal{X}(t_1 - t_{\mathrm{A}})\, \mathcal{X}(t_2 - t_{\mathrm{B}}) + \mathcal{X}(t_1 - t_{\mathrm{B}})\, \mathcal{X}(t_2 - t_{\mathrm{A}}) \right] . \end{aligned} \tag{2.62}$$

The spherical symmetry of the smearing allows one to perform the integration over the angular variables that appear in Eqs. (2.61) and (2.62). On one hand, the integrals in the local term (2.53) straightforwardly evaluate to the surface of the $(n-1)$-sphere, while on the other hand the integrals in the non-local term (2.54) are slightly less straightforward and are computed explicitly in Appendices A.4 and A.5. The resulting expressions for $\mathcal{L}$ and $\mathcal{M}$ are

$$\mathcal{L} = e^2 \int_0^\infty \mathrm{d}|\boldsymbol{k}||\boldsymbol{k}|^{n-2} |\tilde{S}(|\boldsymbol{k}|)|^2 \frac{\pi^{\frac{n}{2}}}{\Gamma\left(n/2\right)} \mathrm{Re}\left[\mathcal{T}_0(|\boldsymbol{k}|, T)\right], \tag{2.63}$$

$$\mathcal{M} = -\, e^2 \int_0^\infty \mathrm{d}|\boldsymbol{k}||\boldsymbol{k}|^{n-2} |\tilde{S}(|\boldsymbol{k}|)|^2 \frac{\pi^{\frac{n}{2}}}{\Gamma\left(n/2\right)} {}_0F_1\left(\frac{n}{2}; -\frac{|\boldsymbol{k}|^2 |\boldsymbol{x}_{\mathrm{A}} - \boldsymbol{x}_{\mathrm{B}}|^2}{4}\right) \mathcal{T}_{t_{\mathrm{BA}}}(|\boldsymbol{k}|, T), \tag{2.64}$$

where ${}_0F_1\left(a; z\right)$ is the confluent hypergeometric limit function [OOL$^+$], $\mathcal{T}_{t_{\mathrm{BA}}}(|\boldsymbol{k}|, T)$ is defined as

$$\begin{aligned}\mathcal{T}_{t_{\mathrm{BA}}}(|\boldsymbol{k}|, T) = \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2\, \theta(t_1 - t_2) e^{-\mathrm{i}|\boldsymbol{k}|(t_1 - t_2)} \\ \times \left[\mathcal{X}\left(t_1\right) \mathcal{X}\left(t_2 - t_{\mathrm{BA}}\right) + \mathcal{X}\left(t_1 - t_{\mathrm{BA}}\right) \mathcal{X}\left(t_2\right)\right],\end{aligned} \tag{2.65}$$

and $t_{\mathrm{BA}} = t_{\mathrm{B}} - t_{\mathrm{A}}$.

A crucial observation to prove that gapless detectors with overlapping interaction time intervals cannot harvest spacelike entanglement is that only the real part of the function $\mathcal{T}_{t_{\mathrm{BA}}}$ contributes to the evaluation of $\mathcal{M}$ when the detectors are spacelike separated. To see this, let us return to the expression of $\mathcal{M}$ in terms of the Wightman function in Eq. (2.54), which for gapless detectors is

$$\begin{aligned}\mathcal{M} = -\, e^2 \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{t_1} \mathrm{d}t_2 \int \mathrm{d}^n\boldsymbol{x}_1 \int \mathrm{d}^n\boldsymbol{x}_2 \\ \times S(\boldsymbol{x}_1 - \boldsymbol{x}_{\mathrm{A}}) S(\boldsymbol{x}_2 - \boldsymbol{x}_{\mathrm{B}}) W_n(t_1, \boldsymbol{x}_1, t_2, \boldsymbol{x}_2) \\ \times \left[\mathcal{X}\left(t_1 - t_{\mathrm{A}}\right) \mathcal{X}\left(t_2 - t_{\mathrm{B}}\right) + \mathcal{X}\left(t_1 - t_{\mathrm{B}}\right) \mathcal{X}\left(t_2 - t_{\mathrm{A}}\right)\right].\end{aligned} \tag{2.66}$$

Given that the smearing and switching functions are real, the only element that can make $\mathcal{M}$ complex is the Wightman function W_n. Remarkably, the imaginary part of the Wightman function $W_n(t, \boldsymbol{x}, t', \boldsymbol{x}')$ is proportional to (the expectation value of) the commutator of the field at the points $(t, \boldsymbol{x})$ and $(t', \boldsymbol{x}')$. Namely [see e.g, Eq. (23) in [MM15]],

$$\langle 0_{\hat{\phi}} | [\hat{\phi}(t, \boldsymbol{x}), \hat{\phi}(t', \boldsymbol{x}')] | 0_{\hat{\phi}} \rangle = 2\mathrm{i}\, \mathrm{Im}\big[W_n(t, \boldsymbol{x}, t', \boldsymbol{x}')\big]. \tag{2.67}$$

The commutator between field observables (and in particular, the field commutator), as mentioned in Section 2.2.2, satisfies microcausality, and so is only

supported inside their respective light cones. Therefore, for spacelike-separated regions the imaginary part of the Wightman function as given by Eq. (2.67) vanishes and the nonlocal term described by Eq. (2.66) is real. This means, from (2.66), that $\mathcal{M}$ is real.

Armed with this information about $\mathcal{M}$, we look at it in the form of Eq. (2.64). Since the hypergeometric functions in Eq. (2.64) are real and $\mathcal{M}$ itself is real, we conclude that only the real part of $\mathcal{T}_{t_{\mathrm{BA}}}(|\boldsymbol{k}|,T)$ contributes to $\mathcal{M}$. This allows us to replace $\mathcal{T}_{t_{\mathrm{BA}}}(|\boldsymbol{k}|,T)$ by $\mathrm{Re}\,[\mathcal{T}_{t_{\mathrm{BA}}}(|\boldsymbol{k}|,T)]$ for any switching and radially symmetric smearing under the condition that the detectors are spacelike separated.

Continuing with the proof, we show in Appendix A.6 that

$$\begin{aligned}\mathrm{Re}\,[\mathcal{T}_{t_{\mathrm{BA}}}(|\boldsymbol{k}|,T)] &= 2\pi|\tilde{\mathcal{X}}(|\boldsymbol{k}|)|^2\cos(|\boldsymbol{k}|t_{\mathrm{BA}})\\ &= \mathrm{Re}\,[\mathcal{T}_0(|\boldsymbol{k}|,T)]\cos(|\boldsymbol{k}|t_{\mathrm{BA}}).\end{aligned} \tag{2.68}$$

Since the confluent hypergeometric limit function satisfies (see 10.14.4 and 10.16.9 in [OOL$^+$])

$$\left|{}_0F_1\left(\alpha;-x^2\right)\right| \leq 1,$$

we obtain

$$\begin{aligned}|\mathcal{M}| &= e^2\left|\int_0^\infty \mathrm{d}|\boldsymbol{k}||\boldsymbol{k}|^{n-2}|\tilde{S}(|\boldsymbol{k}|)|^2\frac{2\pi^{\frac{n}{2}+1}}{\Gamma\left(n/2\right)}\right.\\ &\quad\left.\times\,{}_0F_1\left(\frac{n}{2};-\frac{|\boldsymbol{k}|^2|\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|^2}{4}\right)|\tilde{\mathcal{X}}(|\boldsymbol{k}|)|^2\cos(|\boldsymbol{k}|t_{\mathrm{BA}})\right|\\ &\leq e^2\int_0^\infty \mathrm{d}|\boldsymbol{k}||\boldsymbol{k}|^{n-2}|\tilde{S}(|\boldsymbol{k}|)|^2\frac{2\pi^{\frac{n}{2}+1}}{\Gamma\left(n/2\right)}|\tilde{\mathcal{X}}(|\boldsymbol{k}|)|^2\\ &= \mathcal{L}.\end{aligned} \tag{2.69}$$

This implies that $\mathcal{N}^{(2)}=0$ for gapless, spacelike-separated, spherically symmetric detectors for any—zero or nonzero—overlap between the time intervals of interaction of each detector with the field. Hence, combining the results of this section with those of Section 2.4.1, we see that gapless detectors with finite, spherically symmetric smearings interacting for a finite time with the field can *never* harvest entanglement from spacelike-separated regions, independently of the specific way of interacting with the field or their shape. This, of course, includes as a particular case the use of pointlike detectors, which is the case that is analyzed most often in the literature.

2.4.3 Gapless, hydrogenlike detectors

The results in Sections 2.4.1 and 2.4.2 assumed spherical symmetry on the smearing functions of the detectors. In this section we consider the realistic case of the light-matter interaction of an atomic electron in a hydrogenlike atom with the vacuum state of an electromagnetic field through a dipolar coupling. The orbitals of the hydrogen atom with angular momentum $l>0$ do not present spherical symmetry, so the general result of the previous sections may not apply in this case. Our study becomes particularly relevant for transitions between orbitals of the same quantum number n, which have zero energy gap.

The negativity acquired after interaction is given by Eq. (2.35) where the local $\mathcal{L}$ and nonlocal $\mathcal{M}$ terms become now [note these are just Eqs. (2.29) and (2.30) with $\Omega=0$]

$$\begin{aligned}\mathcal{L}^{\text{EM}} =& e^2 \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2\, \mathcal{X}(t_1)\mathcal{X}(t_2) \\ &\times \int \mathrm{d}^3\boldsymbol{x}_1 \int \mathrm{d}^3\boldsymbol{x}_2\, \boldsymbol{S}^{*\mathbf{t}}(\boldsymbol{x}_2)\mathbf{W}(t_2,\boldsymbol{x}_2,t_1,\boldsymbol{x}_1)\boldsymbol{S}(\boldsymbol{x}_1), \end{aligned} \tag{2.70}$$

$$\begin{aligned}\mathcal{M}^{\text{EM}} =& -e^2 \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{t_1} \mathrm{d}t_2 \int \mathrm{d}^3\boldsymbol{x}_1 \int \mathrm{d}^3\boldsymbol{x}_2 \\ &\times \Big[\mathcal{X}(t_1-t_{\text{A}})\mathcal{X}(t_2-t_{\text{B}})\boldsymbol{S}_{\text{A}}^{\mathbf{t}}(\boldsymbol{x}_1-\boldsymbol{x}_{\text{A}})\mathbf{W}(t_1,\boldsymbol{x}_1,t_2,\boldsymbol{x}_2)\boldsymbol{S}_{\text{B}}(\boldsymbol{x}_2-\boldsymbol{x}_{\text{B}}) \\ &+ \mathcal{X}(t_1-t_{\text{B}})\mathcal{X}(t_2-t_{\text{A}})\boldsymbol{S}_{\text{B}}^{\mathbf{t}}(\boldsymbol{x}_1-\boldsymbol{x}_{\text{B}})\mathbf{W}(t_1,\boldsymbol{x}_1,t_2,\boldsymbol{x}_2)\boldsymbol{S}_{\text{A}}(\boldsymbol{x}_2-\boldsymbol{x}_{\text{A}})\Big], \end{aligned} \tag{2.71}$$

where the spatial smearing vector $\boldsymbol{S}_\nu$ related to the ground and excited wavefunctions as per Eq. (2.17).

In the case of atomic switching functions that do not overlap, the reasoning used in Section 2.4.1 applies: the first summand of Eq. (2.71) evaluates to 0 and in the second summand the integrals in time decouple, making $|\mathcal{M}^{\text{EM}}|$ upper-bounded by $\mathcal{L}^{\text{EM}}$, regardless of the smearing vectors being compactly supported or not. This means that hydrogenlike atoms not interacting simultaneously with the vacuum cannot harvest any entanglement from it at all through transitions of zero energy.

When there is some overlap between the intervals of interaction of each individual atom with the field, the arguments used in Section 2.4.2 would also apply for hypothetical, compactly supported atoms: in this case, and since the electric field also satisfies microcausality, $\mathcal{M}^{\text{EM}}$ would also be real for spacelike separations between the compactly supported atoms. Then, without assuming spherical symmetry of the smearing functions, the hypergeometric function in

Eq. (2.64) is replaced by combinations of spherical Bessel functions. For example, for the zero-energy transition $2s \to 2p$ Eqs. (2.70) and (2.71) read

$$\mathcal{L}^{\text{EM}} = e^2 \frac{3a_0^2}{2\pi^2} \int_0^\infty \mathrm{d}|\boldsymbol{k}||\boldsymbol{k}|^3 \frac{(a_0^2|\boldsymbol{k}|^2-1)^2}{\left(a_0^2|\boldsymbol{k}|^2+1\right)^8} \mathrm{Re}\left[\mathcal{T}_0(|\boldsymbol{k}|,T)\right], \tag{2.72}$$

$$\begin{aligned}\mathcal{M}^{\text{EM}} = &- e^2 \frac{3a_0^2}{2\pi^2} \cos\vartheta \int_0^\infty \mathrm{d}|\boldsymbol{k}|\,|\boldsymbol{k}|^3 \frac{(a_0^2|\boldsymbol{k}|^2-1)^2}{\left(a_0^2|\boldsymbol{k}|^2+1\right)^8} \mathcal{T}_{t_{\text{BA}}}(|\boldsymbol{k}|,T) \\ &\times \left[j_0(|\boldsymbol{k}||\boldsymbol{x}_{\text{A}}-\boldsymbol{x}_{\text{B}}|) + j_2(|\boldsymbol{k}||\boldsymbol{x}_{\text{A}}-\boldsymbol{x}_{\text{B}}|)\right],\end{aligned} \tag{2.73}$$

where ϑ is the angle of the axis of symmetry of atom B's $2p$ orbital with respect to atom A's orbital.

Note that, despite the fact that the hypergeometric function appearing in the scalar nonlocal term [see Eq. (2.64)] has been replaced by a combination of spherical Bessel functions, this combination can still be upper bounded by 1. This means that also in this case the magnitude of the nonlocal term $|\mathcal{M}^{\text{EM}}|$ is upper bounded by the local term $\mathcal{L}^{\text{EM}}$, which means that no entanglement can be harvested from the electromagnetic vacuum with degenerate atomic probes if their radial functions are compactly supported. This argument contains, as a special case, that studied numerically in [Bra02, Bra05] where the atoms were assumed to be pointlike.

One must however note that the atomic wavefunctions of an electron in a hydrogenlike atom do not have compact support. Instead, the radial wavefunctions decay exponentially with the distance to the atomic center of mass. For this reason, one may be tempted to argue that the atoms can never be placed in spacelike-separated regions due to the always-existent overlap of their atomic wavefunctions, which will make the imaginary part of the Wightman function contribute, albeit suppressed by a factor of the overlap between the wavefunctions. For the implementation proposed in [Bra02] with two quantum dots separated by a distance of $d = 10\,\text{nm} \approx 190\,a_0$ (where a_0 is the Bohr radius), the overlap between the wavefunctions is on the order of $\int \mathrm{d}|\boldsymbol{x}||\boldsymbol{x}|^2 \psi^{\text{A}}(|\boldsymbol{x}|)\psi^{\text{B}}(|\boldsymbol{x}|) \approx e^{-190} \approx 10^{-83}$, which is definitely negligible when compared with the entanglement that gapped atoms could harvest at those distance scales. In the examples of Section 2.3.4 the atoms were declared effectively spacelike when separated by 10^4 Bohr radii and their interaction (with Gaussian switching) was short enough so that $10^4 a_0/c$ was more than 9 times the timescale of duration of the interaction. In that example, the overlap between the wavefunctions of the two atoms was of the order of 10^{-4343}, which is effectively 0 for all practical purposes. Since the harvesting of entanglement due to the atomic wavefunctions overlap is negligible, our results carry over to the light-atom interaction.

2.4.4 Instantaneous switchings

Finally, we explore the case in which gapped detectors interact for an infinitesimal amount of time with the field, but with an infinite strength. This case is relevant due to its similarities with a gapless detector case: in the case of a delta-like switching, during the time of interaction the free dynamics of the detectors is effectively halted (roughly speaking, the free Hamiltonian becomes negligible with respect to the infinite strength of the interaction Hamiltonian). This interaction is modeled by Dirac delta switching functions,

$$\mathcal{X}_\nu(t) = \eta\delta(t - t_\nu), \tag{2.74}$$

where η is a constant with dimensions of time. This switching will enable obtaining analytical closed-form expressions even for $\Omega \neq 0$.

For the switching function specified by Eq. (2.74) the local and nonlocal terms Eqs. (2.53) and (2.54) read

$$\begin{aligned}\mathcal{L} &= e^2\eta^2 \int \mathrm{d}^n\boldsymbol{k} \frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|} \left| \int_{-\infty}^{\infty} \mathrm{d}t\, \delta(t) e^{\mathrm{i}(|\boldsymbol{k}|+\Omega)t} \right|^2 \\ &= e^2\eta^2 \int \mathrm{d}^n\boldsymbol{k} \frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|}, \end{aligned} \tag{2.75}$$

$$\begin{aligned}\mathcal{M} = &- e^2\eta^2 \int \mathrm{d}^n\boldsymbol{k}\, \frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|} e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_\mathrm{A}-\boldsymbol{x}_\mathrm{B})} \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{t_1} \mathrm{d}t_2\, e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)} e^{\mathrm{i}\Omega(t_1+t_2)} \\ &\times \left[\delta\left(t_1 - t_\mathrm{A}\right)\delta\left(t_2 - t_\mathrm{B}\right) + \delta\left(t_1 - t_\mathrm{B}\right)\delta\left(t_2 - t_\mathrm{A}\right)\right]. \end{aligned} \tag{2.76}$$

In the case of nonsimultaneous switchings (i.e. when $t_\mathrm{A} \neq t_\mathrm{B}$) the argument in Section 2.4.1 used for evaluating the time integrals of the nonlocal term (2.76) applies: if detector B is switched on after detector A, the first summand evaluates to zero while in the second the integrals decouple. Integration over the time variables then leads to the expression

$$\mathcal{M} = -e^2\eta^2 \int \mathrm{d}^n\boldsymbol{k}\, \frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|} e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_\mathrm{A}-\boldsymbol{x}_\mathrm{B})} e^{-\mathrm{i}|\boldsymbol{k}|(t_\mathrm{B}-t_\mathrm{A})} e^{\mathrm{i}\Omega(t_\mathrm{B}+t_\mathrm{A})}, \tag{2.77}$$

and the magnitude of this expression satisfies

$$\begin{aligned}|\mathcal{M}| &= \left| e^2\eta^2 \int \mathrm{d}^n\boldsymbol{k}\, \frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|} e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_\mathrm{A}-\boldsymbol{x}_\mathrm{B})} e^{-\mathrm{i}|\boldsymbol{k}|(t_\mathrm{B}-t_\mathrm{A})} e^{\mathrm{i}\Omega(t_\mathrm{B}+t_\mathrm{A})} \right| \\ &\leq e^2\eta^2 \int \mathrm{d}^n\boldsymbol{k}\, \frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|} \left| e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_\mathrm{A}-\boldsymbol{x}_\mathrm{B})} e^{-\mathrm{i}|\boldsymbol{k}|(t_\mathrm{B}-t_\mathrm{A})} e^{\mathrm{i}\Omega(t_\mathrm{B}+t_\mathrm{A})} \right| \\ &= e^2\eta^2 \int \mathrm{d}^n\boldsymbol{k}\, \frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|} = \mathcal{L}, \end{aligned} \tag{2.78}$$

so, again in this case $\mathcal{N}^{(2)} = 0$, regardless of the specific shape of the detectors, their relative distance, and additionally now the energy gap.

When the individual interactions of the detectors with the field coincide, i.e. $t_{\mathrm{BA}} = 0$, Eq. (2.77) becomes mathematically ambiguous due to the argument of a Dirac delta coinciding with a limit of the integral. For sufficiently symmetric regularizations of the Dirac deltas, we however have that

$$2\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\, e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}e^{\mathrm{i}\Omega(t_1+t_2)}\delta(t_1-t_{\mathrm{A}})\delta(t_2-t_{\mathrm{A}}) = e^{2\mathrm{i}\Omega t_{\mathrm{A}}}. \tag{2.79}$$

In Appendix A.7 we give two examples of such regularizations. With the interpretation (2.79), the nonlocal term $\mathcal{M}$ becomes

$$\mathcal{M} = -e^2 e^{2\mathrm{i}\Omega t_{\mathrm{A}}}\int \mathrm{d}^n\boldsymbol{k}\,\frac{|\tilde{S}(\boldsymbol{k})|^2}{2|\boldsymbol{k}|}e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}})}. \tag{2.80}$$

Again, the magnitude of this term is bounded from above by the local term $\mathcal{L}$, so $\mathcal{N}^{(2)} = 0$ and entanglement harvesting is not possible in the limit when the switching becomes very short and intense, regardless of the shape or size of the probes, their relative distance or, in this specific case, the size of the gap between the energy levels of the detectors.

2.5 Discussion

In this chapter we have analyzed the harvesting of entanglement from the vacuum state of quantum fields. We have performed two main studies: entanglement harvesting from the electromagnetic vacuum using two fully-featured hydrogenlike atoms, and harvesting from scalar and vector fields with gapless probes.

We have developed a general model that is capable of including any arbitrary transition in the study of atomic entanglement harvesting from the electromagnetic vacuum. With it, we have considered the harvesting through the $1s \to 2p$ transitions of the hydrogen atom, and analyzed the unique features that emerge when considering the vector nature of the electromagnetic interaction in the atom-field dynamics—namely the effect of the relative orientation of the atoms on their ability to harvest entanglement, and the effect of the exchange of angular momentum between the atoms and the electromagnetic field. Furthermore, we have compared the proposed model with previous scalar models consistently used in the literature on entanglement harvesting.

When considering harvesting with zero-gap detectors, we have established several necessary conditions within leading order in perturbation theory for

harvesting to take place. Among other results, we have proved that spherically symmetric, compactly supported, gapless two-level systems in spacelike separation can never harvest any entanglement from scalar field vacua. The same results apply for hydrogenlike atoms harvesting through gapless transitions, such as the $2s \to 2p$ transition.

Our results make significant contributions to the final goal of an experimental observation of entanglement harvesting. Nevertheless, some important questions remain open. Recent studies [SMM18] have reported that one can harvest much more entanglement from highly squeezed states than from the vacuum state of a scalar field, and that, on the contrary (although expectedly, given the results reported in [Bro13]), the amount of entanglement harvested decays with the temperature of the field state. It will be of great importance for proposing an experimentally feasible observation scheme to study how these results generalize to the case of harvesting from electromagnetic field states. On the other hand, having a faithful model of the vacuum-probe interaction allows one to study the role of the vacuum that is implicitly present in all quantum information protocols, and which may have important consequences as demonstrated in [TBMM16]. Our results enable a realistic assessment of the effects of the electromagnetic vacuum in quantum information protocols, of which very little is known so far and thus represent interesting topics for future studies.

3 | Quantum information for thermodynamics

The origin of thermodynamics goes back to the XIX century motivated by the desire to increase the efficiency of early steam engines. In this setting, ideal assumptions could be made that, due to the macroscopicity of the systems involved, would allow one to make predictions that did not deviate significantly from the observations. Since the advent of quantum mechanics, it became of interest to understand the thermodynamic behavior of quantum systems, and it was quickly observed that the ideal assumptions that enabled macroscopic thermodynamics were much more difficult to fulfill in this setting.

This chapter addresses the effects of the removal of idealizations in quantum thermodynamic processes. For doing so, we make use of the formalism of Gaussian quantum mechanics, which we review at the beginning of the chapter. We propose a model for quantum thermodynamic interactions that allows one to go beyond the three canonical approximations in thermodynamics, namely it allows for strong couplings between the individual subsystems, for interactions taking finite time, and for specifying the microscopic properties of thermal baths, such as a finite number of components. In particular, we study the scenarios of a single harmonic oscillator interacting first with a large thermal bath in Section 3.3, and with two different baths in an alternating way, undergoing a strong-coupling adaptation of the Otto cycle in Section 3.4, analyzing the impact of nonideal assumptions and, notably, the inner dynamics of the baths.

3.1 Introduction

The theoretical model for a heat engine is very simple in spirit: it is composed of a working medium (WM) that reciprocates between two thermal baths, pumping heat from a hotter bath at temperature T_h to a colder one at temperature T_c and performing some mechanical work as a result.

The ideal engine converts the internal energy of the hot bath into work with an efficiency given by Carnot's formula, $\eta_C = 1 - T_c/T_h$. In order for the machine to operate at such an efficiency, three ideal requirements must be satisfied, namely: (i) the baths interact with the working medium weakly [Ali79, LL80],

(ii) the cycle is a quasiequilibrium process and hence it takes infinite time to complete [LL80, STH00, SST16], and (iii) the baths are infinitely large [LL80, RW14, WNW19, TH17, RAM18]. It has to be noted, however, that the size of the working medium itself is of no relevance—it can be anything from a two-level quantum system [SSD59, GK92, RDT$^+$16] to a giant steam engine [Cal85].

Strictly speaking, these conditions can never be satisfied: any interaction has a finite strength, any process that can be observed takes a finite time to complete, and infinite baths require of an infinite amount of energy and space to be created. Nevertheless, in the macroscopic regime the time, space and energy scales in which processes take place permit these idealizations to produce accurate predictions.

This is not the case in the microscopic regime where quantum effects play a role. This has not prevented (although not without issues [PLBH$^+$17]) the development of a theory of quantum thermodynamics. Nevertheless, despite the significant attention that finite-time [Rei29, CA75, VRMCH97, AJM08], strong-coupling [MP13, GRE14, FP14, GKAG15, EOG15], and finite-size [AHJM11, IO14, RW14, PSG16] effects have been getting either one by one or in groups of two, a rigorous microscopic analysis of a finite-power thermal machine strongly coupled to finite-sized heat baths has never been carried out. In this chapter we aim to fill this gap by performing a fully microscopic analysis of a heat engine consisting of a harmonic oscillator serving as a WM, reciprocating—by being alternately strongly coupled and decoupled—between two finite, initially thermal harmonic chains serving as thermal baths.

The engine runs a strong-coupling adaptation of the Otto cycle [Cal85]: the two "isochoric" thermalizations are intermediated by two "adiabatic" changes (all properly defined in Section 3.4) of the WM's Hamiltonian. At the beginning of the process, the WM starts uncoupled from the baths and at equilibrium with the cold bath. This makes the initial state of the overall system a Gaussian state. Given that the total Hamiltonian is quadratic at any moment of time, the dynamics of the system can be described within the formalism of Gaussian quantum mechanics (GQM) [AI07]. GQM maps the intractable Schrödinger equation in the infinite-dimensional Hilbert space of the overall system onto a linear evolution of the finite-dimensional phase space. This enables to perform a comprehensive analysis of the machine's operation without the need to adhere to any of the many approximations usually made when dealing with quantum open-system dynamics [Wei99, BP02]. Moreover, by directly simulating the overall system's evolution, we gain access to the states of the baths at any moment of time, which allows us to reveal the physical mechanisms governing the degradation and eventual exhaustion of the initial disequilibrium provided by the baths in the finite-size, finite-time, and strong-coupling regime. This

approach allows one to easily work with baths of size up to 300 times the size of the WM with just a standard table-top computer. Examples of computation of various evolutions and properties can be found in [PKHB17].

3.2 Preliminaries: Gaussian quantum mechanics

The ability of studying the microscopic details of large systems is achieved via the formalism of Gaussian quantum mechanics. In this section we review this formalism, focusing on the aspects necessary for our study, namely the evolution of systems and measures of classical and quantum correlations. For a much broader introduction to the topic, the reader is referred to [AI07]. Throughout this chapter all expressions are given in natural units, i.e., we assume $\hbar = k_B = 1$.

The primary computational advantage of Gaussian quantum mechanics is that it enables the study of interacting systems via a direct system-plus-bath perspective, without having to resort to perturbation theory [BMMMM13] or other open-systems techniques. This provides access to the exact evolution of the bath in addition to the system, a fact we take great advantage of.

3.2.1 The phase-space representation

Consider one or more quantum systems ascribed with bosonic canonical quadrature operators, satisfying the canonical commutation relations (CCRs), $[q_i, p_j] = \mathrm{i}\,\delta_{ij}$, where the indices label the systems (henceforth referred to as oscillators or modes). If one were to think about a harmonic oscillator with Hamiltonian $\frac{P^2}{2\mu} + \frac{\mu\omega^2 Q^2}{2}$, then a convenient choice of quadratures would be $q = Q\sqrt{\omega\mu}$ and $p = \frac{P}{\sqrt{\omega\mu}}$. In terms of the creation and annihilation operators, the quadratures are expressed through $q_i = (a_i + a_i^\dagger)/\sqrt{2}$ and $p_i = \mathrm{i}(a_i^\dagger - a_i)/\sqrt{2}$. For a system of N modes, the quadratures form a phase space that is represented as the vector of operators

$$\boldsymbol{x} = (q_1, p_1, \cdots, q_N, p_N).$$

Due to the CCRs, the phase space is a *symplectic space*, endowed with the structure $[x_a, x_b] = \mathrm{i}\,\Omega_{ab}$. Ω_{ab} are the components of the so-called symplectic form, given by

$$\boldsymbol{\Omega} = \bigoplus_{i=1}^{N} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

In GQM one works with Gaussian states. A state of an N-mode system is Gaussian if and only if it has the form of an exponential of a quadratic form in the quadratures $\{x_a\}_{a=1}^{2N}$. Importantly, thermal states of quadratic Hamiltonians fall within this class. The defining feature of Gaussian states is that they are fully described by the first and second moments of their quadratures, i.e., their mean position and their variances in phase space. The mean quadratures of all the states we consider in this work will be zero, and so the formalism further simplifies. We thus characterize the state of our system via the $2N \times 2N$ covariance matrix $\boldsymbol{\sigma}$, the entries of which are given by

$$\sigma_{ab} = \langle x_a x_b + x_b x_a \rangle = \mathrm{Tr}\left[\rho \left(x_a x_b + x_b x_a\right)\right]. \tag{3.1}$$

An important aspect of GQM is that creating ensembles and performing partial traces is trivial. This is due to working in phase space rather than in a Hilbert space, so partitions are represented as a direct sum rather than as a tensor product. Thus, any combined state of two systems A and B takes the form

$$\boldsymbol{\sigma}_{\mathrm{AB}} = \begin{pmatrix} \boldsymbol{\sigma}_{\mathrm{A}} & \boldsymbol{\gamma}_{\mathrm{AB}} \\ \boldsymbol{\gamma}_{\mathrm{AB}}^{\mathrm{T}} & \boldsymbol{\sigma}_{\mathrm{B}} \end{pmatrix}, \tag{3.2}$$

where $\boldsymbol{\sigma}_{\mathrm{A}}$ and $\boldsymbol{\sigma}_{\mathrm{B}}$ are the reduced states of systems A and B, respectively, and the matrix $\boldsymbol{\gamma}_{\mathrm{AB}}$ specifies the correlations between the systems. Recall from Chapter 2 that the superscript T denotes the operation of transposition.

A fact crucial for GQM is that any unitary evolution generated by a time-dependent Hamiltonian that is quadratic at any moment of time preserves the Gaussianity of a state [Sch86]. Any such unitary, U, that acts on the Hilbert space, corresponds to a linear symplectic transformation on the phase space of quadratures: $\boldsymbol{x} \to U^{\dagger}\boldsymbol{x}U = \boldsymbol{S}\boldsymbol{x}$, with $\boldsymbol{S}$ satisfying

$$\boldsymbol{S}\boldsymbol{\Omega}\boldsymbol{S}^{\mathrm{T}} = \boldsymbol{S}^{\mathrm{T}}\boldsymbol{\Omega}\boldsymbol{S} = \boldsymbol{\Omega}. \tag{3.3}$$

The symplecticity of $\boldsymbol{S}$, expressed by Eq. (3.3), ensures that the CCRs are preserved throughout the change of basis. On the level of the covariance matrix, it is easy to see that this transformation acts as

$$\boldsymbol{\sigma} \to \boldsymbol{\sigma'} = \boldsymbol{S}\boldsymbol{\sigma}\boldsymbol{S}^{\mathrm{T}}. \tag{3.4}$$

3.2.2 Energy and evolution

Another convenient aspect of GQM is that it allows one to compute average energies, evolve systems over time according to some time-dependent quadratic

Hamiltonian, and diagonalize systems into their normal-mode basis without ever referencing a Hilbert space object.

The average energy of a state represented by the covariance matrix $\boldsymbol{\sigma}$, with respect to a purely quadratic Hamiltonian $H = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{x}$, is given by

$$\langle H \rangle = \frac{1}{2}\,\mathrm{Tr}(\boldsymbol{F}\boldsymbol{\sigma}). \tag{3.5}$$

The symplectic (i.e., unitary in the Hilbert space) evolution matrix $\boldsymbol{S}(t)$ generated by this (in general, time-dependent) Hamiltonian obeys a Schrödinger-like equation:

$$\frac{d\boldsymbol{S}(t)}{dt} = \boldsymbol{\Omega}\boldsymbol{F}_s(t)\boldsymbol{S}(t), \tag{3.6}$$

where $\boldsymbol{F}_s = \boldsymbol{F} + \boldsymbol{F}^{\mathrm{T}}$. For a constant Hamiltonian the solution trivially takes the form $S(t) = \exp(\boldsymbol{\Omega}\boldsymbol{F}_s t)$, and for general, driven systems, the equation can be straightforwardly integrated by standard numerical techniques.

3.2.3 Thermal states and thermality

When speaking of a "free" system it is meant that one works in the basis that diagonalizes its Hamiltonian. This is called the normal-mode basis, in which the Hamiltonian takes the form

$$H_{\mathrm{free}} = \sum_{i=1}^{N} \omega_i a_i^\dagger a_i = \sum_{i=1}^{N} \frac{\omega_i}{2}(p_i^2 + q_i^2)$$

where, in the second equality, the zero-point energy is an ignored constant. In this basis, the corresponding matrix in phase space acquires a diagonal form: $\boldsymbol{F}_{\mathrm{free}} = \frac{1}{2}\,\mathrm{diag}(\omega_1, \omega_1, \omega_2, \omega_2, \cdots)$. By definition, the normal modes do not interact with each other, which implies that any thermal state on the entire system is given by the tensor product (in phase space, the direct sum) of the thermal states of the individual normal modes.

In general, the system may have couplings between pairs of modes (for example, between nearest neighbors), which appear as nondiagonal elements in the matrix $\boldsymbol{F}$. The normal-mode basis can be obtained by symplectically diagonalizing this matrix: $\boldsymbol{S}\boldsymbol{F}\boldsymbol{S}^{\mathrm{T}} = \boldsymbol{F}_{\mathrm{free}}$, where $\boldsymbol{S}$ is a symplectic matrix, and $\boldsymbol{F}_{\mathrm{free}}$ is diagonal as above.

In the normal-mode basis, the covariance matrices of the thermal states of a system are given by

$$\boldsymbol{\sigma}_T = \bigoplus_{i=1}^{N} \begin{pmatrix} \nu_i^{(th)} & 0 \\ 0 & \nu_i^{(th)} \end{pmatrix}, \quad \nu_i^{(th)} = \frac{e^{\omega_i/T}+1}{e^{\omega_i/T}-1}, \tag{3.7}$$

where ω_i are the normal frequencies. One can thus find the thermal covariance matrix of any interacting system by first identifying the normal basis, specifying the covariance matrix $\boldsymbol{\sigma}$ as above, and then applying the inverse transformation to this matrix to put it back into the physical-mode basis.

The values $\nu_i^{(th)}$ in Eq. (3.7) are referred to as the symplectic eigenvalues of the thermal state. In general, every Gaussian state of N modes has N symplectic eigenvalues ν_i, which are obtained by symplectically diagonalizing the covariance matrix: there always exists a symplectic matrix $\boldsymbol{S}$ such that

$$\boldsymbol{S}\boldsymbol{\sigma}\boldsymbol{S}^{\mathrm{T}} = \bigoplus_{i=1}^{N} \begin{pmatrix} \nu_i & 0 \\ 0 & \nu_i \end{pmatrix}. \tag{3.8}$$

The symplectic eigenvalues can be directly computed by taking the regular eigenvalues of the matrix $\mathrm{i}\boldsymbol{\Omega}\boldsymbol{\sigma}$, which come in $\pm\nu_i$ pairs.

Due to the strong couplings to the baths that will be considered, the state of the WM will in general acquire nondiagonal terms, thus moving out of thermality. For a single oscillator, however, there is an easy procedure to assign an effective temperature to the system. In order to prescribe an effective temperature to an arbitrary Gaussian state of a single oscillator, which can be described by the covariance matrix

$$\boldsymbol{\sigma}_m = \begin{pmatrix} \nu_1 & \kappa \\ \kappa & \nu_2 \end{pmatrix},$$

one can symplectically diagonalize it to $\hat{\boldsymbol{\sigma}}_m = \mathrm{diag}(\tilde{\nu}, \tilde{\nu})$, and after that compute an effective temperature via

$$T_{\mathrm{eff}} = \frac{\omega_m}{\ln \frac{\tilde{\nu}+1}{\tilde{\nu}-1}}. \tag{3.9}$$

The obtained temperature, which coincides with the real temperature in the case of thermal states, will be the effective temperature attributed to the original state described by $\boldsymbol{\sigma}_m$. This is the definition of T_{eff} that will be used throughout this chapter.

This procedure, however, is only acceptable if the thermal state at temperature T_{eff} is close to the real state of the system. This distance can be measured via the Uhlmann fidelity [NC10] between $\rho(\boldsymbol{\sigma}_m)$, the state described by the covariance matrix $\boldsymbol{\sigma}_m$, and the thermal state at its effective temperature, $\rho(\hat{\boldsymbol{\sigma}}_m)$. The Uhlmann fidelity is given by

$$F[\boldsymbol{\sigma}_m, \hat{\boldsymbol{\sigma}}_m] = \left[\mathrm{Tr}\sqrt{\sqrt{\rho(\boldsymbol{\sigma}_m)}\rho(\hat{\boldsymbol{\sigma}}_m)\sqrt{\rho(\boldsymbol{\sigma}_m)}}\right]^2.$$

This quantity is equal to 1 if and only if $\boldsymbol{\sigma}_m = \hat{\boldsymbol{\sigma}}_m$, and is < 1 otherwise.

For the specific case of Gaussian states, $F[\boldsymbol{\sigma}_m, \hat{\boldsymbol{\sigma}}_m]$ can be directly expressed in terms of the covariance matrices. Indeed, for purely quadratic states, it is given by [Scu98]

$$F[\boldsymbol{\sigma}_m, \hat{\boldsymbol{\sigma}}_m] = \frac{2}{\sqrt{A+B} - \sqrt{B}}, \tag{3.10}$$

where the quantities A and B are

$$A = 4 \det(\boldsymbol{\sigma}_m + \hat{\boldsymbol{\sigma}}_m), \tag{3.11}$$

$$B = (4 \det \boldsymbol{\sigma}_m - 1)(4 \det \hat{\boldsymbol{\sigma}}_m - 1). \tag{3.12}$$

Instead of just the Uhlmann fidelity, in Sections 3.3 and 3.4 we will study the evolution in time of the following measure of athermality:

$$\mathcal{A}(t) = 1 - F[\boldsymbol{\sigma}_m(t), \hat{\boldsymbol{\sigma}}_m(t)]. \tag{3.13}$$

Due to the properties of the Uhlmann fidelity, $0 \leq \mathcal{A}(t) \leq 1$ and $\mathcal{A}(t) = 0$ if and only if $\boldsymbol{\sigma}_m(t) = \hat{\boldsymbol{\sigma}}_m(t)$, i.e., if the state of the system is really thermal.

3.2.4 Entropy and correlations

Consider a two-party state of the form of Eq. (3.2). The off-diagonal matrix $\boldsymbol{\gamma}_{\text{AB}}$ contains the correlation functions between the two systems, and these systems are uncorrelated if and only if $\boldsymbol{\gamma}_{\text{AB}} = 0$. A good quantitative measure of general correlations is the mutual information, defined as

$$I(A, B) = S(\boldsymbol{\sigma}_A) + S(\boldsymbol{\sigma}_B) - S(\boldsymbol{\sigma}_{AB}). \tag{3.14}$$

Here, $S(\boldsymbol{\sigma})$ is the von Neumann entropy of the state with covariance matrix $\boldsymbol{\sigma}$, given by [HW01]

$$S(\boldsymbol{\sigma}) = \sum_{i=1}^{N} f(\nu_i), \tag{3.15}$$

where

$$f(\nu) = \frac{\nu+1}{2} \log \frac{\nu+1}{2} - \frac{\nu-1}{2} \log \frac{\nu-1}{2} \tag{3.16}$$

and $\{\nu_i\}_i$ is the set of symplectic eigenvalues of $\boldsymbol{\sigma}$. This shows that the symplectic eigenvalues of a state—which are invariant under symplectic transformations—give a measure of mixedness for that state. For example, the entropy is zero, i.e., a Gaussian state is pure, if and only if all its symplectic eigenvalues are equal to one. This makes the computation of mutual information across any partition of multipartite systems a very easy task, independently of how many modes each partition contains.

Entanglement across bipartitions is also computable, but, for most situations, it is considerably more difficult. In the particular case of two modes it nevertheless can be easily quantified [Bro13]. In such a case, the logarithmic negativity, given by

$$E_N = \max(0, -\log \tilde{\nu}_-), \tag{3.17}$$

where $\tilde{\nu}_-$ is the smallest of the symplectic eigenvalues of the *partially-transposed* state, can be computed from the formula

$$2\tilde{\nu}_-^2 = \Delta - \sqrt{\Delta^2 - 4 \det \boldsymbol{\sigma}_{AB}}, \tag{3.18}$$

where $\Delta = \det \boldsymbol{\sigma}_{\mathrm{A}} + \det \boldsymbol{\sigma}_{\mathrm{B}} - 2 \det \boldsymbol{\gamma}_{\mathrm{AB}}$.

As it has been discussed already in Chapter 2, quantum correlations are hard to maintain in the presence of a nonzero temperature. In the scenarios studied in the following sections we find that generally entanglement does not play a significant role. Interestingly, this aspect is in accord with (yet by no means logically necessitated by) the fact that, although capable of manifesting many interesting quantum features, GQM is an essentially classical, noncontextual sector of quantum mechanics in that it can be described by a local hidden variable model [BRS12].

3.3 Gaussian interaction with a single bath

Before performing the analysis of the full Otto cycle, we begin by studying some relevant features of the isochoric interaction of the WM with a single thermal bath. In this section we thereby introduce the specific Hamiltonians that describe the components of the Otto engine that will be analyzed in Section 3.4.

From now on, we model thermal baths as collections of harmonic oscillators arranged in one-dimensional, translation-invariant rings with nearest-neighbor interactions. We consider only position-position couplings so that the free Hamiltonian of a bath is given by

$$H_{\mathrm{bath}} = \sum_{i=1}^{N} \frac{\omega_b}{2} \left(p_i^2 + q_i^2\right) + \sum_{i=1}^{N} \alpha q_i q_{i+1}, \tag{3.19}$$

where N is the number of oscillators in the bath, ω_b is the bare frequency of each of them, and α controls the coupling strength between the different oscillators in the chain. Because of the periodic boundary conditions, $q_{N+1} = q_1$.

The phase-space matrix corresponding to this Hamiltonian is

$$\boldsymbol{F}_{\text{bath}} = \frac{1}{2}\begin{pmatrix} \boldsymbol{\omega}_b & \boldsymbol{\alpha} & \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha} & \boldsymbol{\omega}_b & \boldsymbol{\alpha} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\alpha} & \boldsymbol{\omega}_b & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{\alpha} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\alpha} & \boldsymbol{\omega}_b \end{pmatrix}, \tag{3.20}$$

where $\mathbf{0}$ is the 2×2 matrix of zeros, and

$$\boldsymbol{\omega}_b = \begin{pmatrix} \omega_b & 0 \\ 0 & \omega_b \end{pmatrix}, \quad \boldsymbol{\alpha} = \begin{pmatrix} \alpha & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.21}$$

At the beginning of the interaction, the bath is assumed to be in a thermal state at temperature T_b, this is

$$\rho(0) \propto e^{-H_{\text{bath}}/T_b}. \tag{3.22}$$

Due to the interactions between the oscillators that form it, the covariance matrix that characterizes $\rho(0)$ will *not* be given by a simple direct sum as in Eq. (3.7). As described in Section 3.2, one must begin by identifying the normal mode basis that symplectically diagonalizes the Hamiltonian matrix, $\boldsymbol{\omega}_{\text{bath}}/2 = \boldsymbol{S}\boldsymbol{F}_{\text{bath}}\boldsymbol{S}^{\mathrm{T}}$. After that, the thermal state can be created as per Eq. (3.7), and then this state is transformed back to the physical basis in order to obtain the thermal state of the ring, $\boldsymbol{\sigma}_{\text{bath}} = \boldsymbol{S}^{-1}\boldsymbol{\sigma}_T(\boldsymbol{S}^{-1})^{\mathrm{T}}$.

As a WM we employ yet another harmonic oscillator, with bare frequency ω_m. Its coupling to the bath is described by

$$H_{\text{int}} = \gamma\chi(t)\, q_m \sum_{i\in\mathcal{I}} q_i \tag{3.23}$$

where $\mathcal{I}$ is the set of nodes in the bath which will interact with the WM. For most of our analysis, we will choose this set to contain just one and always the same node of the bath, which we label the *first* or *interacting node* q_1. However, the effect of adding more oscillators in $\mathcal{I}$ is studied in Section 3.4.1.

The function $\chi(t)$ is a *switching function* that modulates the interaction in time. In particular, we choose the following compactly supported, smooth switching function

$$\chi(t) = \begin{cases} 0 & t < 0 \\ \frac{1}{2}\left(1 - \tanh\cot\frac{\pi t}{\delta}\right) & 0 \le t < \delta \\ 1 & \delta \le t < \tau - \delta \\ \frac{1}{2}\left(1 + \tanh\cot\frac{\pi(t-\tau)}{\delta}\right) & \tau - \delta \le t < \tau \\ 0 & t > \tau \end{cases}, \tag{3.24}$$

where $\tau \geq 2\delta$ is the total duration of the interaction with the bath and δ is the time that takes to fully switch on (and fully switch off) the interaction. We will refer to δ as the *ramp-up time* of the isochoric interaction.

The phase-space matrix of the overall Hamiltonian will then be

$$\boldsymbol{F}_{\text{tot}} = \begin{pmatrix} \frac{1}{2}\boldsymbol{\omega}_m & \frac{1}{2}\boldsymbol{\gamma} \\ \frac{1}{2}\boldsymbol{\gamma}^{\text{T}} & \boldsymbol{F}_{\text{bath}} \end{pmatrix}. \tag{3.25}$$

Here $\boldsymbol{\gamma}$ is a $2 \times 2N$ matrix containing all zeros except for the first entry $\gamma_{11} = \chi(t)\gamma$, corresponding to the $q_m q_1$ interaction we are imposing. In order to construct the symplectic evolution matrix $\boldsymbol{S}(t)$ of the overall system, which is generated by $\boldsymbol{F}_{\text{tot}}$, we numerically integrate Eq. (3.6). The covariance matrix at instant t will then be simply given by

$$\boldsymbol{\sigma}_{\text{tot}}(t) = \boldsymbol{S}(t)(\boldsymbol{\sigma}_m \oplus \boldsymbol{\sigma}_{\text{bath}})\boldsymbol{S}(t)^{\text{T}}, \tag{3.26}$$

where $\boldsymbol{\sigma}_m = \text{diag}(\nu^{(m)}, \nu^{(m)})$ is the initial state of the WM, and the values of $\nu^{(m)}$ are given by Eq. (3.7).

3.3.1 Thermalization under strong coupling and finite time

During the evolution of the overall—WM plus bath—system, we observe that the state of the WM remains very close to being thermal. That is, at any moment of time, its covariance matrix is very close to that given by Eq. (3.7) for some $\nu^{(th)}$. This can be observed in Fig. 3.1, where it is shown that, with the choice of parameters that will be used throughout, this deviation is negligible, specially at the end of the interaction. Given this, we are able to assign a meaningful effective temperature to the WM by computing the temperature associated with its symplectic eigenvalue.

The evolution of this effective temperature is shown in the green, solid line of Fig. 3.2. Importantly, we notice that at $t \approx 93$ the temperature of the WM becomes equal to that of the bath. It must be noted that this moment is not the thermalization time in the proper sense because the interaction is still on. Rather, the exact thermalization time, defined as the time needed for the bath and the WM to have the same temperature *after the interaction has been turned off*, is $\tau_{th} \approx 98.5$. It turns out that, in order to achieve precise thermalization, one needs to match the frequencies, so that ω_m is also the frequency of the individual bath oscillators [the ω_b in Eq. (3.19)], and the couplings, so that the WM-baths interaction strength is equal to the intra-ring coupling strength, i.e., $\gamma = \alpha$. Intuitively, this matching ensures that the rate of information transfer between the WM and the bath is the same as between oscillators within the

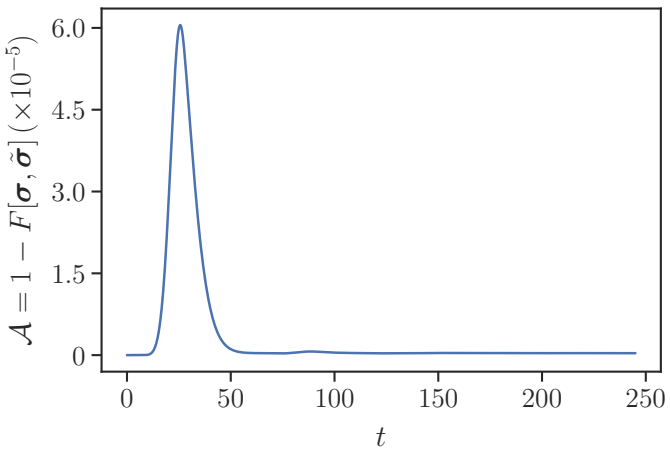

Figure 3.1: Time-dependence of the athermality of the WM, initially thermal at temperature $T_b = 0.5$, for the isochoric interaction with a bath (at temperature $T_b = 4$) consisting of $N = 30$ oscillators with frequencies $\omega_b = \omega_m = 2$. The coupling constants are matched so that $\gamma = \alpha = 0.1$ and the interaction lasts $\tau = 245$ units of time, with the ramp-up time δ being 0.1τ.

bath. Whenever ω_m (resp. γ) is outside a small neighborhood of ω_b (resp. α), the WM does not thermalize with the bath at all. A similar frequency-filtering phenomenon in the classical setting was reported in [SO08]. In view of this, we from now on set $\omega_m = \omega_b$ and $\gamma = \alpha$.

With such a configuration of the parameters, and fixed δ/τ, the thermalization time τ_{th} scales as

$$\tau_{th} \propto \alpha^{-1} \tag{3.27}$$

for $\alpha \ll 1$. In fact, the above scaling is, to a good approximation, preserved also for large α. A remark is in order here. As τ_{th}, we choose the smallest τ that achieves thermalization. Since the bath is finite and the WM couples to it strongly, the temperature of the WM will not approach the bath's temperature T_b in a monotonic way: with passing time, the WM's final temperature will first go slightly above T_b—the maximum being $T_b + \mathcal{O}(\alpha^2)$ [1]—then go below T_b, and continue an oscillatory behavior as that depicted in Fig. 3.2 for $t > 100$.

Another important aspect of thermalization process observed is that, due to the finite duration and the finite strength of the WM-bath interaction, it has a nonzero work cost. More specifically, the extracted work, as quantified by the

[1] This is due to the fact that, because of the interaction, the local effective temperature of a bath node is slightly above the global temperature

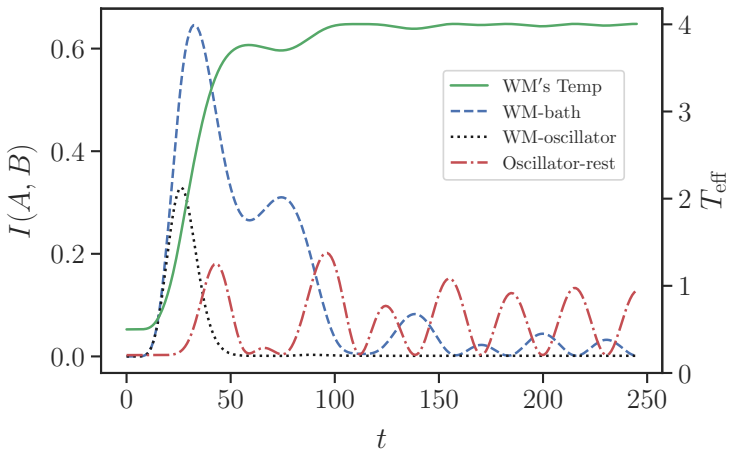

Figure 3.2: Evolution of several quantities during a period of interaction between the WM and a thermal bath. The green solid line is the WM's effective temperature. The other lines represent the mutual information between various partitions: (dashed blue) between the WM and the bath as a whole, (dotted black) between the WM and the specific bath oscillator with which it interacts, and (dot-dashed red) between this oscillator and the rest of the bath. The bath contains $N = 30$ oscillators, all with frequency $\omega_b = 2$, and is initialized in a thermal state at temperature $T_b = 4$. The WM also has frequency $\omega_m = \omega_b$, but is initialized in a thermal state at temperature $T_m = 0.5$. The total time of interaction is $\tau = 245$, the ramp-up time is $\delta = 0.1\tau$, and the interaction strengths are $\alpha = \gamma = 0.1$. For these parameters, the exact thermalization time as defined in the main text is $\tau_{th} = 98.5$. Note that the red curve is not initially zero (and it should not be, because there is initial correlation from the ring couplings). However, since we are working with a relatively hot bath, these correlations are very small (of the order of 10^{-3}), and its magnitude cannot be appreciated in full detail in the figure.

difference between the initial and final average energies of the total system, is not zero. However, despite the strong nonequilibrium character of the process, this amount is small (compared to, e.g., the energy exchanged between the WM and the bath). In fact, for small α, this work cost, W_i (where the subscript $_i$ stands for isochoric), scales as

$$W_i \propto \alpha^2, \tag{3.28}$$

and is almost independent of the ramp-up time, δ. Taking, for example, $N = 30$, $\gamma = \alpha = 0.1$, $\omega_m = \omega_b = 2$, $T_b = 4$, $T_m = 0.5$, $\tau = 100$, and $\delta = 0.1\tau$, we get $W_i \approx -6.2 \times 10^{-3}$, while the exchanged heat is $\approx, 3$. This, together with the fact of exact thermalization discussed above, means that the fine-tuning of the frequencies and couplings provides us with an example of almost work-free thermalization in finite time, resulting from a strong interaction between the WM and the bath. A similar example, where the structure of the bath is known and the Hamiltonian

of the WM is finely-tuned, was constructed in [AHM10]. This is not a standard, exponential relaxation behavior [BP02], and it can be argued that such behavior cannot occur for general baths of unknown structure [AHMG13].

We also note that the fact of almost zero work justifies the usage of the term "isochoric" for this process. Indeed, in this case, most of the energy exchange is heat transfer, which is the characteristic of isochoric processes [Cal85]. In the strong coupling regime, strictly isochoric (or, equivalently, constant-Hamiltonian processes) cannot exist as any nonzero coupling will change the system Hamiltonian, and therefore the term needs to be adapted.

3.3.2 Evolution of correlations

Furthermore, subtle processes such as the evolution of the correlations between the WM and the bath or information exchange between the WM and the bath can be examined in very great detail within the framework of GQM. As an illustration, in Fig. 3.2 we examine the evolution of the correlations between various partitions during the interaction.

We compare the correlations, as measured by the mutual information, between the WM and the whole bath (dashed, blue line), between the WM and the node in the bath q_1 it interacts with (black, dotted line), and between the latter and the rest of the bath (red, dot-dashed line). This provides a number of insights into the nonperturbative interaction of the WM and the bath. First, during the phase of switching on the interaction, the WM and the bath quickly build up strong correlations, which decay later on. This decay is caused by the fact that the bath nodes to which the WM is coupled also interact with the rest of the bath, and the bath, due to its tendency to thermalize, forces these correlations to decay. We can see this process in more detail by examining the other two lines. The correlation between the machine and the interacting node similarly rises and then falls, and the decay occurs exactly as this node becomes significantly correlated with the rest of the bath.

This gives an important intuitive picture. The interaction between the WM and the bath generates correlations between the two (specifically, between the WM and the interacting node). Due to the intra-bath couplings, the WM also becomes correlated with other ring nodes in an outwards-propagating manner. However, these couplings also mean that the correlation between the WM and bath will, over time, be swapped to correlations between different bath nodes, as it can be seen, for example, in the red dot-dashed line in Fig. 3.2. Over the course of many interaction sessions, the bath nodes therefore become more and more intercorrelated, which will eventually result in a halt of the machine when running thermodynamic cycles. We elaborate on this process in the next

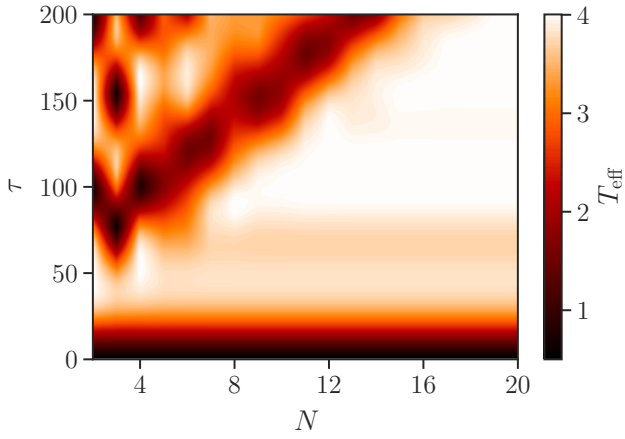

Figure 3.3: Effective temperature of the WM after the interaction with the bath as a function of the bath size N and the time of interaction τ. The parameters for the WM are those used also for Fig. 3.2 ($\omega_m = 2$, $T_m = 0.5$), and similarly for the relevant parameters of the bath ($\omega_b = 2$, $T_b = 4$, $\alpha = \gamma = 0.1$). The ramp-up time of the interaction is $\delta = 0.1\tau$ for every value of τ. Note the two distinct behaviors separated by a straight line $\tau = c \cdot N$.

section, where we discuss the performance of a WM operating cyclically between two finite-sized baths.

3.3.3 Finite-size effects

One does not need to move to the two-bath scenario to observe the effects of having finite-sized baths, though. In fact, one only needs to interact with the bath for a time that is long enough. We do so in Fig. 3.3, where we compute the effective temperature of the machine after interacting with a bath composed of N nodes during a time τ, for different values of N and τ, and all the other parameters being the same as those used for Fig. 3.2.

In Fig. 3.3, we observe two very distinct behaviors that are clearly separated. For $\tau < c \cdot N$, where c indicates the slope of the "causal cone", the temperature of the WM is insensitive to the size of the bath. Indeed, the interaction time in this case is short enough so as to allow the interaction to finish before the perturbations that propagate through the bath return to the region which interacts with the WM (i.e., the interacting node of the ring). Therefore, there is no difference between the temperature that the WM achieves in this case and the temperature that it would achieve from interacting with an infinite bath. The opposite occurs for $\tau > c \cdot N$: in this case, the interaction time is long enough so as to permit the perturbations generated by the interaction with the WM to

return to the interacting node. These perturbations modify the local state of the interacting node, which in turn translates into a response in the WM that diverges from that expected for infinite baths.

It is also worth noting that, for short interaction times, the WM does not have enough time to fully thermalize with the bath. We observe that the effective temperature of the WM increases with the interaction time until the point where thermalization is achieved. After this point, increasing the interaction time further has no major influence on the WM's temperature until, of course, it is long enough for the perturbations to go around the bath.

3.4 The Gaussian Otto cycle

We now study the performance of the WM running an Otto cycle between two thermal baths at temperatures T_h (hot) and T_c (cold), as depicted in Fig. 3.4.

The Otto cycle we consider is composed of two isochoric interactions between the machine and each of the baths (as described in Section 3.3) separated by two sudden changes of the WM's Hamiltonian. Specifically, between subsequent interactions, we instantaneously swap the WM's Hamiltonian,

$$H_m = \omega_c a_m^\dagger a_m \longleftrightarrow H'_m = \omega_h a_m^\dagger a_m, \tag{3.29}$$

so that the WM's state remains unchanged. The fact that the WM is detached from the baths during the swap ensures that the process is adiabatic, i.e., thermally isolated, in the thermodynamic sense[2]. It is important to note that the change in Eq. (3.29) is not equivalent to simply quenching the frequency of the oscillator. Rather, it requires simultaneously changing both the frequency and the mass: $\omega_c \to \omega_h$ and $\mu \to \mu\frac{\omega_c}{\omega_h}$. Here, ω_c and ω_h are the frequencies of the WM used during the interactions with the cold and hot baths, respectively. For a discussion of the case when only the frequency is quenched, see Appendix B.1. As mentioned in Section 3.3, ω_c and ω_h are chosen to coincide with the frequencies of the nodes of, respectively, the cold and the hot baths. Moreover, we also match the interaction strength with the ring coupling strengths, i.e., $\gamma = \alpha_c = \alpha_h$ (we choose $\alpha_c = \alpha_h$ for simplicity only, without loss of generality).

The total work extracted during a cycle is given by the sum of works extracted during each of the four parts of the cycle. As shown in Section 3.3, the work contributions from WM-bath interactions are small, hence most of the work is generated during the adiabats. The work produced by a sudden change in Hamiltonian depicted in Eq. (3.29) is given by a particularly simple

[2] Moreover, the fact that the eigenbasis of the Hamiltonian remains unchanged additionally ensures that the process is adiabatic also in the sense of the quantum adiabatic theorem.

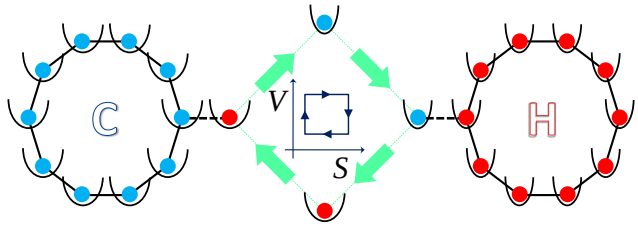

Figure 3.4: Visualization of the Otto cycle. The standard sequence of isochoric thermalizations and adiabatic compressions/expansions defining the Otto cycle in phenomenological thermodynamics are, in our case, implemented as a sequence of $q-q$ interactions (as described in Section 3.3) and sudden changes of the WM's Hamiltonian. More specifically, the cycle consists of the following steps: (i) the WM interacts with the hot bath (the red harmonic chain) by a coupling that is smoothly switched on, kept constant, and smoothly switched off, (ii) the Hamiltonian of the WM is suddenly changed so that the frequency matches the individual frequencies in the cold bath, (iii) the WM is brought into contact with the cold bath (the blue harmonic chain), with the same pattern of interaction as in step (i), and (iv) the Hamiltonian of the WM is suddenly changed back to its original value.

expression. Indeed, since the baths remain intact during the adiabat, the work is given by the energy change of the WM only. For example, in the adiabat after an interaction with the hot bath, the energy of the WM is decreased by $W_{h\to c} = (\omega_h - \omega_c)\,\mathrm{Tr}\,\boldsymbol{\sigma}_m$.

Lastly, an important feature of removing the standard idealizations in thermodynamics is that the engine cycles are not cyclic in the standard thermodynamic sense. Indeed, since the baths are finite and the interaction with WM perturbs them nonnegligibly, at the end of each cycle the state of the WM (and, of course, the state of the baths) will be different from that at the beginning. Nevertheless, we identify a period of "perfect" cycles when the deviation from cyclicity is small. We describe this period below.

3.4.1 Cycle performance

The cycle is set to begin with the interaction with the hot bath, so the starting Hamiltonian of the WM is $H_h = \omega_h a_m^\dagger a_m$, while its state is thermal at temperature T_c with respect to $\omega_c a_m^\dagger a_m$.

Due to the finite size of the baths and the strong perturbations that the interactions with the WM causes in them, one would expect that the performance of the engine dropped over time. This intuition is confirmed in Fig. 3.5, where we plot the work output and efficiency of the engine as a function of the number of cycles of operation. In Fig. 3.5a, each bar represents the work output during

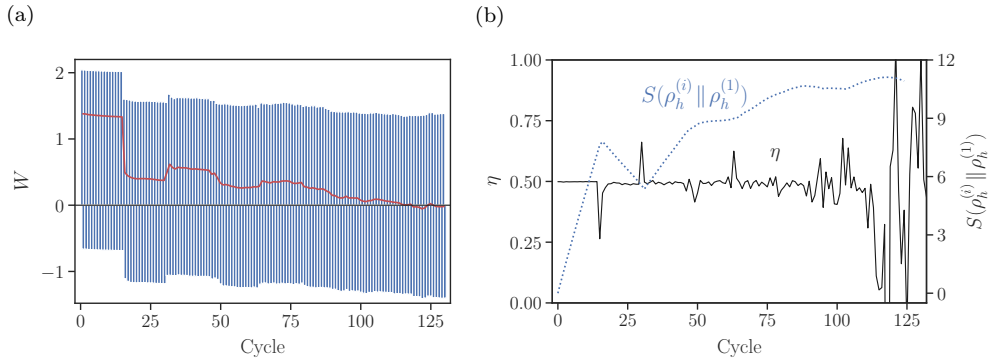

Figure 3.5: (a) Work output during the adiabats. One pair of positive-negative bars represents a full cycle. The red, solid line represents the total work extracted from each cycle (positive bar + negative bar + work during the isochores). The relevant parameters of the system are $N_h = N_c = 300$, $T_h = 4$, $T_c = 0.5$, $\omega_h = 2$, $\omega_c = 1$, $\alpha_h = \alpha_c = \gamma = 0.1$, $\tau = 100$, $\delta = 0.1\tau$. (b) (Solid black) Efficiency of the engine and (dotted blue) relative entropy of the hot bath with respect to its initial state for every cycle of operation. The divergences in the efficiency are caused by the extraction of no heat from the hot bath. Note how the efficiency and work output of the engine remaining virtually constant during the "perfect" cycles is contrasted by the steady increase of the distance, measured by the relative entropy, of the bath's state from its initial value.

an adiabat, and the red line represents the total work output in each cycle (the sum of the works in the adiabats plus the sum of the works in the isochores), as described above. The heat Q is defined as the energy that the hot bath loses per cycle. We define the energy of the hot bath with respect to the Hamiltonian in Eq. (3.19), and, for the n^{th} cycle, the heat is given by

$$Q = -\Delta E_h = \mathrm{Tr}\left\{\boldsymbol{F}_{\text{bath}}[\boldsymbol{\sigma}_h(2n\tau) - \boldsymbol{\sigma}_h((2n+1)\tau)]\right\}, \tag{3.30}$$

where $\boldsymbol{\sigma}_h(t)$ is the covariance matrix describing the state of the hot bath as a function of time. The efficiency of the engine is defined as usual: $\eta = W/Q$.

In Fig. 3.5, two very different regimes can be identified in the engine's performance. First, the work output and absorbed heat are approximately constant, decreasing very slowly, for the first 15 complete cycles. We call these "perfect" cycles.

During this period of "perfect" operation, the perturbations created by the WM-bath interaction propagate through the baths, outwards from the interaction site, in the same way as they would do were the baths infinite. Nevertheless, there is a slight degradation of the engine's performance during these cycles, caused by remnants left behind by the outward-propagating perturbations created by the WM-bath interaction. As the cycles proceed, these small

deviations from the interaction site's equilibrium state accumulate, causing a gradual decrease in work and heat.

One may be tempted to attribute the cause of this degradation to the strong coupling. However, the difference between the work outputs between two consecutive "perfect" cycles does not vanish when the coupling is taken to zero, and thus the perfect-regime degradation can not be a strong-coupling effect. Rather, it strongly depends on the ramp-up time and can be decreased noticeably by increasing δ/τ. However, going to high values of δ/τ prevents the WM from thermalizing with the baths, thereby impairing the functioning of the engine. This means that the degradation, which is a finite-time effect, is inherent to the model. An explicit computation of the correction to the optimal figures of merit due to this degradation can be found in Appendix B.2. It must be noted that, while the dependence of single-cycle characteristics on the ramp-up time is in line with the general intuition that noncommutative, time-dependent interactions generate excitations that cause thermodynamic friction (see, e.g., [RK06, PAA$^+$14]), the important fact of the cycle-to-cycle accumulation of the imperfections caused by finite switching time observed in our model is a separate phenomenon.

The number of "perfect" cycles, N_p, increases asymptotically linearly with $N \equiv N_c = N_h$, the number of nodes in the baths, as is to be expected given the constant, finite speed of propagation of the perturbations in the bath[3]. However, when the interaction time τ is close to τ_{th}, N_p does not depend on α for small α. Indeed, although the thermalization time increases with decreasing α [see Eq. (3.27)] and this requires longer interaction times τ with the baths, the propagation of perturbations within the baths also slows down, and the two effects almost exactly compensate each other. Along with the fact that the degradation is slow, the linear dependence of N_p on N makes the perfect regime relevant for practical engines with large baths.

Differences from the perfect-cycle behavior begin to appear only when the perturbations return to the region of the bath that directly interacts with the WM. This is the point at which the finite-size effects take relevance, and it is marked by the drastic, discontinuous drop in the work output in Fig. 3.5. The performance of the engine becomes unreliable due to large variations that heat and work undergo in both magnitude and sign. The above discontinuous behavior of the engine's figures of merit contrasts with the conventional gradual degradation of the performance of an engine operating between finite reservoirs (see, e.g., [IO14]). The contrast is further sharpened by the observation that,

[3] Recall that, for nearest-neighbor Hamiltonians such as that in Eq. (3.19), the Lieb-Robinson bounds set limitations on the propagation of perturbations [NRSS09, NS10]

as is also the case in the said conventional picture, the baths diverge from their initial states in a gradual, continuous manner. This is illustrated in Fig. 3.5b, where the distance, as measured by relative entropy [NC10], of the state of the hot bath at the beginning of the i^{th} cycle, $\rho_h^{(i)}$, to its initial state, $\rho_h^{(1)}$, is plotted as a function of i (see Appendix B.3 for the definition of relative entropy and its expression in the current scenario). In Fig. 3.5b it can be seen how, during the "perfect" cycles, the characteristics of the engine stay almost constant despite the fact that the bath's state changes at a steady rate.

An important implication of Fig. 3.5 is that, during the perfect cycles, the efficiency of the engine η is very close to $\eta_O = 1 - \omega_c/\omega_h$. The latter is the theoretical maximum for an oscillator running an idealized Otto cycle between two infinite thermal baths to which it is coupled weakly enough for the standard Markovian open quantum system techniques [BP02] to be applicable [RK06, QLSN07, KR17]. Such idealized engines are known to obey the so-called power-efficiency trade-off, which states that the power output of the engine has to approach to zero whenever the efficiency comes close to the reversible maximum (see, e.g., [SST16, KR17, AHMG13]). Our model respects the power-efficiency trade-off for the Otto cycle in the following manner: for $\alpha \ll 1$, the efficiency approaches η_O from below as

$$\eta_O - \eta \propto \alpha^2, \tag{3.31}$$

while for the work output of a perfect cycle we have $W = W_{\alpha=0} - \mathcal{O}(\alpha^2)$. We refer the reader to Appendix B.2 for a more detailed discussion on these quantities. Taking into account Eq. (3.27), this leads us to

$$P \propto \alpha. \tag{3.32}$$

Here P is the power output of the engine: $P = W/\tau_{\text{cycle}}$, where $\tau_{\text{cycle}} = 2\tau$ is the duration of the cycle. We note that, although the setting of our problem is different from that in Ref. [PLWR$^+$18], the scalings in Eqs. (3.31) and (3.32) agree with (and saturate) the optimal scalings derived there.

Additionally, one can also consider coupling the WM to more than one site. For doing so, we have studied the performance of the machine when the WM couples to evenly spaced sites in the ring. It turns out that adding more interacting sites reduces the amount of perfect cycles, which matches the intuitive picture described earlier. Indeed, the reduced distance between the interaction sites leads to a shorter time needed for the perturbations created to reach the nearest site of interaction. Interestingly, the work output of a single perfect cycle is insensitive to the cardinality of the set $\mathcal{I}$, as long as the perturbations generated in one interacting site do not have time to arrive to any other. But,

as expected, the total work output of the engine over several cycles does get reduced by increasing the number of interaction points.

On the other hand, we have observed that the more sites the WM interacts with, the smaller the time necessary for it to thermalize is. This leads to an increased power output for the initial perfect cycles, albeit at the cost of decreasing the number of such cycles.

3.4.2 Propagation of correlations

As noted before, the formalism presented in Section 3.2 allows for an easy way of identifying whether two systems are correlated. In this section we use this property to study how correlations distribute and evolve along the baths and the WM. We consider this as one of the (probably many) paths to obtain a better understanding of the phenomenology presented above.

In Fig. 3.6 one can observe the strength of the correlations between the WM and each of the oscillators in each bath, and how these correlations evolve in time for five consecutive cycles. Throughout this section, each bath is composed of $N = 30$ oscillators, with all other parameters being the same as in Section 3.4.1. The vertical lines denote the instants of time at which the WM stops interacting with one bath and, after the corresponding adiabat, begins interacting with the other.

A feature one immediately observes is the explicit propagation of the perturbations in the form of localized wavepackets at finite speed, in full agreement with the Lieb-Robinson bound [NRSS09, NS10]. While propagating (although it is hard to see in Fig. 3.6), these wavepackets leave residual perturbations behind. The latter are small and, during the first three cycles of operation ($t \in [0, 600]$), the WM appears to interact with almost unperturbed baths. These are the "perfect" cycles described in Section 3.4.1. At time $t = 600$ the perturbations that were generated during the first three cycles manage to intercept the WM as it is *currently interacting* with the bath, leading to the sudden drop of the work output that has been discussed in Section 3.4.1.

It is also noticeable that the propagating correlations quickly fade. This is unsurprising, and carries the same explanation as that given for Fig. 3.2: the transformation of correlations between WM and the nodes in the bath into intra-bath correlations. The computations performed show that, to a surprisingly good approximation, during an interaction the WM becomes correlated with just a single nonlocal mode in the bath—the mode that propagates outwards—as can be appreciated in Fig. 3.6. However, both the WM and this propagating mode are interacting with the rest of the bath as well, and thus this correlation is quickly lost and distributed among bath modes. This also explains the different

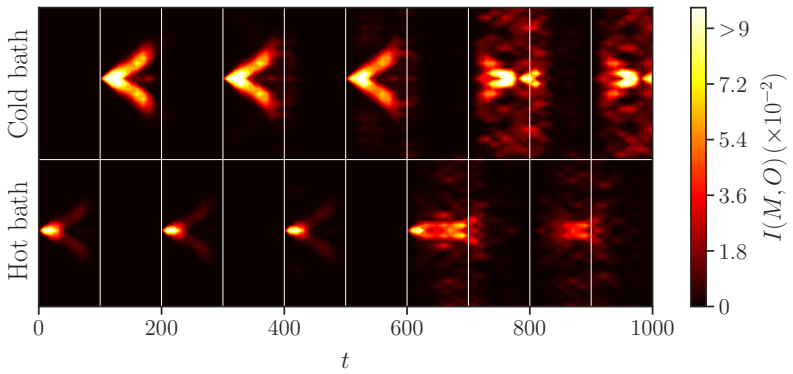

Figure 3.6: Mutual information between the WM and each oscillator in each bath during the five initial cycles of operation. The horizontal line separates the two baths, and the vertical lines separate the interactions of the machine with each of the baths. The baths have $N = 30$ oscillators each, and are initially uncorrelated and at temperatures $T_c = 0.5$ and $T_h = 4$. The rest of relevant parameters are as those used in Fig. 3.4. Note how during the first three cycles the machine observes no differences from the interaction with infinite baths, and after this point the perturbations in the chains arrive back to the interacting oscillator, modifying its local state.

decay rates in the the hot bath (much faster decay, due to the stronger mode-mode interaction) and in the cold bath. Indeed, the hotter the bath, the larger the thermal noise that will break the correlations.

It is also instructive to observe how the correlations are built up and distributed along the baths. This is illustrated in Fig. 3.7, in which the mutual information between each pair of oscillators in each bath is shown at various times. The outward-propagating nature of the correlations is immediately noticeable in this figure.

An important insight into the process of the degradation of the engine is gained by looking at the bath-bath correlations instead. Indeed, given that the WM acts as a carrier of both energy and correlations between the baths, and that the baths gradually evolve away from their initial states, one would expect that, over time, the baths get more and more correlated and end up reaching a global passive state (see [BFH16] for the characterization of passivity within GQM). This intuition is explored in Fig. 3.8, in which it is shown that, surprisingly, the mutual information between the two baths remains close to zero during the ideal cycles, and starts abruptly increasing after the last ideal cycle is complete. This can be explained by noticing in Fig. 3.6 that, during the perfect cycles, the WM is virtually uncorrelated with the baths both at the beginning and at the end (but not in the middle) of each interaction session, which means

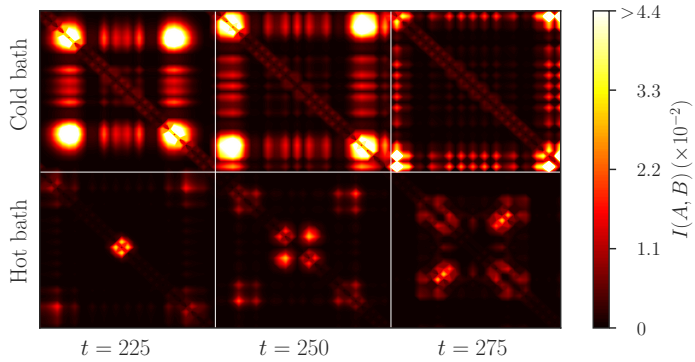

Figure 3.7: Intra-bath correlations at different moments of time, during the second cycle of operation of the engine. The parameters chosen are the same as those in Fig. 3.6. Note that in the cold bath the correlations propagate outwards, away from the interaction point (in the center of the images), while in the hot bath we observe two waves: one propagating outwards—generated by the latest interaction with the WM—and one propagating inwards, generated in the previous interaction of the bath and the WM. The latter wave, due to the boundary conditions of the system, returns to the interaction point. The scripts available in the computational appendix [PKHB17] generate a full animation of this phenomenon.

that the WM does not transmit correlations during these cycles. This picture breaks down once the perturbations reach the interaction site. It is at this point when the WM remains significantly correlated to the baths after each interaction, transmitting those correlations from one bath to the other. This thereby constitutes a clear quantitative link between the correlations among the baths and the optimal performance of the engine.

It is worth noticing that, despite the fact that the mutual information between different elements of the system can be substantially large, for the choice of parameters in the studies performed, none of the correlations built involve entanglement. Indeed, it is well known that entanglement in quantum fields decays very rapidly with temperature, reaching zero at a finite value [Bro13]. However, for a sufficiently cold bath, one could still expect some entanglement to be present, although it is not clear whether it will play a significant role in the engine's performance.

3.5 Discussion

Using the formalism of Gaussian quantum mechanics, in this chapter we have developed a model that is able to circumvent the standard assumptions of weak coupling, slow driving, and infinite size of the baths, usually employed when

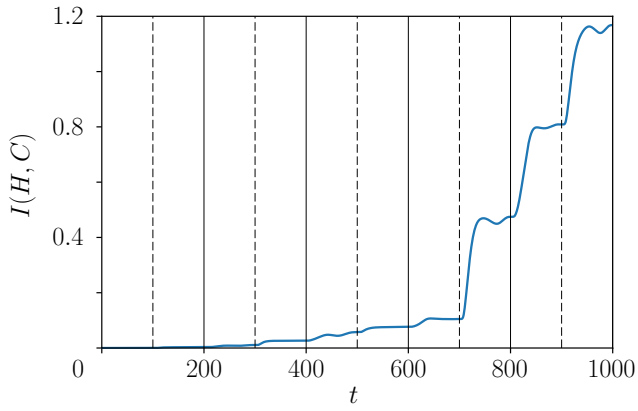

Figure 3.8: Mutual information between the hot and cold baths during the five initial cycles of interaction. The parameters used are the same as for Figs. 3.6 and 3.7. The solid lines denote the end of each cycle, while the dashed lines denote the end of the interaction of the WM with the hot bath and the beginning of its interaction with the cold bath. Note the abrupt increase of the mutual information during the fourth cycle (starting at $t = 600$), which is the first cycle outside the regime of "perfect" operation.

studying thermodynamic phenomena. The focus was put on a single, driven harmonic oscillator interacting with finite harmonic thermal reservoirs, in what constitutes the first model where none of the assumptions is made. With it, we have studied in detail the interaction of a machine with a single bath, and then repeated alternating interactions with two baths in an Otto cycle.

When interacting with a bath, we observe not only how the machine thermalizes, but also how the interaction creates correlations between the WM and the region of the bath that directly interacts with it. These correlations propagate across the bath at a constant speed, but leaving residual perturbations behind that slightly affect the figures of merit when operating cyclically.

In the study of the quantum Otto cycle, we are able to conclude that the crucial element that determines the performance of the engine is the propagation of the perturbations created by the WM-bath interaction. There is always a number of "perfect" cycles in which the figures of merit are close to optimal, followed by a breakdown after the perturbations have travelled around the baths. We expect this picture to also hold beyond the Gaussian regime, provided the speed of sound within the baths is finite [NRSS09, NS10].

Our results constitute an important contribution for assessing finiteness effects in a field that has historically relied on infinite (time, size, and subtlety of the interactions) idealizations. In so doing, it opens new questions of interest both

at the fundamental and the experiment-oriented level.

For instance, we have observed that, remarkably, the bath-bath correlations remain very close to zero during the "perfect" cycles of operation, and start increasing abruptly right after. The reason for this abrupt behavior, and more generally, how the overall system evolves to a more and more passive state, are yet to be understood. Another important question that arises when using the GQM formalism for thermodynamics is that of the characterization of adiabatic processes. While in this chapter we have used the simplest adiabatic transformation, it will be important to know whether there exist Hamiltonian evolutions that lead to adiabatic transformations, and devise means of quantifying deviations from adiabaticity.

Also, it will be important to see how constraints required by experimental realizations can be accommodated in the model. While the finite time in which adiabatic evolutions are implemented in practice has been briefly discussed, it is interesting to understand how the model, and in particular the efficiency in the simulations, is affected by other features such as long-range couplings, dynamic coupling strengths, or having a many-body WM, that may appear when analyzing experimental realizations of quantum thermal engines.

4 Machine learning with quantum information

Parallel to the development of quantum technologies, the last decade has also experienced the rise of big data and the emergence of applications of machine learning in real-world problems. A cross-breeding of the two fields was inevitable, and nowadays there is a big research effort to discern whether processing information encoded in quantum systems provides practical advantages in machine learning, either in terms of computation time, storage space, or by enabling new features. In this chapter we propose a quantum algorithm that performs training of deep neural networks, requiring exponentially less time to run than its classical analogue if an efficient state preparation scheme is available. The algorithm, based on Bayesian techniques, enables an easier estimation of the confidence of the predictions made, which is very hard to achieve when training with standard methods based in gradient descent. We begin by reviewing the classical theory that allows for Bayesian training of deep neural networks. After that, we describe the algorithm proposed and analyze its theoretical runtime. Finally, we test the core component of the algorithm on both classical simulators and real quantum computers, analyzing the impact of noise in the executions.

4.1 Introduction

In recent years the world has witnessed a revolution in machine learning. A revolution that was enabled by two critical factors: the development of the deep learning framework [Hin07], and the realization that Graphical Processing Units could be employed to perform relevant calculations much more efficiently than conventional CPUs. These allowed for algorithms to reach, and even surpass, human-level performance in the automation of tasks such as image recognition [TL19], game playing [SSS+17, JCD+19], or natural language processing [vdODZ+16, ASWdF16]. Nowadays, deep learning algorithms underlie recommendation systems like those of Amazon and Netflix, the control systems of autonomous vehicles, or the face recognition devices in airports. Deep learning has also been applied as a tool in various other areas of research, with notable results such as early diagnosis of medical conditions [GPC+16] or the

discovery of exoplanets [SV18].

One of the main characteristics of deep learning algorithms is the large amount of parameters used to model the data under scrutiny. This, combined with the training techniques employed—which rely on the theory of functional optimization—makes the task of identifying the specific characteristic that triggers a prediction very complicated in practice. In fact, deep learning algorithms are often regarded as *black boxes*, that only provide results, but not explanations of these. Furthermore, there exist training techniques that lack a solid foundation, with confirmations being done empirically rather than theoretically [Tie08]. In this context, questions such as "what is the confidence that the prediction given is actually correct?" or "what is the difference in confidence between the first and second choices for the prediction?" are very difficult to answer. Given that it is expected that in the coming years machine learning algorithms will have more and more relevance in one's life, it is important to develop tools that allow one to "open the box" and obtain information about the process of formulating a prediction [XAI].

In this sense, the Bayesian approach to machine learning provides a clear advantage over traditional techniques, as it provides information about the uncertainty in their predictions. It does so by transitioning from the theory of functional optimization to that of probability, where the quantity of interest is the probability distribution of possible predictions conditioned on the previous observation of known data. But not only that, they have further advantages, including automated ways of learning structure and avoiding overfitting, and robustness to adversarial attacks [BMG17, GPSB17].

In fact, the Bayesian framework has already been making advances in various deep architectures [BCKW15, GG16]. Nevertheless, the methods for Bayesian training of machine learning algorithms have a high computational cost that prevent their application to real-world problems which require the processing of thousands, or even millions, of datapoints. In the quest for faster and more efficient ways to process the information, it was suggested the idea that one could use properties offered by quantum mechanics, such as superposition, entanglement and interference, to perform calculations of prohibitive cost in classical computers.

In fact, quantum technologies have been already making advances in machine learning. A new breed of quantum neural networks is aimed at current and near-future quantum computers [VBB17, TGR19, KVD$^+$18, FN18, VPB18], which is in stark contrast with attempts in the past [SSP14]. Some constraints must be observed that, however, are unusual in classical machine learning algorithms. Importantly, quantum protocols must be coherent, that is, we require from a quantum machine learning algorithm that it is described by a unitary transfor-

mation that maps input nodes to output nodes. While the common wisdom is that a nonlinear activation is a necessary component in neural networks, a linear, unitary mapping between the inputs and outputs actually reduces the vanishing gradient problem [ASB16, HR17]. Training a hierarchical representation in a unitary fashion is also possible on classical computers [LRW+19, Sto18]. So while this constraint is unusual, it is not entirely unheard of in classical machine learning, and it is the most common setting in quantum-enhanced machine learning [BWP+17].

With the aim of obtaining the best of both worlds, namely the classical knowledge of optimization algorithms and the quantum advantage provided by superposition, recently various hybrid variational methods [PMS+14, FGG14, MRBAG16, KVD+18] have been developed, which perform a classical learning loop of the parameters of a quantum circuit.

4.2 Preliminaries

The algorithm that we present combines, on one hand, the theory of Gaussian processes for Bayesian inference and its connections to training of deep neural networks, and on the other hand, an existing protocol for quantum-assisted Gaussian process regression, based on a quantum algorithm for solving linear systems of equations. In this section we review these topics, focusing on the aspects that are relevant for the chapter.

4.2.1 Gaussian processes for machine learning

Before the development of deep learning, the most popular algorithms in machine learning belonged to a category denominated "kernel machines". Among these, the two most studied were the support-vector machine [CV95] and the Gaussian process [RW06]. While both can be used for addressing the same task of classification via supervised learning, they do so in different ways. Support-vector machines approach classification as an optimization problem, where the task is to define a boundary that separates the datapoints with different labels, and has the largest distance to them. Then, the label assigned to a new datapoint corresponds to the side of the boundary it falls in. On the other hand, Gaussian processes take a probabilistic approach: having observed the training instances with their respective labels, and given a new datapoint, the task is to know which is the probability distribution over labels conditioned on all previous information and the new observation. The label assigned in this case corresponds to that with highest probability.

It is immediate from the probabilistic approach that it allows for predictive uncertainty quantification, this is, to know not only which is the label assigned to a datapoint, but also measures of confidence on the prediction and its distance to the alternatives. But not only that, they have further advantages. In addition to its Bayesian foundation, these include the avoidance of overfitting, and robustness to adversarial attacks [BMG17, GPSB17], which represent important problems in backpropagation-based machine learning.

Mathematical description

The concept of a *process* can be seen as the generalization to functions of probability distributions over random variables. This is, a process defines a probability distribution over functions, understood as infinite-dimensional vectors specifying the value of the function for every input. Then, a Gaussian process is defined as

Definition 1. A collection of random variables is a Gaussian process if every finite collection of them has a multivariate normal distribution.

More formally, a (possibly uncountably-infinite) set of random variables $\{f_i \triangleq f(\boldsymbol{x}^{[i]})\}_i$ is a Gaussian process if for any countable subset, $\{f_1, \ldots, f_n\}$,

$$(f_1, \ldots, f_n) \sim \mathcal{N}(\boldsymbol{\mu}, K) \propto \exp\left[-\frac{1}{2}(\boldsymbol{f}-\boldsymbol{\mu})^{\mathrm{T}} K^{-1}(\boldsymbol{f}-\boldsymbol{\mu})\right], \tag{4.1}$$

where the parameters of the process, $\boldsymbol{\mu}$ and K, are known as the vector of means and the covariance matrix. The former is constructed from a mean function, $\boldsymbol{\mu} = (\mu(\boldsymbol{x}^{[1]}), \ldots, \mu(\boldsymbol{x}^{[n]}))$, that can be assumed to be $\mathbf{0}$ without loss of generality, while the latter is built from a kernel function $k(\boldsymbol{x}, \boldsymbol{y})$ evaluated on every pair of points, $K_{ij} = k(\boldsymbol{x}^{[i]}, \boldsymbol{x}^{[j]})$. It is standard to denote the fact that $\{f_i\}_i$ is a Gaussian process by $f \sim \mathcal{GP}(\mu, k)$

The distribution of a Gaussian process is the joint distribution of all the (infinitely many) random variables in $\{f_i\}_i$ and, as such, it is a distribution over functions with continuous domain.

A Gaussian process defines a prior over functions, which can be converted to a posterior after some data has been observed. In the context of supervised learning, given a dataset $\mathcal{D} = (X, \boldsymbol{y})$ with $X = \{\boldsymbol{x}^{[i]}\}_{i=1}^{M}$ being the training instances and $\boldsymbol{y} = \{y_i\}_{i=1}^{M}$ the corresponding labels, and a new datapoint $\boldsymbol{x}_*$, the quantity of interest is the distribution of possible labels y_* conditioned on the observed data, $p(y_* | \boldsymbol{x}_*, X, \boldsymbol{y})$. For Gaussian processes this distribution is again

Gaussian, with mean and variance estimates given, respectively, by

$$\bar{y}_* = \boldsymbol{k}_*^{\mathrm{T}}(K + \sigma_n^2 \mathbb{1})^{-1}\boldsymbol{y}, \tag{4.2}$$

$$\mathbb{V}[y_*] = k(\boldsymbol{x}_*, \boldsymbol{x}_*) - \boldsymbol{k}_*^{\mathrm{T}}(K + \sigma_n^2 \mathbb{1})^{-1}\boldsymbol{k}_*. \tag{4.3}$$

Here $K + \sigma_n^2\mathbb{1}$ denotes the covariance matrix of the model when one inserts the extra assumption that the values of the observations may have some Gaussian noise (so $y_i = f(\boldsymbol{x}_i) + \epsilon_i$) with variance σ_n^2, and $\boldsymbol{k}_*$ denotes the vector created by evaluating the kernel function in the target point and all the training points, $\boldsymbol{k}_* = (k(\boldsymbol{x}_*, \boldsymbol{x}^{[1]}), \ldots, k(\boldsymbol{x}_*, \boldsymbol{x}^{[M]}))$.

Connections to deep learning

The correspondence between Gaussian processes and a neural network with a single hidden layer is well-known [Nea94] and, in fact, it constitutes one of the most central results in the theoretical foundations of deep learning[1]. Indeed, let a single-hidden-layer neural network with input $\boldsymbol{x} \in \mathbb{R}^{d_{in}}$ and output $\boldsymbol{z}(\boldsymbol{x}) \in \mathbb{R}^{d_{out}}$ be defined by

$$z_i(x) = b_i^1 + \sum_{j=1}^{d_1} W_{ij}^1 x_j^1(\boldsymbol{x}), \quad x_j^1(\boldsymbol{x}) = \phi\left(b_j^0 + \sum_{k=1}^{d_{in}} W_{jk}^0 x_k\right), \tag{4.4}$$

where $\boldsymbol{x}^1(\boldsymbol{x}) \in \mathbb{R}^{d_1}$ is the state of the neurons in the hidden layer, $\phi(x)$ is an activation function, and the weights and biases of the hidden and output layers are denoted as W and b, respectively.

As an initial assumption, take that all weight and bias parameters are drawn from independent and identically distributed (i.i.d.) random variables. If such is the case, for a fixed input $\boldsymbol{x}$, the value of each output neuron $z_i(\boldsymbol{x})$ will be a sum of i.i.d. terms. Therefore, the Central Limit Theorem implies that, when $d_1 \to \infty$, $z_i(\boldsymbol{x})$ is a random variable that follows a Gaussian distribution. An analogous reasoning shows that for any set of k input data points, the corresponding outputs $\{z_i(\boldsymbol{x}^{[1]}), z_i(\boldsymbol{x}^{[2]}), \ldots, z_i(\boldsymbol{x}^{[k]})\}$ will follow a joint multivariate Gaussian distribution, with a covariance matrix given by $K(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[x_i^1(\boldsymbol{x}) x_i^1(\boldsymbol{x}')]$. This covariance matrix is, in fact, the same for all i due to the i.i.d. assumption on the priors. Additionally, the mean of the distribution can be set to zero if the prior is assumed to have zero mean. Therefore, each neuron in the output layer, z_i, is a Gaussian process, $z_i \sim \mathcal{GP}(\mu, K)$.

[1] *The connection between deep learning and Gaussian processes legitimates deep learning as an appropriate learning framework, since it can be understood as an instance of approximate Bayesian inference.* Seth Lloyd, Benasque Quantum Information, 2019.

The Bayesian training of the neural network corresponds to computing the posterior distribution of the given GP model, that is, calculating the mean and variance of the predictive distribution from inverting the covariance matrix as per Eqs. (4.2) and (4.3).

This argument can generalized to deep neural network architectures in a recursive manner [LBN+18, MHR+18]. Letting z_i^l denote the i^{th} component of the output of the l^{th} layer, by induction it follows that $z_i^l \sim \mathcal{GP}(0, K^l)$. K^l, the covariance matrix on the l^{th} layer, is given by

$$K^l(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[z_i^l(\boldsymbol{x}) z_i^l(\boldsymbol{x}')] = \sigma_b^2 + \sigma_w^2 \mathbb{E}[\phi(z_i^{l-1}(\boldsymbol{x}))\phi(z_i^{l-1}(\boldsymbol{x}'))], \tag{4.5}$$

where the variance on the weight and bias parameters, σ_w^2 and σ_b^2, as well as the nonlinearity ϕ, are typically known or adjusted to minimize the logarithmic marginal likelihood [ZFO+19], and $z_i^{l-1} \sim \mathcal{GP}(0, K^{l-1})$.

Remarkably, numerical experiments suggest that the infinite-width neural network trained with Gaussian priors outperforms finite deep neural networks trained with stochastic gradient descent in many cases [LBN+18, MHR+18].

4.2.2 Quantum-assisted Gaussian process regression

The computation of the parameters of the posterior, given by Eqs. (4.2) and (4.3), requires the inversion of the covariance matrix $K + \sigma_n^2 \mathbb{1}$. This operation has a time complexity of $\mathcal{O}(M^{2.373})$ with the best current algorithms [LG14], where M is the size of the matrix to invert. In machine learning, the number of datapoints, and therefore the size of the covariance matrix, can be of the order of millions, making the application of Gaussian processes to real problems impossible in practice.

In the quest for faster solutions, a quantum algorithm for Gaussian process regression was introduced in [ZFF19]. The main idea of the method is to encode the vectors $\boldsymbol{y}$ and $\boldsymbol{k}_*$ in states of quantum systems, $|\boldsymbol{y}\rangle$ and $|\boldsymbol{k}_*\rangle$, and use the covariance matrix as a Hamiltonian on such systems. With these, the state of the system can be evolved to $(K + \sigma_n^2 \mathbb{1})^{-1} |\boldsymbol{b}\rangle$, where $|\boldsymbol{b}\rangle = |\boldsymbol{y}\rangle$ or $|\boldsymbol{k}_*\rangle$ depending on whether one wants to compute the mean or the variance predictor, with the quantum linear systems algorithm developed in [HHL09]. Finally, the inner product with $\langle \boldsymbol{k}_*|$ can also be performed in an efficient way using known methods [TMGB19, SK18, ZFF19].

The quantum linear systems algorithm, also known as HHL algorithm, has a runtime of $\mathcal{O}(\log(M))$ when K is sparse—it has at most $s \ll M$ nonzero entries

per row—and well-conditioned[2,3]. As such, it is a very important primitive in quantum computing, being at the core of many quantum algorithms [WBL12, Ber14, RML14]. Due to its central role, both in the quantum-assisted GP algorithm and in the one we propose in Section 4.3, let us review it briefly. Assume a Hamiltonian operator A and a quantum state $|\boldsymbol{b}\rangle = \sum_{j=1}^{N} \beta_j |u_j\rangle$, where $|\boldsymbol{b}\rangle$ is written in the basis that diagonalizes A (this is done for simplicity in the notation, without losing any generality). The HHL algorithm consists of the following steps:

- Perform quantum phase estimation [Kit95] to extract estimates of the eigenvalues of A, and store them to precision $1/T$ in a quantum register as a weighted superposition. This is achieved by appending an ancillary register of size $\log(T)$ initialized in a—not necessarily uniform—superposition of all its possible states, applying the conditional Hamiltonian evolution $\sum_{\tau=0}^{T-1} e^{\mathrm{i}A\tau t_0/T} \otimes |\tau\rangle\langle\tau|$ to the combined system of initial register and ancilla, and then performing an inverse quantum Fourier transform in the ancillary register. The state of the full system after this step is

$$\sum_{j=1}^{N} \sum_{k=0}^{T-1} \beta_j c_{k|j} |u_j\rangle |k\rangle ,$$

  where $|c_{k|j}|$ is large only if $\lambda_j \approx \frac{2\pi k}{t_0}$, $\{\lambda_j\}_{j=1}^{N}$ being the eigenvalues of A.

- Append an additional qubit as a second ancillary register. Initialized in the state $|0\rangle$, rotate this qubit, via a controlled rotation conditional on the value stored in the eigenvalue register, an angle proportional to the inverse of such eigenvalue. This leads to an entangled state between the initial and both ancillary registers, namely

$$\sum_{j=1}^{N} \beta_j |u_j\rangle |\lambda_j\rangle \left(\sqrt{1 - \frac{C^2}{\lambda_j^2}} |0\rangle + \frac{C}{\lambda_j} |1\rangle \right) ,$$

  where C is a problem-dependent constant to be set and we have assumed[4] that $c_{k|j} = \delta_{jk}$. Due to this and for simplicity, we have also changed notation from $|k\rangle$ to $|\lambda_j\rangle$[5].

[2] New proposals [Amb10, CKS17, WZP18] have the same runtime in size, but improved bounds in sparsity and condition number.

[3] A thorough analysis on the runtime of the HHL algorithm can be found in [Aar15].

[4] This is the case when $T \to \infty$ and the qubits in the register are prepared in a uniform superposition. In realistic implementations, these coefficients must be taken into account when assessing the performance of the protocol [HHL09].

[5] Loosely speaking, if $c_{k|j} \approx \delta_{jk}$, then the labels k and λ_j can be used interchangeably to denote a same quantum state, since $\lambda_j \approx \frac{2\pi k}{t_0}$, with t_0 being a constant.

- Uncompute all operations performed in the initial ancillary register, and measure the last register, and discard it. Conditioned on the outcome 1, the final state after the algorithm is

$$\sum_{j=1}^{N} \frac{\beta_j}{\lambda_j} |u_j\rangle |1\rangle = \left(A^{-1} |\boldsymbol{b}\rangle\right) \otimes |1\rangle .$$

Due to the use of HHL, the quantum GP algorithm also runs in $\mathcal{O}(\log(M))$ time when K is sparse and well-conditioned. An important step is the simulation of the Hamiltonian evolution in the phase estimation subroutine. Efficient quantum algorithms exist already for performing this operation [Llo96, Chi10, BC12]. Furthermore, Hamiltonian simulation can be addressed by tapering the covariance function using a compactly supported function [FGN06]. A similar methodology is also known in kernel methods [WT11]. However, as noted by Aaronson in [Aar15], sparsity and well-conditioning may not be sufficient to ensure the exponential speedup of the HHL algorithm. If preparing the state $|\boldsymbol{b}\rangle$ takes M^c steps for some constant c, the exponential speedup is lost. The quantum Bayesian training algorithm described in the following section solves partially this issue by not requiring a dynamical data structure such as a quantum random access memory, but a single circuit that can be queried as many times as needed.

Subsequently to the quantum GP algorithm, a corresponding quantum method for enhancing the training and model selection of GPs was introduced in [ZFO+19].

4.3 Quantum Bayesian training of neural networks

Now, using both the connection between GPs and training of deep neural networks, and the quantum GP algorithm, we develop a way of conducting Bayesian training of deep neural networks using a Gaussian prior.

According to the connection described in Section 4.2.1, Bayesian training of a deep neural network of L layers requires sampling the values of the neurons in the final layer from the Gaussian process $\mathcal{GP}(0, K^L)$, where K^L can be computed in a recursive manner beginning from K^0, the covariance matrix of the training data, following Eq. (4.5). If we had classical access to the elements of the covariance matrix K^0, one possibility could be to classically compute K^L and then resort to the simulation of the Hamiltonian evolution generated by K^L to obtain the mean predictor $\bar{y}_*$ and variance predictor $\mathbb{V}[y_*]$ needed in the quantum Gaussian process algorithm of Section 4.2.2 [ZFF19]. This procedure

would require simulating the Hamiltonian evolution from a classical encoding of K^L, which may hinder the speedup expected from the algorithm in this case.

The algorithm we propose makes use of the following observation: the quantum GP algorithm does not require a complete knowledge of the covariance matrix. In reality, one just needs to know the time evolution operator under the covariance matrix encoded as a Hamiltonian. We propose a way of constructing such time evolution operator given access to a quantum encoding of the covariance matrix of the training dataset, K^0, given as a density matrix of a multi-qubit system. Once the time evolution operator is simulated, our algorithm, as the quantum Gaussian process algorithm, needs sampling from only one Gaussian process, that corresponding to the last layer in the network.

The condition that K^0 can be encoded in a density matrix requires of some properties to be satisfied: it must be a symmetric, positive semi-definite matrix, normalized by its trace in order to qualify as a quantum state [RML14]. The two first are satisfied by the definition of covariance matrix, and the last one can be achieved with an appropriate rescaling. This can be implemented by an adequate choice of the kernel function $k(\boldsymbol{x}, \boldsymbol{y})$.

A requirement of the algorithm is, as in the classical case, a functional expression of the covariance matrix in the last layer in terms of the base case K^0. For general nonlinear activation functions, this can only be done with numerical integration, which is impossible to implement coherently in contemporary quantum computers. A complete quantum protocol would require a large number of qubits and at least polynomial-size quantum circuits, which remains out of reach with current technology. However, different works have shown activation functions which yield kernels and recursion relations that can be analytically calculated or approximated [CS09, DFS16]. A particularly useful special case is that using the ReLU nonlinear activation on every layer. The ReLU activation function is $\phi(x) = \max(0, x)$, and has been crucial in addressing issues such as the vanishing gradient problem in deep learning [GBB11]. For this case, the covariance matrix of the l^{th} layer has an analytical formula [LBN+18]:

$$K^l(\boldsymbol{x}, \boldsymbol{x}') = \sigma_b^2 + \frac{\sigma_w^2}{2\pi}\sqrt{K^{l-1}(\boldsymbol{x}', \boldsymbol{x}')K^{l-1}(\boldsymbol{x}, \boldsymbol{x})} \times \left[\arcsin(\theta^{l-1}_{x,x'}) - (\pi - \theta^{l-1}_{x,x'})\arccos(\theta^{l-1}_{x,x'})\right], \tag{4.6}$$

where

$$\theta^l_{x,x'} = \arccos\left(\frac{K^l(\boldsymbol{x}, \boldsymbol{x}')}{\sqrt{K^l(\boldsymbol{x}, \boldsymbol{x})K^l(\boldsymbol{x}', \boldsymbol{x}')}}\right). \tag{4.7}$$

The nonlinear functions featured in Eq. (4.6) can be approximated by polynomial series with some convergence conditions.

Note that the factor $K^l(\boldsymbol{x},\boldsymbol{x})K^l(\boldsymbol{x}',\boldsymbol{x}')$ represents the outer product between two copies of the vector of diagonal entries in K^l. Therefore, the computation of Eq. (4.6) can be decomposed into such outer product operations combined with element-wise matrix (also known as Hadamard) multiplications. In the following we provide a construction for simulating the evolution under the Hamiltonians generated by these operations on the matrix elements of a quantum state.

Our approach is inspired by the quantum principal component analysis algorithm [RML14] where the density matrix ρ of a quantum state is treated as a Hamiltonian and used to construct the desired controlled unitary $e^{it\rho}$ acting on a target quantum state for a time period t. This is an unusual concept for classical machine learning and classical algorithms: a high-dimensional vector becomes an operator on itself to reveal its own eigenstructure. A throughout description of this density-matrix-based Hamiltonian simulation procedure is presented in [KLL+17]. We give below an overall description of the method, while the detailed analysis is presented in Appendix C.1.

In order to apply density matrix-based Hamiltonian simulation using the kernel of the l^{th} layer, we need to incorporate methods to compute certain element-wise matrix operations between two density matrices. It is convenient to define the following operators:

$$S_{\odot} = \sum_{j,k} |j\rangle\langle k| \otimes |j\rangle\langle k| \otimes |k\rangle\langle j| , \tag{4.8}$$

$$S_{\oslash} = \sum_{j,k} |j\rangle\langle j| \otimes |k\rangle\langle k| \otimes |k\rangle\langle j| . \tag{4.9}$$

With an augmented density matrix exponentiation scheme, $S_{\odot}$ computes the exponential of the Hadamard product of two density matrices, while $S_{\oslash}$ computes the exponential of the outer product between the diagonal entries of two density matrices. Specifically, we have that

$$\mathrm{Tr}_{1,2}\left[e^{-\mathrm{i}S_{\odot}\delta}(\rho_1\otimes\rho_2\otimes\sigma)e^{\mathrm{i}S_{\odot}\delta}\right] = e^{-\mathrm{i}(\rho_1\odot\rho_2)\delta}\sigma e^{\mathrm{i}(\rho_1\odot\rho_2)\delta} + \mathcal{O}(\delta^2), \tag{4.10}$$

where $\rho_1 \odot \rho_2$ denotes the Hadamard product between ρ_1 and ρ_2, and $\mathrm{Tr}_{1,2}$ denotes tracing out the first and second subsystems and thus only working with the register that encodes σ. The factor δ represents a small evolution time with the operator in the exponents. We also have that

$$\mathrm{Tr}_{1,2}\left[e^{-\mathrm{i}S_{\oslash}\delta}(\rho_1\otimes\rho_2\otimes\sigma)e^{\mathrm{i}S_{\oslash}\delta}\right] = e^{-\mathrm{i}(\rho_1\oslash\rho_2)\delta}\sigma e^{\mathrm{i}(\rho_1\oslash\rho_2)\delta} + \mathcal{O}(\delta^2), \tag{4.11}$$

where $\rho_1 \oslash \rho_2$ denotes taking the outer product between the diagonal entries of ρ_1 and ρ_2. The derivation of Eqs. (4.10) and (4.11) are presented in Appendix C.1.

Both $S_\odot$ and $S_\oslash$ are sparse and thus efficiently simulable as a Hamiltonian with methods based on quantum walks [BC12, BCK15]. A similar method of using a modified version of the SWAP operator combined with density matrix exponentiation was used in [RSML18] for a quantum singular value decomposition algorithm.

In order to approximately compute the nonlinear function of Eq. (4.6) we make use of a polynomial series in $K^0(x, x')$. Note that due to the structure of Eq. (4.6) the products involved in this polynomial series are just the Hadamard and diagonal outer products. We will denote the polynomial in K^0 to the order $n(l)$ which approximates the covariance matrix of the l^{th} layer as $P^n_{(\odot,\oslash)}(K^0)$.

We note that by using a generalized $\tilde{S}$ operator which combines the components in $S_\odot$ and $S_\oslash$ one can implement n operations $\odot$ and $\oslash$ in arbitrary orders. In Appendix C.1 we show that this simply amounts to summing over the tensor product of the projectors $|j\rangle\langle j|$, $|j\rangle\langle k|$, and $|k\rangle\langle k|$. Similar problems of simulation of polynomial series were addressed in [KLL$^+$17] and [RSW$^+$19], but the type of product considered was standard matrix multiplication instead of element-wise operations.

The quantum technique described above, combined with using the series expansions of the nonlinear functions in Eq. (4.6), gives us a way to approximate $e^{\mathrm{i}tK^l}\sigma e^{-\mathrm{i}tK^l}$, where σ is an arbitrary input state. Hence, given multiple copies of a density matrix which encodes the initial layer covariance matrix, K^0, the unitary operator, $\exp[-\mathrm{i}tP^n_{(\odot,\oslash)}(K^0)]$ can be constructed to act on an arbitrary input state, as required by applying the quantum GP algorithm described in Section 4.2.2. Note that, in contrast with previous approaches where one needed a quantum random access memory to perform oracle queries of the elements of K^0, the requirement of having K^0 encoded as a density matrix is much more feasible given current technology since the desired state preparation can be encoded in a quantum circuit and run as many times as needed without requiring any dynamical structure. Nevertheless, as described in [Aar15], it may still be the case that the circuit for preparing the density matrix takes a time polynomial in the size of K^0, potentially hindering the exponential speedup of the rest of the algorithm.

4.4 Experiments

The central part of the algorithm described in Section 4.3 is the intricate quantum protocol of linear systems of equations for computing the predictors in Eqs. (4.2) and (4.3). As described in Section 4.2.2, this algorithm is probabilistic, meaning that it only succeeds conditioned on obtaining specific results

after measuring specific qubits. Therefore, it is not assured that the protocol will succeed in a particular run, and it has to be repeatedly performed until it succeeds in obtaining the correct solution. Moreover, computations on state-of-the-art quantum computers are subjected to imprecisions in the gates applied to the qubits, readout errors, and losses of coherence in the state of the system [Pre18], and its limited size prevents the application of quantum error correction schemes.

Therefore, when thinking about a realistic application of the quantum Bayesian algorithm presented in Section 4.3, the important questions to ask are how experimentally feasible it is, and how far we are from running it on real quantum computers. With this goal in mind, we have performed two sets of experiments: on the one hand, we have run simulations of the quantum linear systems protocol on two different quantum virtual machines with various noise models that affect real quantum computers, and analyzed their impact on the output—the final quantum state after the protocol. On the other hand, we have run scaled-down versions of the protocol in two real, state-of-the-art quantum processing units to gauge how far we are from implementations of practical relevance.

We have implemented the complete quantum linear systems protocol in the Rigetti Forest API using PyQuil and Grove [SCZ16]. This implementation can perform approximate eigenvalue inversion on a Hermitian matrix of arbitrary size [PKW18]. The PyQuil framework has advanced gate decomposition features and provides a way to perform arbitrary unitary operations on a multi-qubit quantum state. Furthermore, Rigetti's classical simulator of quantum circuits (referred to as a *quantum virtual machine*) provides a variety of noise models that can affect computations in real quantum architectures, allowing a detailed analysis of how noise affects accuracy and computational overhead.

In addition, we have implemented reduced, 2×2 problems in both PyQuil—to be run in Rigetti's Quantum Processing Unit—and in IBM's QISKit software stack [AAB+19]—to be run in IBM's Quantum Experience computers. QISKit also provides a noisy classical simulator, of which we also make use to contrast the performance of the quantum linear systems algorithm run in the real QPUs against simulations with different noise models.

The quantum processing units employed in the experiments are IBM's 16-qubit Rueschlikon (IBMQX5) [WLYZ18] and Rigetti's 8-qubit 8Q-Agave. While the number of available physical qubits is in both cases higher than the number of qubits required for the implementation (a total of six for the 2×2 reduced version), the depth of the circuit is much higher for larger matrices, and the current noise levels in the QPUs would not allow obtaining meaningful results when solving larger problems.

4.4.1 Simulations on quantum virtual machines

We first report the results of the simulations performed in Rigetti's quantum virtual machine. We have conducted two sets of experiments to analyze the sensitivity of the protocol to different noise types that appear in real quantum computers. In the first, we restrict ourselves to the simplest possible scenario of solving the problem

$$\begin{aligned} 3x_1 + 2x_2 &= 2 \\ 2x_1 + 3x_2 &= 0. \end{aligned}$$

This problem can be written in the form $A\boldsymbol{x} = \boldsymbol{b}$, with $\boldsymbol{x} = (x_1, x_2)$ and

$$A = \frac{1}{2}\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}, \qquad \boldsymbol{b} = (1, 0). \tag{4.12}$$

For solving it, we employ the problem-specific circuit in [CDFK12]. This circuit is much shallower than the full protocol detailed in [CPP$^+$13], making it more realistic to implement on current and near-future quantum computers due to its reduced depth. The second case is the complete implementation of the full quantum linear systems protocol [HHL09, CPP$^+$13]. This version requires a large number of ancillary qubits to perform the calculations, in particular for the computation of the reciprocals of the eigenvalues. We simulate the solution of a 4×4 problem with four bits of precision, as this is the largest example that could fit on the largest Rigetti QPU (19Q-Acorn, which has a total of 19 qubits).

We have studied the impact of two noise models, both being instances of parametric depolarizing noise. The first one, known as *gate noise*, applies a Pauli X operator—which swaps the states $|0\rangle$ and $|1\rangle$ of the qubit it acts upon—with a certain probability on each qubit after *every* gate application. The probability of application of the operator indicates the noise level. The second type of noise that we study is known as *measurement noise*. In this case, a Pauli X operator is applied with certain probability only on every qubit that is measured, just before the measurement takes place. Therefore, it can also be understood as a readout error that, with a certain probability, instead of recording the result of a measurement, y, it records $NOT(y)$.

The circuits we implement have a much larger number of gates (∼20 for the 2×2 reduced version, increasing with increasing size of the problem being solved) than measurements (just one, that which certifies the success of the postselection). This is the reason why in all the experiments run the gate noise has a stronger impact on the results than the measurement noise.

In Fig. 4.1 we show the results for the solution of the 2×2 problem in Eq. (4.12). We analyze the two critical factors of the protocol, namely how

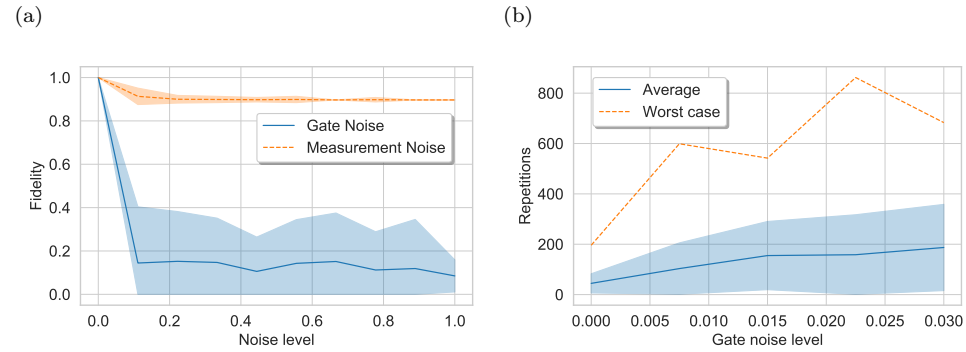

Figure 4.1: Simulated gate and measurement noise on a specialized circuit for solving the problem in Eq. (4.12), run in Rigetti's quantum virtual machine. The shaded regions represent the standard deviations for 100 independent executions. (a) The fidelity shows the overlap with the expected correct state after the computation. A fidelity of 0 means that the output state (the result of the computation) is completely orthogonal to the correct solution, while a fidelity of 1 means that the output state coincides with the expected one. The average fidelity of a random state with the correct solution is $1/2^N$, where N is the number of qubits in the system. (b) The number of repetitions expresses the average of how many times the probabilistic program is executed before it succeeds. We define a successful run with two conditions: the qubit in which the controlled rotation is performed is in the state $|1\rangle$ after its measurement, and the final state of the remaining qubits has a fidelity greater than 0.9 with the expected outcome of an ideal run of the protocol. The dashed line represents the maximum number of runs observed in the simulations before having a successful one.

different the expected result and the output from the noisy simulator are when we know that the postselection has succeeded, and how many times the protocol is needed to run in order to obtain a successful execution. As expected, the measurement noise has a much smaller impact in the result than the gate noise, which for reasonably low noise levels already renders the output state (and hence the result of the protocol) with low overlap with the expected result.

The number of repetitions needed for the algorithm to succeed, understood as the average number of times the algorithm must be run in order to obtain the outcome associated to the state $|1\rangle$ when measuring the qubit to which the conditional rotation is applied, is a fragile quantity that, on its own, does not provide meaningful insights when dealing with noise. In the case of measurement noise, an error either discards a successful run of the algorithm or accepts as successful a failed run, deeming further computations useless in both cases. For gate noise, even if the measurement succeeds and therefore the state of the flag qubit is $|1\rangle$, the remaining computations on the other qubits may lead to a final state that deviates from the expected result.

In order to obtain a good estimation of the number of runs needed to de-

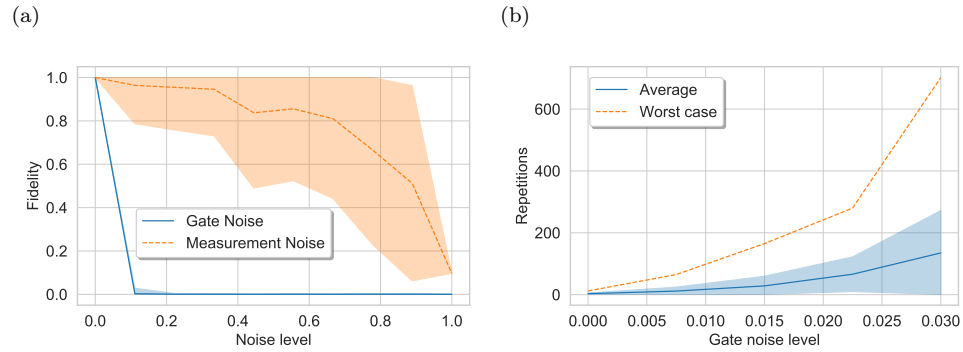

Figure 4.2: (a) Simulated gate and measurement noise, and (b) number of runs until a successful postselection, on the generic circuit for solving linear systems of equations. The matrix in the benchmark was 4×4, and the eigenvalues were represented by four bits of precision. Together with the ancilla qubits in the calculations, making a total of 12 qubits, this is the largest system that can be simulated with less than 19 qubits, which is the size of Rigetti's largest QPU (19Q-Acorn).

tect a final state that encodes the desired solution, in Fig. 4.1b we show the number of repetitions of the algorithm needed in order to have a successful run according to the flag qubit (i.e., that its state after the measurement is $|1\rangle$ [HHL09]), in which the overlap of the final state and the desired state is higher than a specific value. We measure such an overlap with the fidelity, given by $\mathcal{F} = |\langle\psi_{\text{real}}|\psi_{\text{ideal}}\rangle|^2$, where $|\psi_{\text{real}}\rangle$ and $|\psi_{\text{ideal}}\rangle$ determine the state of the qubits after a noisy simulation and a noiseless successful run, respectively.

Given that the protocol is probabilistic, the number of repetitions needed to have a successful run depends on the actual problem being solved even in the case of a noiseless run, as can be observed by comparing Figs. 4.1b and 4.2b, and grows fast with the gate noise level in the qubits. It is important not only to track the average behavior of the protocol (in solid blue in Fig. 4.1b), but also to estimate worst-case scenarios (in dashed orange), where we observe the protocol must be run up to more than five times the average in order to have a successful execution. Nevertheless, worst-case performance scales with the noise level in a similar way as the average performance.

In Fig. 4.2 we perform the same studies for the implementation of the general algorithm solving a well-defined, random problem with four unknowns, where the matrix A involved always has the eigenvalues $\boldsymbol{\lambda} = (1/2, 1/4, 1/8, 1/16)$. It is immediately apparent that increasing the circuit depth makes the protocol more sensitive to noise, and the fidelity drops to almost zero with lower variance in the case of the gate noise. However, the noise level for which the fidelity of the output of the circuit with the expected state drops abruptly is approximately

equal in both the 2×2 and 4×4 cases, and it would be interesting to see whether it remains constant for larger problems. We still observe better robustness to measurement noise, but the impact of this type of noise in the resulting state is stronger than in the problem-specific algorithm of Fig. 4.1. The number of repetitions for a successful run now has a nonlinear behavior with the level of gate noise in the simulation, although the ratio of the worst-case scenario to the average is the same as in the case of solving the 2×2 problem.

4.4.2 Evaluation on quantum processing units

In this section we study the implementation of the restricted 2×2 algorithm for solving the problem in Eq. (4.12) in two real quantum computing architectures. The reason of choosing the restricted algorithm is that current quantum computers have a small number of qubits, limited qubit-qubit connectivity, and most importantly, short coherence times, which implies that only shallow quantum circuits can be implemented. The restricted algorithm needs a much simpler circuit than the general one, resulting in about 20 gates for the full protocol [CDFK12].

In the case of runs on real QPUs one does not have direct access to the whole output state of the circuit, but only to samples of measurements on it. This makes computing the fidelity with the expected state difficult, and instead, we perform a *swap test* [GC01]. The test runs as follows: the expected result of the algorithm is encoded manually in auxiliary qubits, and after performing some joint operations between the qubits encoding the output and those encoding the expected result, a flag qubit indicates whether the states of both sets of qubits are equal, in which case the state of the flag qubit is $|0\rangle$, or not, in which case the state is a weighted superposition of $|0\rangle$ and $|1\rangle$. The figure of merit is now the probability of success in the test, $P(success) = P(0)$, which can then be related to the fidelity by the expression $\mathcal{F} = |2P(success) - 1|$. Note that this success probability is different from the probability that the eigenvalue inversion subroutine succeeds, which is the quantity that has already been studied in Figs. 4.1b and 4.2b.

We have implemented the protocol to be run in both Rigetti's 8Q-Agave and IBM's IBMQX5 quantum processing units. The IBM QISKit software [AAB+19] also provides a classical simulator to run noisy experiments, and we use these to benchmark the performance of the runs on the real chips. The results of the experiments can be found in Fig. 4.3.

As in the case of the simulations in Rigetti's software stack, the measurement noise produces a smaller impact in the protocol than the gate noise. Note that the qubit that encodes the success of the swap test is also subject to readout

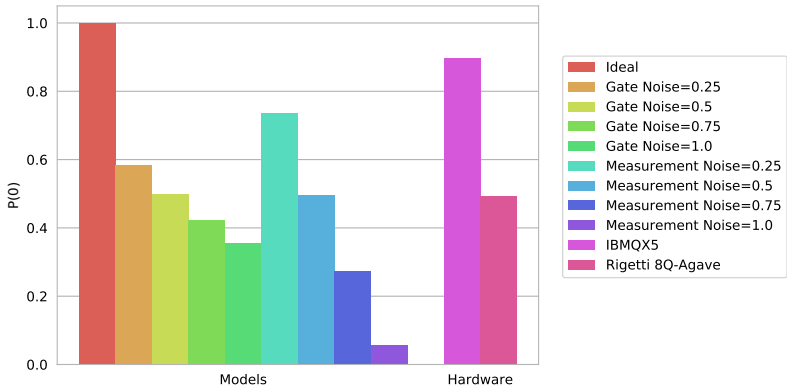

Figure 4.3: Probability of obtaining the outcome 0 in the qubit encoding the result of the swap test after success in the eigenvalue inversion subroutine, for different classical noisy simulations, and executions on the IBM and Rigetti quantum processing units (rightmost bars). The noise models involve faulty gate operations—gate noise—and faulty readout errors—measurement noise—with different probabilities of failure. The algorithm is run 8192 times for each instance, after which $P(0)$ is estimated from the readout results.

error when simulating measurement noise. Therefore, for large measurement noise levels, the fact that $P(0)$ is very low means that the actual state of the flag qubit is $|0\rangle$ (i.e., the protocol has succeeded, and the output state is the desired one), but due to the noise the result that is recorded after performing the measurement is 1. Thus, one should therefore associate $P(success)$ with $P(1)$ instead.

Gate noise has a stronger impact in the final state. This kind of error, unlike the measurement noise, directly affects the computations in the circuit, so lower success probabilities now represent a real discrepancy between the output and desired states. In this case, the success probabilities lie in the range of $[0.35, 0.6]$, which translates into fidelities in the range of $[0, 0.3]$.

Turning to executions in the real QPUs, the probability of protocol success is higher in IBMQX5. This is mostly due to its improved coherence time[6,7], that allows keeping the state in the circuit better isolated from external perturbations during computation. The probability of protocol success is 89%, which translates into a fidelity with the expected state of 0.78. This is a very encouraging result, despite the size of the problem being solved. In contrast, $P(0) \approx 1/2$ when the protocol is run in 8Q-Agave. This result can be explained by the large

[6] Information about chip architecture and performance measures of Rigetti's QPUs can be found in http://docs.rigetti.com/en/1.9/qpu.html

[7] Information about performance measures of IBM's QPUs can be found in http://www.research.ibm.com/ibm-q/technology/devices/

gate cost of the embedding of the controlled-SWAP gate [YY15], needed for the swap test, in the chip architecture[6]. The number of implementable gates required to perform the controlled-SWAP may be large enough to exceed the coherence time of the qubits, rendering the flag qubit in a completely mixed state by the end of the test. This qubit loses all its information about the success of the protocol, making the quantity $P(0)$ deprived of its relation to the actual success of the swap test, and to the quality of the execution of the HHL algorithm.

4.5 Discussion

As quantum computers become available and continue improving in scale and noise tolerance, it is an exciting question to ask whether they can make a qualitative difference in machine learning applications. However, an important fact is that for at least the next decade quantum computers will remain limited in both properties, and this must be factored in when constructing quantum-enhanced algorithms. Furthermore, simple machine learning methods are already efficiently executed on classical hardware, so efforts should be directed towards the development of quantum algorithms for solving problems that have a prohibitive computational cost in classical computers.

With this motivation, in this chapter we studied a complex, Bayesian approach to training deep learning architectures, that is difficult to perform on digital hardware. We developed a quantum algorithm for learning Gaussian processes that can be applied for training arbitrarily deep neural networks. In order to analyze the feasibility of a real use of the algorithm, we implemented its core routine, the quantum linear systems protocol, to be executed in both classical simulators of quantum circuits and real state-of-the-art quantum processors.

Although promising, the experimental results obtained do not completely prove that the full protocol will be efficiently implementable in near-term quantum technologies. Full implementation in architectures with limited coherence time and sparse connectivity is an interesting avenue for future research. Moreover, the implementation used performs a rough estimation of the reciprocals of the eigenvalues involved in the HHL algorithm. After this work, other proposals [AAB+19] were developed that performed this step with a higher precision, at the expense of a large increase of the circuit depth. An interesting problem to pursue is the design of a fully-coherent, low-depth algorithm for eigenvalue inversion, potentially using a variational classical-quantum form [MRBAG16]. This would additionally have important applications in quantum simulations

and quantum control.

On the theoretical side, the sparsity and well-conditioning requirements of the HHL algorithm may restrict the form of the nonlinear activation function or other parameters in the network architecture. An analysis of such constraints will be of importance for assessing which problems can benefit from the application of the algorithm proposed.

5 | Quantum information for understanding causality

Causality is a seminal concept in science: any research discipline, from sociology and medicine to physics and chemistry, aims at understanding the causes that could explain the correlations observed among some measured variables. In this chapter we address the question of identifying the causal explanation underlying some observed correlations, including the case where the causes have quantum degrees of freedom. We propose two different methods for analyzing correlations achievable in quantum causal scenarios. After an introduction to the general frameworks of causal compatibility and quantum correlations, we describe both methods along with results obtained regarding device-independent certification of multipartite nonlocality. We show how both techniques identify correlations incompatible with quantum realizations in specific causal structures, and use them to characterize the quantum correlations that are achievable in simple scenarios. Finally, we contrast the methods against each other and discuss their application to the study of causality in the classical realm.

5.1 Introduction

It can be argued that one of the main challenges in any scientific discipline is to identify which causes are behind the correlations observed among some measured variables, encapsulated by their joint probability distribution. Understanding this problem is crucial in many situations, such as, for example, the development of medical treatments, taking data-based social policy decisions, the design of new materials or the theoretical modeling of experiments.

A celebrated feature of quantum theory is the paradigm shift it necessitates in our fundamental conceptions of probability and causality [H$^{\otimes 4}$09, BCP^{+}14, PMA^{+}15, CRGB18]. As proven by the violation of Bell inequalities [Bell64], quantum theory is nonlocal, in the sense that there exist correlations between outcomes of measurements performed on distant entangled quantum systems that are incompatible with any explanation involving just local hidden variables (LHV). Quantum nonlocality is a powerful resource that grounds protocols for secure cryptography [Eke91, BHK05, ABG^{+}07], randomness certi-

fication [Col07, PAM+10], self-testing [MY04, ŠB19] or distributed computing [BCMdW10]. Therefore, it is crucial to develop ways to test the incompatibility of a correlation with models involving only classical resources, that is, to detect whether the correlation contains nonlocal features that can be harnessed.

When presented with some information-theoretic task (such as commmunication, coordination, or computation), a frequent research desideratum is to assess what quantum advantage, if any, may be found by contrasting the quantum and classical variants of that task. Thus, one can isolate two separate research directions, both of which are critical to advancing quantum information theory. The first is the derivation of classical constraints, which are used to delineate the boundary separating classical statistics from their nonclassical counterparts [BCP+14, CF12]. The second is the derivation of constraints for quantum resources, in order to quantify the degree to which access to quantum correlations is actually valuable [Cir80, WY12, PV09]. An example of the first research direction is the derivation of Bell inequalities [WW01, ŻB02], whereas calculating the maximal violation of a Bell inequality in quantum theory [WY12, PV09] is exemplary of the second research direction. A related subject of study is the optimization of tasks in quantum theory while constraining the types of states considered [MBL+13, LRB+15, SAT+17], which offers insight into the value of different states for different tasks.

On the practical side, fast progress towards advanced demonstrations of quantum communication networks nowadays requires to go beyond the two-party scenario and characterize networks of growing complexity, providing the tools to witness the nonclassicality of quantum correlations. To that aim, the framework of causal networks [Pea09] and its quantum generalizations [Fri12, HLP14, CMG15, RAV+15, PB15, CS16, ABH+17] have played an insightful role. Causal networks not only allow one to derive Bell's theorem from a causal inference perspective [WS15, CKBG15] but also provide generalizations to more complex scenarios such as quantum networks with several sources [Cha16, BGP10, BRGP12, Fri12] or involving communication among the parties [CCA+18, vHBP+18, CCA17].

Despite all recent advances, the understanding of the structure of correlations in networks remains very limited. The most general method to characterize classical network correlations relies on algebraic geometry [GM99] and, in practice, is restricted to very simple cases. Motivated by that, alternative methods have been proposed [CLM+14, Cha16, KvPÅ+17, WSF19] that, however, do not have a clear path for a quantum generalization. The development of a general and numerically-efficient method for analyzing the quantum correlations achievable in causal networks is very much desired, both for the fundamental understanding of quantum theory and its application in technology.

5.2 Preliminaries

The characterization, on the one hand, of correlations achievable in classical causal networks, and on the other hand, of correlations achievable within quantum mechanics, are both very active research topics. In this section we review their fundamentals, which we make use of in the remaining of the chapter.

5.2.1 Causal scenarios

A natural tool used for reasoning about correlations in causal scenarios is Bayesian networks [Pea09, Fri12]. These are directed acyclic graphs (DAGs) that encode hypotheses on the causal structure underlying the correlations observed among different variables. Formally, a DAG is defined as

Definition 2 (DAG)**.** A directed acyclic graph is a pair (V, E) of a set of nodes $V = \{X^{(i)}\}_i$ and a set of directed edges $E \subseteq V \times V$ with no cycles, where a cycle is a sequence of edges of the form $X^{(k_1)} \to X^{(k_2)} \to \cdots \to X^{(k_n)} \to X^{(k_1)}$.

In Bayesian networks, nodes represent random variables and directed edges represent causal relations: an arrow $X^{(1)} \to X^{(2)}$ indicates that the variable $X^{(1)}$ has an influence on the value of the variable $X^{(2)}$. Therefore, given the notion of time implied by a causal order, the acyclicity property just means that one variable cannot influence itself, preventing the existence of closed timelike curves [Deu91] or, more colloquially, "kill-your-grandfather"-like paradoxes.

There are two possible types of nodes in causal networks. On the one hand, *observed variables* are those that are measured in an experiment. For instance, these would be the drug dose or the cholesterol levels in a clinical trial for a new, cholesterol-reducing drug, or the readouts of a measurement device in a quantum experiment. On the other hand, there can also be *unobserved* or *latent variables*, which are not measured but can ultimately affect the values of the observed variables. Examples of these can be some genetic bias of the patient in the trial, or the actual quantum state measured in the experiment.

Interestingly, typical experiments in quantum information science such as a Bell test or entanglement swapping, can easily be represented in the language of Bayesian networks. Sources preparing states, which are not directly empirically observable, are represented by latent variables, while measurement choices and outputs define the observed variables, as represented in Fig. 5.1.

Correlations achievable in a specific causal scenario have a strong dependence in the properties of the latent variables, for instance, its cardinality [RGW18]. Throughout this chapter, we will focus on the nature of the variables instead, and whether having classical or quantum latent variables allows for different

correlations to be achieved. It is thus important to introduce the concept of networks with a *classical-quantum separation*.

Definition 3. A network provides a classical-quantum separation whenever there exist quantum sources producing correlations among the observed variables that are impossible to attain by classical means in the same network.

The scenario underlying a Bell test is an example of a network with a classical-quantum separation, but not the only one [BGP10, BRGP12, Fri12, HLP14, Pie17, CCA+18, vHBP+18].

Causal inference in classical scenarios

When all variables in a network are observed, there exist general procedures to address whether a probability distribution over the variables could have been generated in the network. In fact, all constraints imposed by the network structure are captured by the notion of *Markovianity*, this is, that the value of any variable in the graph is only dependent on the value of the variables that have a direct influence on it. For instance, the application of the Markov condition to the causal structure underlying the Bell scenario in Fig. 5.1a imposes the following factorization of compatible probability distributions[1]:

$$P(A,B,X,Y,\lambda) = P(A|X,\lambda)P(B|Y,\lambda)P(X)P(Y)P(\lambda). \tag{5.1}$$

This factorization already encodes relevant information such as the no-signaling conditions, namely $P(A|X,Y) = P(A|X)$ and $P(B|X,Y) = P(B|Y)$. While the factorization (5.1) would be a sufficient condition for a distribution $P(A,B,X,Y,\lambda)$ to be compatible with the causal structure were λ a visible variable, in practice it is latent, and thus one only has access to the marginal $P(A,B,X,Y) = \sum_\lambda P(A,B,X,Y,\lambda)$. In fact, it is usual to factor out $P(X)$ and $P(Y)$—which describe the distributions over input choices for each party—and just consider the marginal distribution

$$P(A,B|X,Y) = \sum_\lambda P(\lambda)P(A|X,\lambda)P(B|Y,\lambda). \tag{5.2}$$

It was first noted by Bell that probability distributions of this form satisfy certain algebraic inequalities [Bell64] and therefore it is possible to discern,

[1] Throughout this chapter we will use P to denote both probability distributions over random variables and probabilities for specific assignments of values of the variables. Thus, we will write $P(a,b,\dots)$ as a shorthand for $P(A{=}a, B{=}b, \dots)$.

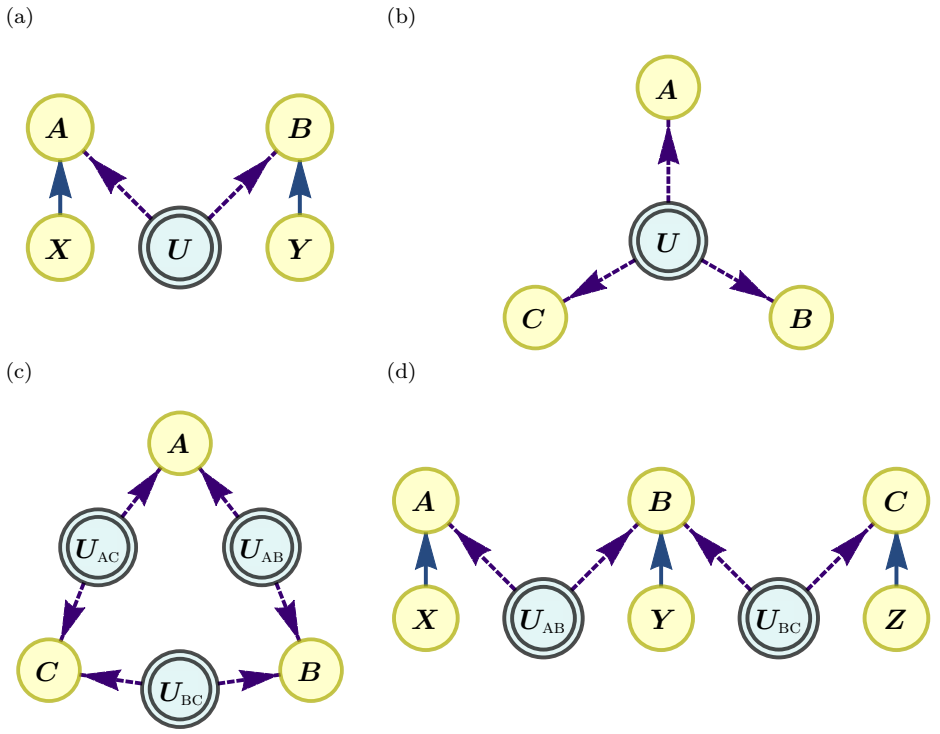

Figure 5.1: DAG representation of different causal scenarios. The blue, double-lined circles are latent variables, and the yellow, single-lined circles denote visible variables. (a) The Bell scenario, which is one of the simplest causal structures exhibiting a classical-quantum gap. The classical, parentless variables X and Y represent measurement choices, while the classical, childless variables A and B represent measurement outcomes. The term *Bell scenarios* describes all causal structures where every childless variable is influenced by a corresponding parentless visible variable (representing a measurement choice) and global latent variable. (b) depicts the most general way of correlating three variables that does not involve communication between them. If every visible variable was influenced by a measurement choice, then this would represent the tripartite Bell scenario. A first restriction to the general scenario in (b) is the *triangle* scenario in (c). Here, every two parties are influenced by a common bipartite latent variable. Removing one of these latent variables leads to the *line* scenario of (d) (depicted with setting variables), where now parties A and C are causally disconnected.

even only with the partial information given by Eq. (5.2), distributions that are compatible with the causal scenario of Fig. 5.1a.

On the other hand, given a probability distribution over some random variables, the Inductive Causation algorithm of Pearl and Verma generates the smallest DAG with only visible nodes that can give rise to it [PV91]. The algorithm consists in finding all conditional independences between the variables in the distribution and using these to recover the graph according to some specified rules.

As briefly outlined above, assessing whether a given correlation is compatible with a causal network with latent variables is, in general, much harder than in the case where all variables are visible. The reason for this is, mainly, that constraints that are possible to examine experimentally can only refer to observable variables. This requires, in general, solving nonlinear quantifier elimination problems. While the Tarski-Seidenberg theorem provides an algorithm for obtaining exact solutions to this type of problems, its computational complexity is far too large to be practical, except in the case of particularly simple scenarios [GM98, LS17]. Over the years, several approaches have been attempted [GM99, CLM$^+$14, Cha16, KvPÅ$^+$17, WSF19], whose main idea was to perform different relaxations of the original problem into more tractable ones. A notable example is the inflation technique of [WSF19], which provides a hierarchy of necessary (later also proven to be asymptotically sufficient [NW17]) conditions for a probability distribution to be compatible with a given causal structure. Due to its importance, and its relevance to the current work, we review this technique in the following section.

5.2.2 The inflation technique

The inflation technique [WSF19] primarily addresses the problem of *classical causal compatibility*: given a causal structure and a probability distribution over its visible variables, inflation allows one to discern whether the distribution could have been generated in the causal structure when the latent variables are classical.

For doing so, it proceeds by contradiction. Assume that there exists an assignment of causes according to the candidate structure topology explaining the observed correlations P_{obs}. Then, [WSF19] proved that there should exist probability distributions P_{inf} defined on larger causal structures, termed *inflations*, that should satisfy some conditions that relates them to P_{obs}. Whenever an inflation is found for which the associated P_{inf} cannot satisfy all required constraints, P_{obs} is identified as not compatible with the original causal scenario.

An inflation of a causal structure consists of another causal structure that

contains copies of the nodes in the original. These nodes are connected in such a way that every copy of a visible variable keeps the *parent structure* of the original node. This is, every copy of a visible variable is only influenced by one copy of each of the variables that influence the original one. Fig. 5.2 depicts some inflations of simple causal structures. In the inflated structures, one should be able to define probability distributions P_{inf} that satisfy two main constraints. On the one hand, since all copies of a same variable should behave in the same way, P_{inf} should be invariant under permutation of copies of the variables. On the other hand, when marginalized over subsets of variables that are equivalent to the original scenario, the corresponding marginal distribution should coincide with P_{obs}. Both of these types of constraints can be written as linear equalities of the elements of P_{inf}. This makes that the problem of determining the existence of a P_{inf} satisfying such constraints can be cast as a linear program, which can be efficiently solved.

An illustration may be useful at this point. Consider the triangle scenario of Fig. 5.2b. In order to test whether a given $P_{\text{obs}}(A, B, C)$ is compatible with the triangle scenario, one can consider the problem of the existence of a P_{inf} in the inflated structure of Fig. 5.2d. Note that this structure contains two copies of each latent variable and four copies of each visible variable, since this is the maximum number of variables that can be affected by a different pair of copies of latent variables. The copies of a visible node are precisely indexed by the pair of copies of the latent variables that influences it. P_{inf} is defined over all visible variables $\{A^{i,j}\}_{i,j}$, $\{B^{k,l}\}_{k,l}$, $\{C^{m,n}\}_{m,n}$. The first constraint, known as *automorphic symmetry*, is expressed as

Automorphic symmetry

$$\begin{aligned} &P_{\text{inf}}(\{A^{i,j}{=}a^{i,j}, B^{k,l}{=}b^{k,l}, C^{m,n}{=}c^{m,n}\}) \\ &\quad = P_{\text{inf}}(\{A^{i,j}{=}a^{\pi(i),\pi'(j)}, B^{k,l}{=}b^{\pi'(k),\pi''(l)}, C^{m,n}{=}c^{\pi''(m),\pi(n)}\}), \end{aligned} \tag{5.3}$$

where π, π' and π'' are three independent permutations of the indices denoting the copies of the latent variables. These constraints impose the restrictions associated to the inflated causal structure being an inflation of the original, by requiring invariance under the swap of copies of a same latent variable. As such, they can be used not only for addressing problems of causal compatibility, but also when one desires to obtain the optimal value of functions of the correlations that can be achieved in a given causal structure. Eq. (5.6) shows an example of this type of problems.

Then, the second type of constraints impose the relations to the observed distribution, in the case of causal compatibility. In the case of the inflation in Fig. 5.2d, one can observe that the marginalization over the visible variables

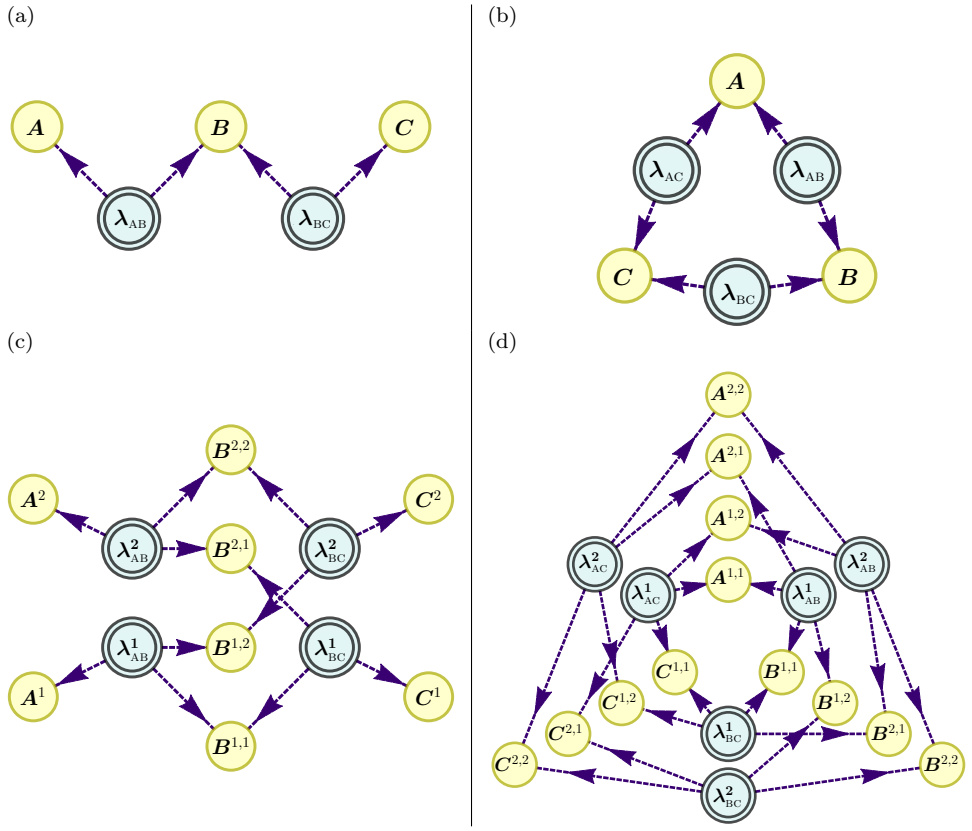

Figure 5.2: Examples of classical inflation of causal structures. (a) The tripartite-line causal structure of Fig. 5.1d, and (c) its second-order classical inflation. Note that every copy of the visible variables has the same parent structure than the original variable in (a). (b) The triangle causal structure with classical latent variables, and (d) its second-order classical inflation.

$A^{i,i}$, $B^{i,i}$ and $C^{i,i}$ is a graph composed of two copies of the original triangle scenario. Therefore, this marginalization should satisfy

Scenario identification

$$P_{\text{inf}}(\{A^{i,i}{=}a^{i,i}, B^{i,i}{=}b^{i,i}, C^{i,i}{=}c^{i,i}\}) = \prod_i P_{\text{obs}}(a^{i,i}, b^{i,i}, c^{i,i}). \tag{5.4}$$

In fact, there are more constraints of this kind that can be imposed, involving products of marginals of P_{obs} of greater degree than the number of copies of a latent variable. An example in the inflation of Fig. 5.2d is $P_{\text{inf}}(A^{1,2}{=}a, B^{1,2}{=}b, C^{1,2}{=}c) = P_{\text{obs}}(a)P_{\text{obs}}(b)P_{\text{obs}}(c)$.

The way of creating inflations we have followed, where there are n copies of each latent variable and the maximum number of copies of the visible variables possible, gives rise to a hierarchy in n of inflations of a causal structure. Built in this way, a level-n inflation graph of a given causal structure contains the inflation graph of level $n{-}1$. As was proven in [NW17], the set of distributions that can be written as marginals of a P_{inf} over a copy of the original scenario approximates the set of compatible distributions with the original scenario from the outside, more and more tightly with increasing n, until it converges for $n \to \infty$. Therefore, the inflation techniques provides a necessary and sufficient hierarchy of conditions that allow one to solve the problem of classical causal compatibility.

Moreover, inflation can be used to bound optimal values of polynomials of elements of probability distributions compatible with a causal structure. These problems are written as

$$\begin{aligned} g_* \;=\; & \underset{P}{\text{maximize}} && g(P) \\ & \text{subject to} && P \in \aleph, \end{aligned} \tag{5.5}$$

where $\aleph$ is the set of distributions compatible with a given classical causal scenario. The key is relaxing this problem to P being the appropriate marginalization of a P_{inf} compatible with an inflation. For the example of the triangle scenario, this would be written as

$$\begin{aligned} g_n \;=\; & \underset{P}{\text{maximize}} && g(P) \\ & \text{subject to} && P = P_{\text{inf}}(A^{1,1},B^{1,1},C^{1,1}), \\ & && \text{equation (5.3)}. \end{aligned} \tag{5.6}$$

For increasing n, the hierarchy gives a sequence of upper bounds satisfying

$$g_1 \geq g_2 \geq \cdots \geq g_\infty = g_*.$$

However, it must be noted that this procedure can be used for optimizing only polynomials $g(P)$ that can be written as linear combinations of the elements of P_{inf}. For more general functions $g(P)$, the problem (5.6) is not a linear program any more, and thus an efficient algorithm for obtaining the solution may not exist.

Inflation with quantum latent variables

It has been already mentioned that the inflation technique enables solving the *classical* causal compatibility problem. Unfortunately, a straightforward adaptation to the quantum case is not possible. The reason lies in the concepts of monogamy of correlations and no-broadcast of information, which are unique to nonclassical probability theories such as quantum mechanics.

Consider, for instance, the tripartite-line scenario of Fig. 5.2a and its inflation in Fig. 5.2c, or the triangle scenario of Fig. 5.2b and its inflation in Fig. 5.2d. Imagine that the node λ_{AB} sends two maximally correlated bits to nodes A and B. Then, the node λ^1_{AB} in the inflations should be such that both bits going to $B^{1,1}$ and $B^{1,2}$ are perfectly correlated with the bit sent to A^1 ($A^{1,1}$, in the case of the triangle scenario). However, if λ_{AB} distributed a maximally entangled quantum state to A and B, the monogamy of entanglement [CKW00] prevents the existence of a state that is maximally entangled between A^1 (or analogously $A^{1,1}$) and $B^{1,1}$, and simultaneously between A^1 (analogously $A^{1,1}$) and $B^{1,2}$. This is, in fact, the property that allows inflation to solve classical causal compatibility, as it draws a separation between classical probability theory and more general theories.

It is possible to construct inflations that do not suffer of this problem, known as *non-fanout* inflations. An example for the triangle scenario is given in Fig. 5.3. Given that quantum and more general realizations can be built in these inflated networks, non-fanout inflations are not capable of detecting nonclassical correlations that would otherwise be compatible with the original scenario. Therefore, they cannot be used to distinguish quantum from more general correlations, and only be applied to discern if a distribution satisfies the constraints imposed by the causal graph itself.

Furthermore, recent results on quantum correlations [Slo19] imply that the problem of quantum causal compatibility is undecidable. Yet, this does not preclude the existence of a method similar to inflation to tackle the question. We present such method in Section 5.4.

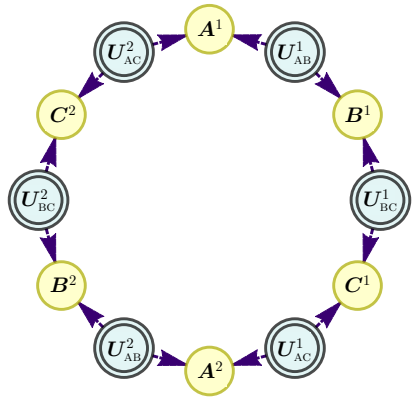

Figure 5.3: Second-order ring inflation of the triangle scenario. Since no information is broadcast, this is a non-fanout inflation. While imposing constrints related to the causal structure, this type of inflations is not capable of identifying correlations not compatible with realizations in terms of classical, quantum, or more general latent variables.

5.2.3 Local, quantum and post-quantum correlations

One of the central tasks of quantum information consists in determining whether some given observations of measurement outcomes are genuinely quantum or, in the contrary, they can be reproduced using classical systems. This question has a very intuitive formulation when taking a geometrical approach. This approach considers probability distributions $P(o_1, \ldots, o_n|i_1, \ldots, i_n)$ as multidimensional vectors in real space $\mathbb{R}^D$, and then studies the subsets of $\mathbb{R}^D$ that are determined by the different physical theories[2]. Historically, three sets of correlations have been extensively studied, which we describe below and are pictorially represented in Fig. 5.4. For simplicity, we assume that the underlying causal scenario is the multipartite generalization of the Bell scenario in Figs. 5.1a, 5.1b, where the different parties are isolated of each other, and the only source of correlation is some globally-shared resource.

The local set $\mathcal{L}$

Local correlations (also commonly referred to as *classical*) are those that can be reproduced by a *local hidden variable* model,

$$P(o_1, \ldots, o_n|i_1, \ldots, i_n) = \sum_\lambda p_\lambda \prod_{k=1}^{n} P(o_k|i_k, \lambda), \quad p_\lambda \geq 0, \quad \sum_\lambda p_\lambda = 1. \quad (5.7)$$

[2] In a scenario with n parties, each of which can perform m measurements obtaining on each one out of d possible outcomes, $D = (md)^n$.

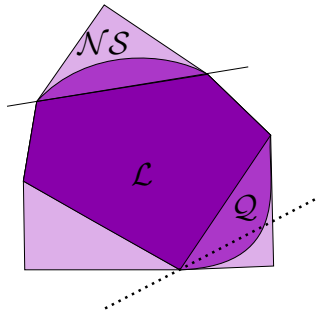

Figure 5.4: Pictorial representation of the different sets of correlations of interest in quantum information. The local set $\mathcal{L}$ contains probability distributions generated by LHVs through Eq. (5.7). The quantum set $\mathcal{Q}$ contains distributions achievable via the Born's rule in Eq. (5.9). The nonsignaling set $\mathcal{NS}$ contains all correlations that are achievable by just imposing the notion of spacelike separation, encoded by Eq. (5.12). The solid line represents a Bell inequality, that separates $\mathcal{L}$ from $\mathcal{Q}$. The dashed line represents a nonlocality witness, whose violation certifies nonlocal correlations, but unlike Bell inequalities its satisfaction does not certify the existence of a local model.

This model—which we derived for two parties in Section 5.2.1 from the Markov condition—represents the action of a global source of correlations that sends the (classical) information λ to the parties with probability p_λ, and they use this information and its respective input i_k to generate an outcome o_k. Thus, the local set contains all correlations that are achievable with classical systems. Distributions that lie outside the local set are termed *nonlocal*. It is precisely the nonlocal distributions those that are of interest in quantum information theory, as they can not be generated with classical resources.

The local set is convex, as it can easily be seen from Eq. (5.7). Moreover, it can be shown that it is in fact a polytope. To do so, it suffices to notice that the functions $P(o_k|i_k,\lambda)$ can be chosen without loss of generality to be deterministic functions $D(o_k|i_k,\lambda)=\delta_{o_k,\lambda(i_k)}$, where $\lambda(\cdot)$ is a deterministic function $\lambda:\{0,\ldots,m-1\}\to\{0,\ldots,d-1\}$.

Interpreting the deterministic functions $D(o_k|i_k,\lambda)$ as vectors in $\mathbb{R}^{mn}$, and considering their Cartesian products for a fixed λ, $D_\lambda=\prod_k D(o_k|i_k,\lambda)$, Eq. (5.7) tells us that any point $P\in\mathcal{L}$ can be written as a convex combination of the *deterministic strategies* D_λ. These vectors are the vertices of the polytope $\mathcal{L}$, and can be enumerated by counting the different possible deterministic assignments of the d outcomes of each party for each of the possible m inputs, which sum up to d^{mn}.

Therefore, to know whether a particular P is within the local set, one solves the problem of writing P as a convex combination of the vertices D_λ. This is a

feasibility problem that can be addressed with linear programming. Although an enumeration of the vertices of $\mathcal{L}$ is possible for any Bell scenario, the number of elements to be enumerated grows exponentially with both the number of parties and the number of inputs of each party. In fact, it has been proven that determining whether a point P belongs to $\mathcal{L}$ is in general NP-complete [Avi04] and, due to the above, determining whether a specific P is local via vertex enumeration is actually efficiently implementable only for cases with small values of n and m.

An alternative approach can be taken using the fact that any polytope admits a dual description in terms of a finite set of facets. These facets are linear inequalities in terms of the probabilities

$$\sum_{o_1,\ldots,o_n} \sum_{i_1,\ldots,i_n} \alpha_{i_1,\ldots,i_n}^{o_1,\ldots,o_n} P(o_1,\ldots,o_n|i_1,\ldots,i_n) \geq \beta_C \tag{5.8}$$

which are satisfied by all $P \in \mathcal{L}$. These facets are usually referred to as *Bell inequalities* and the quantity β_C is the *local bound*. Bell inequalities are widely used in quantum information theory, as they typically allow for intuitive interpretations as games for which the parties want to maximize the score [BCP+14]. Geometrically, a Bell inequality is a hyperplane that conforms a boundary of $\mathcal{L}$. Importantly, any nonlocal distribution violates at least one Bell inequality.

Using the characterization of $\mathcal{L}$ in terms of its facets instead of in terms of its vertices does not reduce the complexity of determining if a point is local or not. Nevertheless, many Bell inequalities are known that can detect nonlocal correlations produced by quantum states, including scenarios consisting of an arbitrary number of inputs, outputs or parties [BCP+14].

An object similar to Bell inequalities that are useful in detecting nonlocal correlations are *nonlocality witnesses*. These are inequalities like those in Eq. (5.8), that do not intersect the interior of $\mathcal{L}$, i.e., $\mathcal{L}$ is contained in one of the half-spaces defined by the witness. Nonlocality witnesses, if violated, certify that a point is nonlocal, since a violation proves that the point belongs to the half-space that does not contain $\mathcal{L}$. An example of such a witness is represented as a dashed line in Fig. 5.4.

The quantum set $\mathcal{Q}$

Quantum correlations are defined as those that can be obtained through the application of the Born rule. That is, a multipartite probability distribution $P(o_1,\ldots,o_n|i_1,\ldots,i_n)$ has a quantum realization if and only if there exists a quantum state $|\psi\rangle$ and measurement projectors $\{\Pi_{o|i}^p\}_{i,o}$ for $p = \mathrm{A}, \mathrm{B}\ldots$ such

that

$$P(o_1,\dots,o_n|i_1,\dots,i_n) = \langle\psi| \bigotimes_{k=1}^{n} \Pi_{o_k|i_k}^{p_k} |\psi\rangle \,. \tag{5.9}$$

In Eq. (5.9), the state $|\psi\rangle$ is a normalized vector in a Hilbert space $\mathcal{H}$, while the measurement projectors are positive-semidefinite operators that satisfy the additional following conditions:

i) Orthogonality and idempotency: $\Pi_{o|i}^{p}\Pi_{o'|i}^{p} = \delta_{o,o'}\Pi_{o|i}^{p} \quad \forall\, p, o, o', i,$

ii) Hermiticity: $\Pi_{o|i}^{p} = \left(\Pi_{o|i}^{p}\right)^{\dagger} \quad \forall\, p, o, i,$

iii) Normalization: $\sum_o \Pi_{o|i}^{p} = \mathbb{1} \quad \forall\, p, i.$

When no restriction is imposed on the dimensions of the local Hilbert spaces where each set of projectors acts on, the set $\mathcal{Q}$ is convex. It is however not a polytope, and in fact its faces exhibit interesting characteristics [GKW+18]. The same set can be obtained when fixing the dimension of the local Hilbert spaces, and using as sets of operators positive-operator valued measures (POVMs), which satisfy conditions (ii) and (iii), but not (i).

The quantum set is a strict superset of $\mathcal{L}$, since any multipartite probability distribution admitting a LHV model can be written in terms of the Born rule. To see this, it is useful to write Eq. (5.9) in terms of density matrices. A completely equivalent form of Eq. (5.9) is

$$P(o_1,\dots,o_n|i_1,\dots,i_n) = \mathrm{Tr}\left[\rho \bigotimes_{k=1}^{n} \Pi_{o_k|i_k}^{p_k}\right], \tag{5.10}$$

where the state ρ is a Hermitian, positive-semidefinite matrix with trace 1. The most general form of a state without genuinely quantum correlations is $\rho = \sum_\lambda p_\lambda \bigotimes_{k=1}^{n} \rho_{p_k,\lambda}$, with $p_\lambda > 0$ and $\sum_\lambda p_\lambda = 1$. These states are known as *separable* states, for which Eq. (5.10) reads

$$\begin{aligned} P(o_1,\dots,o_n|i_1,\dots,i_n) =& \mathrm{Tr}\left[\sum_\lambda p_\lambda \left(\bigotimes_{k=1}^{n} \rho_{p_k,\lambda}\right)\left(\bigotimes_{k=1}^{n} \Pi_{o_k|i_k}^{p_k}\right)\right] \\ =& \sum_\lambda p_\lambda \prod_{k=1}^{n} \mathrm{Tr}\left(\rho_{p_k,\lambda}\Pi_{o_k|i_k}^{p_k}\right). \end{aligned}$$

This is exactly Eq. (5.7) upon defining $P(o_k|i_k,\lambda) = \mathrm{Tr}(\rho_{p_k,\lambda}\Pi_{o_k|i_k}^{p_k})$. In quantum theory there exist also nonseparable or *entangled* states, that can not be written as convex combinations of product states. These states give rise, via Eq. (5.9), to distributions that do not admit LHV models, and thus $\mathcal{L} \subsetneq \mathcal{Q}$.

The no-signalling set $\mathcal{NS}$

The set $\mathcal{Q}$ encompasses all possible correlations we expect to find in nature. However, from a fundamental point of view, it is interesting to wonder about more general correlations, and study how physical principles constrain them to recover $\mathcal{Q}$ [NW09, FSA+13, SGAN18].

The most general multipartite distributions satisfy three basic constraints. To begin with, in order to be proper probability distributions, they have to be normalized and their elements must be nonnegative numbers,

$$P(o_1, \dots, o_n | i_1, \dots, i_n) \geq 0 \quad \forall\, o_1, \dots, o_n,\, i_1, \dots, i_n, \tag{5.11a}$$

$$\sum_{o_1,\dots,o_n} P(o_1, \dots, o_n | i_1, \dots, i_n) = 1 \quad \forall\, i_1, \dots, i_n. \tag{5.11b}$$

The last constraint is motivated by physics, and can be interpreted as derived from Special Relativity. If the parties are not allowed to communicate with each other—this can be enforced by having the parties physically separated in space and requiring them to perform their measurements simultaneously—then the local observations of one of the parties should be the same regardless the measurements the other parties perform, since the information that a measurement has been performed travels at finite speed. This is,

$$\sum_{o_k} P(\dots, o_k, \dots | \dots, i_k, \dots) = \sum_{o_k} P(\dots, o_k, \dots | \dots, i'_k, \dots), \tag{5.12}$$

for all i_k, i'_k, all $o_1, \dots, o_{k-1}, o_{k+1}, \dots, o_n$, all $i_1, \dots, i_{k-1}, i_{k+1}, \dots, i_n$, and all $k = 1, \dots, n$.

Conditions (5.11a), (5.11b) and (5.12) are the smallest set of conditions any reasonable physical theory must satisfy. In particular, all points in either $\mathcal{L}$ or $\mathcal{Q}$ satisfy them, among with the remaining constraints described earlier. The set of correlations restricted only by conditions (5.11a), (5.11b) and (5.12) is called the *no-signaling set* and typically denoted $\mathcal{NS}$. It is easy to see that $\mathcal{NS}$ is also a polytope, since it is defined by a finite amount of linear equalities and inequalities.

There exist points in $\mathcal{NS}$ that achieve values of Bell inequalities not attainable by any point in $\mathcal{Q}$ [PR94], and thus, $\mathcal{Q}$ is a strict subset of $\mathcal{NS}$.

5.2.4 The NPA hierarchy

Finding a complete characterization of the quantum set $\mathcal{Q}$ is a very important question that remains still open in general [Slo19]. To date, the most effective method to study the properties of quantum correlations is based on a hierarchy of outer approximations to $\mathcal{Q}$, known as the NPA hierarchy. It is based in the following observation: assume a set $\mathcal{S} = \{S_i\}_i$ of operators on the Hilbert space $\mathcal{H}$, defined in Section 5.2.3. The positivity of the quantum state ρ implies that the matrix Γ, with elements given by[3]

$$\Gamma_{i,j} = \mathrm{Tr}\left(\rho S_i^\dagger S_j\right) \qquad S_i, S_j \in \mathcal{S} \tag{5.13}$$

is always positive-semidefinite, for any choice of $\mathcal{S}$.

A possible way to test whether a given probability distribution admits a quantum realization is choosing a set $\mathcal{S}$ composed of products of the parties' measurement operators and building a moment matrix of the form (5.13). In this matrix, some entries will be computable from the distribution under scrutiny, while the rest remain as variables related by linear constraints that arise from the properties i–iv. The nonexistence of a variable assignment that makes Γ positive-semidefinite is a proof that the distribution is not compatible with a quantum realization.

This is the logic behind the so-called Navascués-Pironio-Acín (NPA) hierarchy [NPA07, NPA08]. There, the proposal is to use a hierarchy of moment matrices $\Gamma^{(n)}$, each associated to the corresponding operator set

$$\mathcal{S}_n = \mathcal{S}_{n-1} \cup \{\text{products of } n \text{ projectors}\}, \tag{5.14}$$

with $\mathcal{S}_0 = \{\mathbb{1}\}$. The sets of probability distributions for which there exists a $\Gamma^{(n)} \succeq 0$ (typically denoted as $\mathcal{Q}_n$) constitute outer relaxations of the set of quantum correlations that satisfy $\mathcal{Q}_1 \subseteq \mathcal{Q}_2 \subseteq \ldots \subseteq \mathcal{Q}_\infty = \mathcal{Q}$. Fig. 5.5 contains a pictorial representation of this hierarchy.

Another complete hierarchy that we will use throughout this chapter is that of *local* levels. The local level n, $\mathfrak{L}_n$, is built from the products of operators that contain at most n operators of a same party. For instance, in a two-party scenario, $\mathfrak{L}_2$ would be given by

$$\begin{aligned}\mathfrak{L}_2 = \{\mathbb{1}\} \cup \{A_i\}_i \cup \{B_j\}_j \cup \{A_iB_j\}_{i,j} \cup \{A_iA_j\}_{i,j} \cup \{B_iB_j\}_{i,j} \\ \cup \{A_iA_jB_k\}_{i,j,k} \cup \{A_iB_jB_k\}_{i,j,k} \cup \{A_iA_jB_kB_l\}_{i,j,k,l}.\end{aligned}$$

[3] Throughout this thesis, we will indistinctly use the trace notation, $\mathrm{Tr}(\rho\,\cdot)$, and expectation value notation, $\langle\cdot\rangle_\rho$, to denote averages of operators with respect to quantum states.

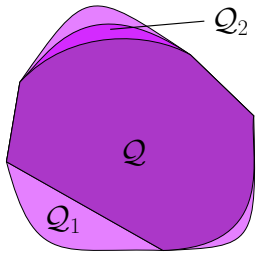

Figure 5.5: Pictorial representation of the first sets in the NPA hierarchy. As shown, $\mathcal{Q}_r \subseteq \mathcal{Q}_s$ for any $r \leq s$. Nevertheless, there might be regions where a finite level already gives the tight characterization of $\mathcal{Q}$, as $\mathcal{Q}_2$ already does in the lower part of the figure.

It turns out that $\mathcal{S}_n \subset \mathfrak{L}_n$, but $\mathfrak{L}_{n-1} \not\subset \mathcal{S}_n$. Because of this, the use of local levels may provide tighter characterizations in certain situations. We will see some of these cases throughout the following sections.

The general question of a distribution admitting a quantum realization in the sense of Eq. (5.9) is undecidable in the general case, as implied by recent results on quantum correlations [Slo19]. In contrast, determining the existence of a positive-semidefinite Γ whose elements are related to each other through a set of linear constraints is a typical instance of semidefinite programming, for which optimal solutions can be obtained efficiently. As such, the NPA hierarchy has been extensively used for many different applications in quantum information theory, ranging from the calculation of maximal quantum violations of Bell inequalities [Weh06] to estimating the amount of randomness achievable in device-independent protocols [PAM+10] or providing insights on the geometrical structure of the quantum set itself [GKW+18].

Hierarchy for local correlations

It is possible to modify the NPA hierarchy in order to bound the set of multipartite classical correlations, instead of the set of quantum correlations. Any local correlations can be realized by performing commuting local measurements on a quantum state. Therefore, if one adds to the conditions i–iv the commutation of all measurement operators of a same party, namely

v) Commutation for different measurements: $[\Pi^p_{o|i}, \Pi^p_{o'|i'}] = 0 \quad \forall\, p, o, o', i, i',$

the associated NPA hierarchy bounds the set of classical correlations from the outside. This observation was used, for instance, in [BCWA17] to provide efficiently computable entanglement witnesses and Bell inequalities for multipartite quantum systems.

A causal interpretation

It must be noted that, aside from nonnegativity and trace-one—which are constraints any quantum state must satisfy—no further properties are required from the state ρ. This means that, from a causal point of view, the NPA hierarchy makes no restrictions on the causal structure underlying the correlations, and thus in general considers the least restrictive, where a multipartite state is shared among all the parties, as in Fig. 5.1b.

However, in general multipartite scenarios multiple sources distribute states to different parties, enforcing a tensor-product factorization of ρ. An example of this can be seen in Figs. 5.1c or 5.1d, where the quantum states involved would satisfy, respectively, $\rho = \rho_{\mathrm{A_1B_1}} \otimes \rho_{\mathrm{B_2C_1}} \otimes \rho_{\mathrm{A_2C_2}}$ and $\rho = \rho_{\mathrm{AB_1}} \otimes \rho_{\mathrm{B_2C}}$. These may give rise to nonlinear constraints between the elements of a specific Γ. While the reasoning behind the NPA hierarchy still applies in this case, the constraints that arise in Γ due to the structure of the quantum state cannot be included in semidefinite programming, leaving the procedure orphaned of an efficient method for numerical computations. In this sense, the main results of this chapter can be understood as NPA-like hierarchies for characterizing the correlations achievable in nontrivial quantum causal structures.

5.3 Scalar extension

The first technique we describe allows for enforcing, in a way compatible with semidefinite programming, relaxations of the conditions that two variables in a causal structure are conditionally independent of each other.

For illustration, consider the line scenario, a generalization of the Bell scenario underlying quantum repeaters where parties are arranged in a line and two consecutive parties are linked together by a latent variable. The simplest of these instances corresponds to an entanglement swapping experiment involving three parties, as represented in Fig. 5.1d. Correlations compatible with line scenarios are such that the marginal distributions resulting after discarding any nonextremal party factorize,

$$\sum_{o_k} P(o_1 \dots o_k \dots o_n | i_1 \dots i_k \dots i_n) = P(o_1 \dots o_{k-1} | i_1 \dots i_{k-1}) P(o_{k+1} \dots o_n | i_{k+1} \dots i_n).$$

where i_k and o_k denote the input and output variable for party k.

These constraints are nonlinear and nonconvex, and thus cannot be imposed in semidefinite programs. However, the first method we develop—termed *scalar*

extension—partially overcomes this limitation by imposing relaxations of factorization conditions that are compatible with semidefinite programming.

5.3.1 Description of the method

The main idea behind scalar extension is complementing the moment matrices involved in the steps of the NPA hierarchy with monomials that represent factorized quantities. More concretely, given a set of products of measurement operators, $\mathcal{S}$, which produces a specific moment matrix Γ, scalar extension consists in enlarging $\mathcal{S}$ with extra operators of the form $S_i\langle S_j\rangle$, $S_i\langle S_j\rangle\langle S_k\rangle\ldots$, where S_i, S_j, $S_k\ldots$ are products of operators not necessarily belonging to the original set $\mathcal{S}$. These additional generating operators give rise to matrix elements that represent factorized quantities, which can then be related via equality constraints—allowed in semidefinite programming—to elements in the original moment matrix that should factorize.

In the same spirit as the original NPA hierarchy, some of the matrix elements in $\tilde{\Gamma}$ can be computed from the given probability distribution, the rest remaining as variables. If the given probability distribution is compatible with the causal scenario, then there exists a variable assignment such that $\tilde{\Gamma}\succeq 0$, which can be found via semidefinite programming.

Since Γ is a principal submatrix of $\tilde{\Gamma}$, the positivity of the latter implies the positivity of the former. If a positive-semidefinite completion of $\tilde{\Gamma}$ cannot be found, it follows that the correlation under scrutiny is not compatible with the proposed causal explanation. Furthermore, if it is also the case that a nonnegative completion for Γ does not exist, per the standard NPA hierarchy not even with having access to a global multipartite quantum state (in the spirit of Fig. 5.1b) the correlation could have been generated.

To illustrate the method let us consider the tripartite line scenario of Fig. 5.1d. The network geometry imposes that all entries in the moment matrix generated by operator strings that only contain operators of the extreme parties A and C factorize. Assume measurement operators for each of the parties A_0, A_1, C_0 and C_1, that square to the identity operator[4], this is, $P_i^2=\mathbb{1}$. We can consider the moment matrix generated by the extended set of operators

[4] This is the standard setting when one considers measurements with only two possible outcomes, where $P_i=\Pi^p_{0|i}-\Pi^p_{1|i}$.

$\{\mathbb{1}, A_0A_1, C_0C_1, \langle A_0A_1\rangle\mathbb{1}\}$:

$$\tilde{\Gamma} = \begin{array}{c} \\ \mathbb{1} \\ (A_0A_1)^\dagger \\ (C_0C_1)^\dagger \\ \langle A_0A_1\rangle^*\mathbb{1} \end{array} \begin{array}{c} \begin{array}{cccc} \mathbb{1} & A_0A_1 & C_0C_1 & \langle A_0A_1\rangle\mathbb{1} \end{array} \\ \begin{pmatrix} 1 & v_1 & v_2 & v_3 \\ & 1 & v_4 & v_5 \\ & & 1 & v_6 \\ & & & v_7 \end{pmatrix} \end{array} \tag{5.15}$$

(the lower triangle has been omitted since $\tilde{\Gamma}$ is Hermitian).

The insertion of the extra column labeled by the operator $\langle A_0A_1\rangle\mathbb{1}$, which represents an operator equal to the identity times the unknown scalar factor $\langle A_0A_1\rangle$, gives rise to a number of equality constraints that relate elements in the top-left 3×3 submatrix (which is the corresponding Γ that one obtains in the application of the standard NPA hierarchy) with elements in the last column.

On the one hand, there are the "trivial" constraints $v_1 = v_3$ and $v_5 = v_7$. These equate different expressions of the variable $\langle A_0A_1\rangle^2$, and thus do not encode any information on the causal structure to be tested. On the other hand, only if it is wanted to contrast the distribution against the tripartite-line scenario, one can impose that $v_4 = v_6^*$. This is a causal constraint, that encodes the factorization $\langle A_0A_1C_0C_1\rangle = \langle A_0A_1\rangle\langle C_0C_1\rangle$.

Additional constraints can be set in $\tilde{\Gamma}$, namely $v_4 = v_1^*v_2$ (which also arises from the tripartite-line causal structure, and thus in principle should only be imposed when testing compatibility with it) and $v_5 = |v_1|^2$. These constraints are however nonlinear and will not be enforced, since it is desired to keep the problem solvable via semidefinite programming. Furthermore, as shown below, they are not necessarily required for providing useful information about correlations in networks.

The independence constraints exploited in scalar extension apply both to classical and quantum networks, and thus, as we show below, it can be employed to study both classical and quantum correlations in networks.

5.3.2 Applications

We illustrate the power of scalar extension with two examples, one regarding compatibility with quantum causal structures, and the other regarding compatibility with classical structures. The latter constitutes a new form of nonlocality activation: we find quantum measurements that cannot produce nonlocal correlations in a standard Bell scenario, but give rise in the tripartite-line scenario to correlations that cannot be reproduced with two independent classical sources, i.e., they can generate *nonbilocal* correlations.

Identifying supra-quantum correlations

Simple scalar extensions allow recovering the results of [BRGP12] in the tripartite line scenario. In that work, nonlinear Bell-like inequalities were obtained to discern in a device-independent manner whether correlations observed in tripartite scenarios are compatible or not with *bilocal hidden variable models* of the form

$$P(abc|xyz) = \int_{\Lambda_1} \mathrm{d}\lambda_1 \int_{\Lambda_2} \mathrm{d}\lambda_2\, P(\lambda_1)P(\lambda_2)P(a|x,\lambda_1)\, P(b|y,\lambda_1,\lambda_2)\, P(c|z,\lambda_2). \tag{5.16}$$

In fact, similar queries can be done on whether a particular distribution can have a *biquantum model*, i.e., if it can be written in the form

$$P(abc|xyz) = \mathrm{Tr}\left(\Pi^{\mathrm{A}}_{a|x}\Pi^{\mathrm{B_1B_2}}_{b|y}\Pi^{\mathrm{C}}_{c|z}\rho_{\mathrm{AB_1}} \otimes \rho_{\mathrm{B_2C}}\right). \tag{5.17}$$

As an example of use of the scalar extension technique, we recover two known results about the entanglement swapping configuration where each party has a binary input and output, $x, y, z, a, b, c \in \{0, 1\}$. Consider the one-parameter family of correlations

$$P_v = vP^{22} + (1-v)P_0, \tag{5.18}$$

where $0 \leq v \leq 1$, $P_0(abc|xyz) = 1/8\ \forall\, a, b, c, x, y, z$ is a noise term, and P^{22} is [BRGP12]

$$P^{22}(abc|xyz) = \frac{1}{8}\left[1 + (-1)^{a+b+c+xy+yz}\right]. \tag{5.19}$$

The distribution P_v does not have a biquantum model for any $v > 1/2$, and a bilocal hidden variable model for any $v > 1/4$. Using the scalar extension construction we are able to reproduce these results. The explicit calculations performed in this section are shown in the computational appendix [PK19].

For comparing against biquantum models, we consider an extension of the NPA level $\mathcal{S}_3$ [NPA08] with a minimal set of operators needed to impose constraints on all elements in the moment matrix that should factorize. This minimal set has the form $\{\langle S_i\rangle\mathbb{1}\}_i$, where S_i is any product of operators of party A of length $2 \leq \ell \leq 5$. We generate the necessary moment matrices $\tilde{\Gamma}$ with Ncpol2sdpa [Wit15], and optimize their smallest eigenvalues λ using Mosek [MOS], SeDuMi [Stu99] and SDPT3 [TTT99]. The specific problem solved is

$$\begin{aligned} &\underset{\vartheta}{\text{maximize}} && \lambda \\ &\text{subject to} && \tilde{\Gamma}(\vartheta) - \lambda\mathbb{1} \succeq 0, \end{aligned} \tag{5.20}$$

where ϑ represents the variables in the unknown entries of $\tilde{\Gamma}$.

For visibilities $v > 1/2$ the optimal smallest eigenvalues found are negative[5]. This renders the family of distributions $P_{v>1/2}$ incompatible with biquantum models. For $v = 1/2$ the distribution is equal to $P_{1/2} = P_Q^{22}$, defined in [BRGP12], which is known to have a quantum realization. To generate $P_{1/2}$, the sources must generate maximally entangled states $\rho_{AB} = \rho_{BC} = |\phi^+\rangle\langle\phi^+|$, the central party performs a Bell state measurement, the extreme parties perform the measurements $A_i = C_i = [X + (-1)^i Z]/\sqrt{2}$.

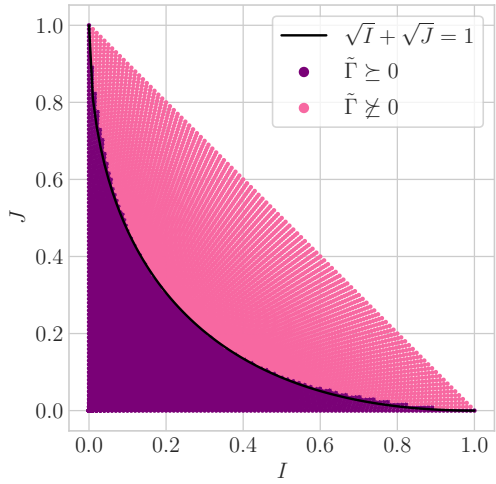

Figure 5.6: Results of scalar extension in the $I-J$ plane. The black line denotes the analytical bilocality inequality, which is known to be tight [BRGP12]. The light dots represent probability distributions that are identified, through scalar extension, not to be compatible with a bilocal model in the form of Eq. (5.16). The numerical results are obtained with the generating set $\mathcal{S}_3$ extended with the only monomial $\langle \Pi^A_{0|0}\Pi^A_{0|1}\rangle\mathbb{1}$. In a three-party, two-input, two-output scenario, $I = \frac{1}{4}\sum_{x,z=0,1}\langle A_x B_0 C_z\rangle$, $J = \frac{1}{4}\sum_{x,z=0,1}(-1)^{x+z}\langle A_x B_1 C_z\rangle$.

In order to compare P_v against bilocal models, we impose the additional constraint that operators representing different measurements of a same party commute (see Section 5.2.4 and Appendix D.1). Using the corresponding generating set $\mathcal{S}_3$, in this case an extension with the single operator $\langle \Pi^A_{0|0}\Pi^A_{0|1}\rangle\mathbb{1}$ suffices to discard the existence of a bilocal model of the correlations whenever the vis-

[5] The stopping criterion for the solvers is set to be that the absolute gap between the primal and dual objectives is smaller than 10^{-12}. The optimal smallest eigenvalue is considered negative whenever both the primal and dual objectives are negative. The output of an optimization problem is accepted as valid whenever the results given by all solvers coincide up to a relative error of 10^{-6}.

ibility satisfies $v > 1/4$. This is the same result than that shown in [BRGP12] for noisy versions of the quantum distribution $P_Q^{22}(V)$, since $P_{1/4} = P_Q^{22}(V{=}1/2)$. When $v \leq 1/4$, P_v is known to have a bilocal model.

The construction depicted above—the generating set $\mathcal{S}_3$ complemented with $\langle \Pi^{\mathrm{A}}_{0|0} \Pi^{\mathrm{A}}_{0|1} \rangle \mathbb{1}$—can not only tightly identify the bounds of P_v. In fact, it accurately approximates the results of the analytic *bilocality inequality* of [BRGP12] in the so-called $I{-}J$ plane. In Fig. 5.6 we show explicitly how the bilocality inequality and scalar extension compare.

Activation of measurement nonlocality

In the following we use scalar extension to demonstrate how the nonlocal power of measurement devices can be activated in a network structure: measurements that do not lead to nonlocal correlations in the standard Bell scenario, do it when arranged in a network. Similar effects are in fact known for quantum states [CASA11].

Consider a scenario in which one has access to three measuring devices. One implements a single four-output measurement. The remaining two devices each implement two measurements of binary outputs with limited detection efficiency: for all measurements there is a probability $1 - \eta$ that a third outcome is observed, indicating the fact that the measurement has failed. We denote by η_1 and η_2 these efficiencies. What are their values so that nonlocal correlations can be certified with these three devices acting on a quantum state? Since one of the devices implements only one measurement, the only possibility is to run a standard bipartite Bell test with the other two, possibly conditioned on one of the measurement outputs of the first. Critical values for the detection efficiencies such that no nonlocal correlations can be observed with two-output measurements can be obtained from [MP03]: in case one device is perfect, say $\eta_1 = 1$, it is impossible to observe nonlocal correlations whenever $\eta_2 < 1/2$; if both devices have the same efficiency, a local model for the correlations always exists if $\eta_1 = \eta_2 \leq 2/3$.

We now arrange these devices in a tripartite-line scenario and make use of scalar extension to determine detection efficiencies for which nonclassical correlations can be observed in the network. In particular, we focus on the case where the sources send partially-entangled states $|\theta_{ij}\rangle = \cos\theta_{ij} |00\rangle + \sin\theta_{ij} |11\rangle$, the measurement device of party B makes a perfect, 4-outcome Bell state measurement in the standard basis $\{|\phi^+\rangle, |\phi^-\rangle, |\psi^+\rangle, |\psi^-\rangle\}$, and the measurements

Assumptions		η_{min}	θ_{min}	$(\eta, \theta)_{ex}$
$\eta_{\mathrm{A}} = 1,$	$\theta_{\mathrm{AB}} = \pi/4$	$< 10^{-5}$	$< 10^{-4}$	(0.0001, 0.1250)
$\eta_{\mathrm{A}} = 1,$	$\theta_{\mathrm{AB}} = \theta_{\mathrm{BC}}$	$< 10^{-5}$	$< 10^{-4}$	(0.0444, 0.1000)
$\eta_{\mathrm{A}} = \eta_{\mathrm{C}},$	$\theta_{\mathrm{AB}} = \pi/4$	0.6085	0.0010	(0.6389, 0.6545)
$\eta_{\mathrm{A}} = \eta_{\mathrm{C}},$	$\theta_{\mathrm{AB}} = \theta_{\mathrm{BC}}$	0.5291	0.0070	(0.5626, 0.1751)

Table 5.1: Upper bounds to the smallest detection efficiency and entanglement needed to generate nonbilocal correlations in the tripartite-line scenario. The values on the second and third columns $(\eta_{min}, \theta_{min})$ are not related, so in general one needs $\theta > \theta_{min}$ to be able to discard bilocal models with $\eta = \eta_{min}$ and vice-versa. The last column shows an example of small combined values of the parameters for which nonbilocal correlations can be generated.

performed by parties A and C have the form [BSS14]

$$A_0 = \cos\alpha_0 Z - \sin\alpha_0 X, \qquad C_0 = \cos\alpha_0 Z + \sin\alpha_0 X,$$
$$A_1 = \cos\alpha_1 Z + \sin\alpha_1 X, \qquad C_1 = \cos\alpha_1 Z - \sin\alpha_1 X.$$

With probability $1-\eta_i$, party $i \in \{\mathrm{A}, \mathrm{C}\}$ produces the third, "measurement failed" result. Therefore, the measurements can be modeled by the three-outcome POVMs

$$\left\{ \frac{\eta_i}{2}\left(\mathbb{1} + i_j\right), \frac{\eta_i}{2}\left(\mathbb{1} + i_j\right), (1-\eta_i)\mathbb{1} \right\}_{i,j} = \left\{ \eta_i\, \Pi^i_{0|j}, \eta_i\, \Pi^i_{1|j}, (1-\eta_i)\mathbb{1} \right\}_{i,j}.$$

The task is to find the smallest values for the efficiencies of the extreme parties A and C that are able to produce correlations which cannot be explained by a bilocal hidden variable model of the type (5.16). The results are shown in Table 5.1, where we also vary the entanglement of the prepared pure states. All results have been achieved with the generating operator set corresponding to the (commuting) NPA level $\mathcal{S}_3$, extended with the set of four operators $\{\langle \Pi^{\mathrm{A}}_{a_1|0} \Pi^{\mathrm{A}}_{a_2|1} \rangle \mathbb{1} : a_1, a_2 \in \{0,1\}\}$, where $\Pi^{\mathrm{A}}_{o|i}$ is the projector on the outcome o of the i^{th} measurement of party A. This set is sufficient for imposing nontrivial constraints on all the monomials that should factorize in the original Γ.

When party A's measurement device is perfect (the case in the two first rows of Table 5.1), the observed three-party correlations fail to have bilocal realizations even in cases when party C detects as few as 0.001% of all the particles that receives. This value is well below 50%, the critical value needed for certifying standard nonlocality. In fact, we believe that nonclassical correlations can be observed for all $\eta_{\mathrm{C}} > 0$, and that the obtained critical value is a consequence of numerical issues when the detection efficiency is very low.

We also observe that very low entanglement is needed to create nonbilocal correlations. This last finding is similar to the result known for bipartite Bell

scenarios that any partially-entangled state is capable of generating nonlocal correlations [Gis91].

It must be noted that, albeit nonbilocal, the correlations generated can still be generated through a single three-partite local hidden variable. This means that the nonbilocality does not arise from the correlations being nonlocal in the conventional sense.

In the case of two inefficient devices, if the state prepared by the source between parties A and B is fixed to be maximally entangled, nonbilocal correlations can be established for any detection efficiency higher than $\eta_{min} = 0.6085$. Nonbilocal correlations at η_{min} are generated when $\theta_{\text{BC}} = \pi/4$—so the source between parties B and C also distributes maximally entangled states—and the measurements performed are $A_x = [Z - (-1)^x X]/\sqrt{2}$, $C_z = [Z + (-1)^z X]/\sqrt{2}$. For higher detection efficiencies, nonbilocal correlations can be generated with decreasing amounts of entanglement and measurements whose aperture (the relative angle $|\alpha_1 - \alpha_0|$) decreases, suggesting that this is the optimal configuration for detecting nonbilocality in a loss-resistant manner.

Varying also the entanglement in θ_{AB}, correlations without bilocal models can be certified for efficiencies above $\eta_{min} = 0.5291$, again well below $2/3$, the value necessary for certifying standard Bell nonlocality.

5.4 Quantum inflation

As discussed, scalar extension is a technique only applicable when studying causal structures with causally independent nodes. Importantly, there exist nontrivial structures without causally independent variables, the simplest example being the triangle scenario of Fig. 5.1c. In the following we describe a method to study the quantum correlations achievable in any causal scenario, that can be seen as a quantum analogue of the inflation technique which does not present issues with information broadcasting.

5.4.1 Description by an example

We first explain how quantum inflation works by means of an example and provide more details about the general construction in Section 5.4.2.

Consider the quantum causal network depicted in Fig. 5.7a, whereby three random variables a, b, c are generated by conducting bipartite measurements over the ends of three bipartite quantum states ρ_{AB}, ρ_{BC}, ρ_{AC}. We are handed a distribution $P_{\text{obs}}(a, b, c)$ of observed variables and asked if it is compatible with this model.

To answer this question, suppose that there existed indeed bipartite states $\rho_{\text{AB}}, \rho_{\text{BC}}, \rho_{\text{AC}}$ of systems $A'B'$, $B''C'$, $A''C''$, and commuting measurement operators E_a, F_b, G_c, acting on systems $A'A''$, $B'B''$, $C'C''$, respectively, which were able to reproduce the distribution $P_{\text{obs}}(a,b,c)$, i.e.,

$$\begin{aligned} P_{\text{obs}}(a,b,c) =& \text{Tr}\left(E_a F_b G_c \rho_{\text{AB}} \otimes \rho_{\text{BC}} \otimes \rho_{\text{AC}}\right) \\ =& \langle E_a F_b G_c \rangle_{\rho_{\text{AB}} \otimes \rho_{\text{BC}} \otimes \rho_{\text{AC}}}. \end{aligned}$$

Now imagine how the scenario would change if n independent copies $\rho^i_{\text{AB}}, \rho^i_{\text{BC}}, \rho^i_{\text{AC}}$, $i = 1, \ldots, n$ of each of the original states were distributed instead, as in Fig. 5.7b. Call ρ the overall quantum state before any measurement is carried out. For any $i, j = 1, \ldots, n$ we can, in principle, implement the measurement $\{E_a\}_a$ on the i^{th} copy of ρ_{AC} and the j^{th} copy of ρ_{AB}. Let us denote by $\{E^{i,j}_a\}_a$ the corresponding measurement operators, and similarly, call $\{F^{i,j}_b\}_b$, $(\{G^{i,j}_c\}_c)$ the measurement $\{F_b\}_b$ $(\{G_c\}_c)$ over the states $\rho^i_{\text{AB}}, \rho^j_{\text{BC}}$ $(\rho^i_{\text{BC}}, \rho^j_{\text{AC}})$.

The newly defined operators and their expectation values under state ρ satisfy nontrivial relations with respect to the original operators that reproduce P_{obs}. For example, for $H = E, F, G$ and $i \neq k, j \neq l$ the operators $H^{i,j}_a$ and $H^{k,l}_b$ act on different Hilbert spaces, and hence they commute, i.e., $[H^{i,j}_a, H^{k,l}_b] = 0$. Similarly, expressions such as $\langle E^{1,1}_a E^{1,2}_{a'} F^{2,2}_b \rangle_\rho$ and $\langle E^{1,2}_a E^{1,1}_{a'} F^{1,2}_b \rangle_\rho$ can be shown identical, since one can arrive at the second one from the first one just by exchanging ρ^1_{AB} with ρ^2_{AB}, which are copies of the same bipartite quantum state. More generally, for any function $Q(\{E^{i,j}_a, F^{k,l}_b, G^{m,n}_c\})$ of the measurement operators and any three permutations π, π', π'' of the indices $1, \ldots, n$, one should have

$$\left\langle Q(\{E^{i,j}_a, F^{k,l}_b, G^{m,n}_c\}) \right\rangle_\rho = \left\langle Q(\{E^{\pi(i),\pi'(j)}_a, F^{\pi'(k),\pi''(l)}_b, G^{\pi''(m),\pi(n)}_c\}) \right\rangle_\rho. \tag{5.21}$$

Finally, note that, if we conducted the measurements $\{E^{i,i}, F^{i,i}, G^{i,i}\}_{i=1}^n$ at the same time (we can do this, as they all commute with each other), then we would be measuring the over exact copies of the original causal structure. Therefore, the measurement outcomes $a^1, \ldots, a^n, b^1, \ldots, b^n, c^1, \ldots, c^n$ would be distributed according to

$$\left\langle \prod_{i=1}^n E^{i,i}_{a^i} F^{i,i}_{b^i} G^{i,i}_{c^i} \right\rangle_\rho = \prod_{i=1}^n P_{\text{obs}}(a^i, b^i, c^i). \tag{5.22}$$

If the original distribution $P_{\text{obs}}(a,b,c)$ is compatible with the network in Fig. 5.7a, then there should exist a Hilbert space $\mathcal{H}$, a state $\rho \in \mathcal{H}$ and operators $\{E^{i,j}_a\}_{i,j,a}$, $\{F^{k,l}_b\}_{k,l,b}$, $\{G^{m,n}_c\}_{m,n,c}$ satisfying the above relations. If such is the case, we say that $P_{\text{obs}}(a,b,c)$ admits an n^{th} *order quantum inflation* in

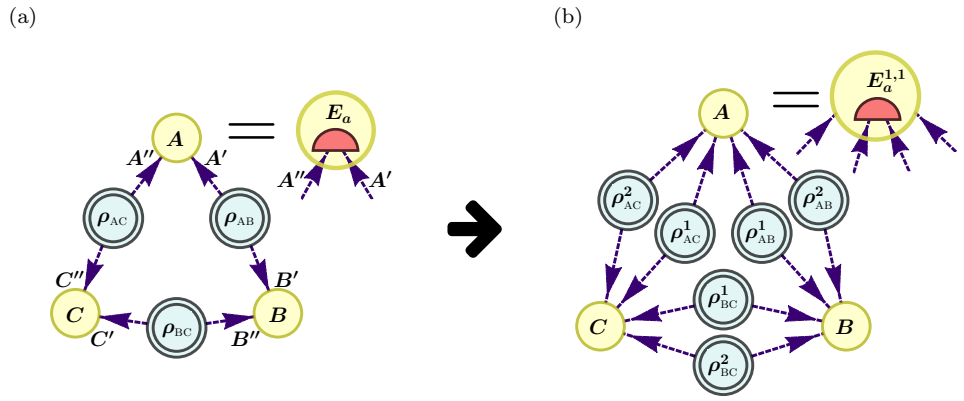

Figure 5.7: Illustration of the quantum inflation technique in the triangle scenario. (a) In the original scenario, by probing systems A', A'' with the quantum measurement $\{E_a\}$, depicted as a red semi-circle, a value a for the random variable A is generated. The values b, c for the random variables B and C are produced similarly. (b) In quantum inflation, we distribute n (in the case shown, $n = 2$) independent copies of the same states to the parties, which now use the original measurement operators on different pairs of copies of the states they receive. For instance, the measurement operators $\{E_a^{1,1}\}_a$ act on the states corresponding to copies ρ^1_{AB} and ρ^1_{AC}, and the measurements with other superindices are defined in an analogous way. If a probability distribution has a quantum realization in the original scenario, there should exist a global state and measurement operators in the inflated scenario that would reproduce the original probability distribution when considering suitable combinations of measurements, as given by Eq. (5.22). Additionally, the measurement operators should be such that any function of them should be invariant under permutation of the indices denoting the copies of the systems, as described in Eq. (5.21). The existence of such a global state and measurement operators is a typical instance of the problems solved with the NPA hierarchy. Therefore, quantum inflation imposes a hierarchy of constraints, each one more restrictive than the previous, that correlations generated in quantum causal networks must satisfy.

the causal scenario. By increasing the index of n, we arrive at a hierarchy of conditions, each of which must be satisfied by any compatible distribution $P_{\text{obs}}(a, b, c)$.

The problem above is precisely of the type that the NPA hierarchy can address. If for some n we were able to certify, via the use of the NPA hierarchy or general noncommutative polynomial optimization (NPO) theory [PNA10]—of which the NPA hierarchy is an example—, that $P_{\text{obs}}(a, b, c)$ does not admit an n^{th} order quantum inflation, then we would have proven that $P_{\text{obs}}(a, b, c)$ does not admit a realization in the quantum network of Fig. 5.7a.

5.4.2 Detailed explanation

The method just described can be easily adapted to bound the statistics of any network in which the observed variables are defined by measurements on the quantum latent variables. The essential premise of quantum inflation is to ask what would happen if multiple copies of the original (unspecified) quantum states were simultaneously available to each party. In this gedankenexperiment the parties use copies of their original measurement apparatus, allowing a party to perform n simultaneous measurements on the n copies of the original quantum states now available to her. There are different ways in which a party can align their measurements to act on the states now available, thus must explicitly specify upon which unique set of quantum states a given measurement operator acts upon nontrivially. Let us therefore denote measurement operators by

$$\hat{O}^{\boldsymbol{s}|k}_{i|m} \equiv \hat{O}\big(\text{Party}{=}k,\ \text{States}{=}\boldsymbol{s},\ \text{Setting}{=}m,\ \text{Outcome}{=}i\big)$$

where the four indices specify

1. k, the *party* performing the measurement,
2. $\boldsymbol{s}$, the Hilbert spaces the given operator acts on,
3. i, the measurement *setting* being used,
4. m, the *outcome* associated with the operator.

For example, using an $n = 2$ quantum inflation of the triangle scenario one would find that, for instance, $\boldsymbol{s}$ for party $k = \text{A}$ may be sampled from precisely four possibilities, each value being a different tuple:

$$\boldsymbol{s} \in \Big\{ \{\rho^1_{\text{AB}}, \rho^1_{\text{AC}}\}, \{\rho^1_{\text{AB}}, \rho^2_{\text{AC}}\}, \{\rho^2_{\text{AB}}, \rho^1_{\text{AC}}\}, \{\rho^2_{\text{AB}}, \rho^2_{\text{AC}}\} \Big\}.$$

These operators will be regarded as the noncommuting variables of an NPO problem [PNA10] where there will be four families of constraints. On one hand, there will be polynomial constraints coming both from projection axioms and

from commutation rules. On the other, statistical constraints will follow from symmetry under permutation of the state indices and consistency with the observed probabilities.

To begin with, for fixed $\boldsymbol{s}, k, m$, the noncommuting variables $\{\hat{O}^{\boldsymbol{s}|k}_{i|m}\}_i$ must correspond to a complete set of measurement operators. Since we do not restrict the dimensionality of the Hilbert space where they act upon, we can take them to be a complete set of projectors without loss of generality. That is, they must obey the relations

Projection axioms

$$\hat{O}^{\boldsymbol{s}|k}_{i|m} = (\hat{O}^{\boldsymbol{s}|k}_{i|m})^\dagger, \qquad \forall\, \boldsymbol{s}, k, i, m, \tag{5.23a}$$

$$\hat{O}^{\boldsymbol{s}|k}_{i|m} \cdot \hat{O}^{\boldsymbol{s}|k}_{i'|m} = \delta_{ii'} \hat{O}^{\boldsymbol{s}|k}_{i|m}, \qquad \forall\, \boldsymbol{s}, k, i, i', m, \tag{5.23b}$$

$$\sum_i \hat{O}^{\boldsymbol{s}|k}_{i|m} = \mathbb{1}, \qquad \forall\, \boldsymbol{s}, k, m. \tag{5.23c}$$

These relations imply, in turn, that each of the noncommuting variables is a bounded operator. Hence, by [PNA10], the hierarchy of SDP programs provided by NPO is complete: if the said distribution does not admit an n^{th} order inflation, then one of the NPO SDP relaxations will detect its infeasibility.

Commutation rules encode the property that operators acting on different Hilbert spaces commute. These include the standard commutations for measurement operators of different parties [Eq. (5.24a)], which are used in standard NPO theory for studying quantum correlations. In the case of quantum inflation, it additionally applies to those measurement operators of a same party that act on disjoint subsets of the Hilbert space associated to the party [Eq. (5.24b)]. Formally,

Commutation rules

$$\hat{O}^{\boldsymbol{s}_1|k_1}_{i_1|m_1} \cdot \hat{O}^{\boldsymbol{s}_2|k_2}_{i_2|m_2} = \hat{O}^{\boldsymbol{s}_2|k_2}_{i_2|m_2} \cdot \hat{O}^{\boldsymbol{s}_1|k_1}_{i_1|m_1} \qquad \text{if} \quad k_1 \neq k_2, \tag{5.24a}$$

$$\text{or if} \quad \boldsymbol{s}_1 \cap \boldsymbol{s}_2 = \emptyset. \tag{5.24b}$$

As was also done in scalar extension (see Section 5.3), these commutation rules can be modified by inserting the constraints defined in Appendix D.1 so as to construct an alternative SDP for constraining the correlations of *classical* causal structures.

The ingredient that relates the inflated network structure to the original network is that all moments are invariant under any permutation of the source indices. Call ρ the overall quantum state of the inflated network (since we do

not cap the Hilbert space dimension, we can assume that all state preparations in the original network are pure, and so is ρ). Then these constraints are expressed as

Copy-index invariance (automorphic symmetry)

$$\left\langle \hat{O}^{\boldsymbol{s}_1|k_1}_{i_1|m_1}\cdot \hat{O}^{\boldsymbol{s}_2|k_2}_{i_2|m_2}\cdot \ldots \cdot \hat{O}^{\boldsymbol{s}_n|k_n}_{i_n|m_n} \right\rangle_\rho = \left\langle \hat{O}^{\pi(\boldsymbol{s}_1)|k_1}_{i_1|m_1}\cdot \hat{O}^{\pi(\boldsymbol{s}_2)|k_2}_{i_2|m_2}\cdot \ldots \cdot \hat{O}^{\pi(\boldsymbol{s}_n)|k_n}_{i_n|m_n} \right\rangle_\rho . \tag{5.25}$$

An example of such statistical constraints imposed in the triangle scenario was given in Eq. (5.21). Another example for that scenario, for an inflation level $n=3$, can be the following:

$$\begin{aligned}
&\left\langle E_0^{\{\rho^1_{\mathrm{AB}},\rho^1_{\mathrm{AC}}\}} E_1^{\{\rho^2_{\mathrm{AB}},\rho^2_{\mathrm{AC}}\}} F_0^{\{\rho^3_{\mathrm{BC}},\rho^1_{\mathrm{AB}}\}} G_0^{\{\rho^1_{\mathrm{AC}},\rho^2_{\mathrm{BC}}\}} \right\rangle_\rho \\
&\qquad\qquad \text{apply } \rho^1_{\mathrm{AB}}\leftrightarrow\rho^2_{\mathrm{AB}} \\
&= \left\langle E_0^{\{\rho^2_{\mathrm{AB}},\rho^1_{\mathrm{AC}}\}} E_1^{\{\rho^1_{\mathrm{AB}},\rho^2_{\mathrm{AC}}\}} F_0^{\{\rho^3_{\mathrm{BC}},\rho^2_{\mathrm{AB}}\}} G_0^{\{\rho^1_{\mathrm{AC}},\rho^2_{\mathrm{BC}}\}} \right\rangle_\rho \\
&\qquad\qquad \text{apply } \rho^1_{\mathrm{AB}}\leftrightarrow\rho^3_{\mathrm{AB}} \\
&= \left\langle E_0^{\{\rho^2_{\mathrm{AB}},\rho^1_{\mathrm{AC}}\}} E_1^{\{\rho^3_{\mathrm{AB}},\rho^2_{\mathrm{AC}}\}} F_0^{\{\rho^3_{\mathrm{BC}},\rho^2_{\mathrm{AB}}\}} G_0^{\{\rho^1_{\mathrm{AC}},\rho^2_{\mathrm{BC}}\}} \right\rangle_\rho \\
&\qquad\qquad \text{apply } \rho^1_{\mathrm{BC}}\leftrightarrow\rho^3_{\mathrm{BC}} \\
&= \left\langle E_0^{\{\rho^2_{\mathrm{AB}},\rho^1_{\mathrm{AC}}\}} E_1^{\{\rho^3_{\mathrm{AB}},\rho^2_{\mathrm{AC}}\}} F_0^{\{\rho^1_{\mathrm{BC}},\rho^2_{\mathrm{AB}}\}} G_0^{\{\rho^1_{\mathrm{AC}},\rho^2_{\mathrm{BC}}\}} \right\rangle_\rho .
\end{aligned}$$

Up to here, all constraints bound the set of feasible solutions to comply with constraints imposed by quantum mechanics and the causal structure. As such, they can also be imposed in problems aimed to optimize quantities—for instance, Bell-like inequalities, or guessing probabilities in a cryptographic scenario—under the set of correlations compatible with a quantum causal structure.

Finally, as described by Eq. (5.22) in the example of the triangle scenario, in the case of compatibility problems certain product averages can be related to products of the probabilities of the distribution P_{obs} that one wishes to test. This is due to the fact that, when restricting to copies of variables that reproduce (multiple copies of) the original scenario, the marginal distribution must coincide with (products of) P_{obs}. This is,

Moment identification

$$\left\langle \prod_{\substack{c=1\ldots|\mathcal{S}|\\ k\in\{\mathrm{A},\mathrm{B},\ldots\}}} \hat{O}^{\boldsymbol{s}_c|k}_{i_{k_c}|m_{k_c}} \right\rangle_\rho = \prod_{c=1}^{|\mathcal{S}|} P_{\mathrm{obs}}(i_{\mathrm{A},c} i_{\mathrm{B},c}\ldots|m_{\mathrm{A},c} m_{\mathrm{B},c}\ldots), \tag{5.26}$$

where $\boldsymbol{s} \cap \boldsymbol{s}' = \emptyset$ for any $\boldsymbol{s}, \boldsymbol{s}' \in \mathcal{S}$, $\boldsymbol{s} \neq \boldsymbol{s}'$. In fact, further constraints of this nature can be imposed when not all factors in the moments reproduce copies of the original scenario. We elaborate on this type of constraints in Appendix D.2.

Arbitrary classical causal scenarios

The technique presented above is a systematic method to characterize the correlations achievable in arbitrary two-layer causal structures where no node has both parents and children. Now we extend those ideas to characterize arbitrary causal structures where classical variables, both visible and latent, can have both parents and children. The case of arbitrary quantum structures is given later on.

For the case of a visible node having both parents and children, we follow a procedure called *interruption* (also called *unpacking* in [NW17]). Graphically, interruption modifies a graph as follows: for every observed node A_i which is neither parentless nor childless, introduce a new variable $A_i^{\#}$ and replace all edges formerly originating from A_i by edges originating from $A_i^{\#}$. In the interruption graph, A_i becomes a childless node and $A_i^{\#}$ is a parentless node. Then, the distribution defined on the interruption, P_{int}, should be such that

$$P_{\text{obs}}(A_i{=}a_i, \dots) = P_{\text{int}}(A_i{=}a_i, \dots | A_i^{\#}{=}a_i). \tag{5.27}$$

The linear nature of these constraints allows one to impose them easily in the semidefinite programs associated to quantum inflation. Some graphical examples of interruptions are shown in Fig. 5.8.

Conceptually, interruption has extensive precedent in literature regarding classical causal inference. It is closely related to the Single-World Intervention Graphs (SWIGS) pioneered by [RR13], as well as the e-separation technique introduced by [Eva12]. Interruption previously has been used to show Tsirelson inequalities constraining the set of quantum correlations compatible with the Bell scenario can be ported to quantum constraints pertaining to the instrumental scenario [vHBP$^+$18] (see also [HLP14, Thm. 25]).

On the other hand, when a causal structure has classical latent variables with parents, the exogenization procedure by Evans [Eva18] eliminates this parental relation. The procedure consists in replacing all arrows from any type of variable to a latent variable with arrows from the former variable to the children of the latter. This operation is repeated until no latent variables have parents. Fig. 5.9 contains an example of the application of this technique.

Proceeding in this fashion, combining interruption and Evans' exogenization, any classical causal structure can be converted into a two-layer structure, which can be analyzed using inflations as described above.

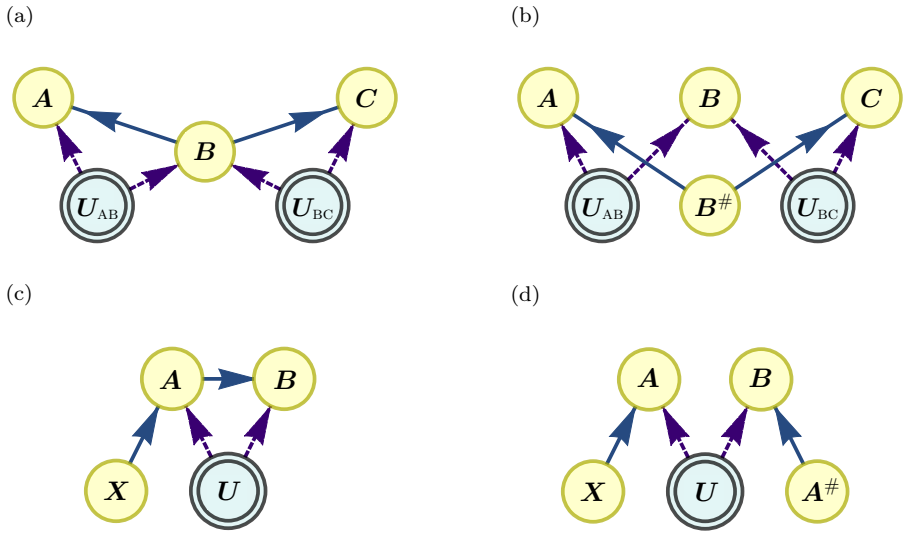

Figure 5.8: Examples of interruption of generic causal structures. (a) The confounding causal structure, and (b) its interruption. Note that in (a) node B has parents and children. The constraint that must be imposed on the interruption is $P_{\text{obs}}(A{=}a, B{=}b, C{=}c) = P_{\text{int}}(A{=}a, B{=}b, C{=}c|B^{\#}{=}b)$. (c) The instrumental causal structure, and (d) its interruption. The constraint to impose in order for the interruption to behave as the original scenario is $P_{\text{obs}}(A{=}a, B{=}b|X{=}x) = P_{\text{int}}(A{=}a, B{=}b|X{=}x, A^{\#}{=}a)$.

Arbitrary quantum causal scenarios

Unfortunately, when applied to quantum latent variables, exogenization results in a new quantum structure that, in general, does not generate the same distributions of observed events as its predecessor. This is made explicit in Fig. 5.9.

In the causal structure of Fig. 5.9a the classical variable S serves as a setting, which adjusts the state ρ_{BC} before it is sent to B and C. Thus, it is possible for $P(A, B|X, Y, S{=}0)$ to maximally violate a Bell inequality for A and B, and at the same time $P(A, C|X, Z, S{=}1)$ to maximally violate a Bell inequality for A and C. No quantum state prepared independently of S can do so, due to the monogamy of quantum correlations [CKW00]. Evans' exogenization procedure leads to Fig. 5.9b as the two-layer equivalent structure, where the node S has been disconnected from the state ρ. Given that in this structure S has no effect on ρ, the monogamy of quantum correlations prohibits to reproduce the maximal violation of Bell inequalities between A and B and between A and C.

In order to properly exogenize quantum latent variables, we treat observable variables with unobserved children as random variables indicating the classical control of a quantum channel. Thus, in Fig. 5.9a we treat the root variable S

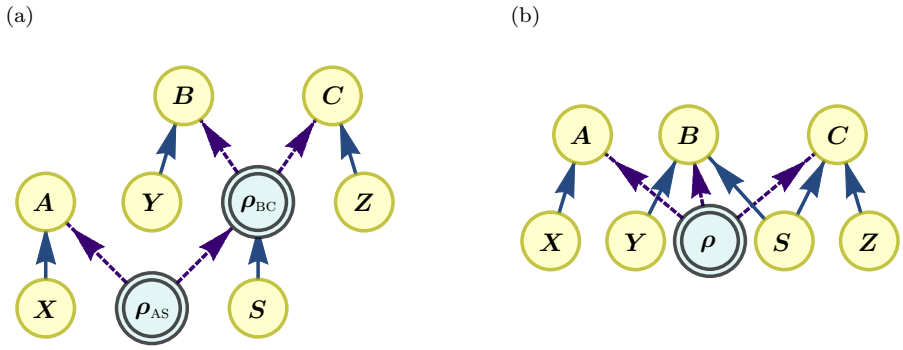

Figure 5.9: The structure in (a) is a causal structure with ρ_{BC} being a nonexogenous unobserved quantum node. Exogenization via Evans' procedure leads to the causal network in (b). While these two graphs would be equivalent if the latent nodes were classical, they are demonstrably inequivalent when the latent nodes represent quantum states being distributed.

as the classical control for a quantum channel acting on the subsystems B and C. That is, one understands $\rho_{BC} = \hat{S}_s \rho_{AS} \hat{S}_s^\dagger$, where $\hat{S}$ is a unitary operator that commutes with any operator acting solely on A's subsystem. As such, the joint distribution of the values of the visible variables A, B, C conditioned on the root visible nodes can be understood as generated by

$$\begin{aligned} P(A{=}a, B{=}b, C{=}c | X{=}x, Y{=}y, Z{=}z, S{=}s) &= \left\langle \hat{S}_s^\dagger \hat{A}_x^a \hat{B}_y^b \hat{C}_z^c \hat{S}_s \right\rangle_{\rho_{AS}} \\ &= \left\langle \hat{A}_x^a \hat{S}_s^\dagger \hat{B}_y^b \hat{C}_z^c \hat{S}_s \right\rangle_{\rho_{AS}} . \end{aligned} \tag{5.28}$$

This interpretation can be made without loss of generality, since the subspace S of the complete Hilbert space can be understood as containing the subspaces corresponding to B and C. Furthermore, the constraints of the form (5.28) can still be implemented in the framework of NPO theory.

Interruption, classical exogenization and quantum exogenization cover all possible nontrivial causal influences in arbitrary quantum causal structures. Due to this, quantum inflation is a technique of full applicability that can characterize the quantum correlations achievable in any causal scenario.

5.4.3 Applications

As in the case of the scalar extension introduced in Section 5.3, we now illustrate the power of quantum inflation with several examples.

It must be noted that, in practice, the implementation of quantum inflation comprises two different hierarchies: the one of inflations, and for each inflation,

the NPO hierarchy used to determine whether a distribution admits such an inflation. Under a different point of view, for addressing a specific problem one can choose to either enforce stricter causality constraints by increasing the level in the inflation hierarchy, or enforce stricter quantum constraints by increasing the level in the NPO hierarchy.

Known incompatible correlations

We begin by showing that quantum inflation recovers known results for various simple quantum causal structures.

The first result is that already discussed in Section 5.3.2 of the P^{22} distribution in the tripartite-line scenario. Recall, P^{22} was defined as [note that this is the same as Eq. (5.19)]

$$P^{22}(abc|xyz) = \frac{1}{8}\left[1 + (-1)^{a+b+c+xy+yz}\right].$$

The distribution P^{22} does not admit a second-order quantum inflation, and this incompatibility is identified already at the second level [recall Eq. (5.14)] of the corresponding NPO hierarchy. This is a proof that P^{22} cannot be realized in the tripartite-line scenario when the latent variables are taken to be quantum. Moreover, this same relaxation (second-order inflation, along with NPO level 2) allows one to identify that, when mixing P^{22} with white noise as in Eq. (5.18), the resulting distribution P^{22}_v is not compatible with the quantum tripartite-line causal structure whenever the visibility is above $v = 1/2$. Despite the fact that the asymptotic convergence of quantum inflation is still to be proven, this example shows that the method is capable of giving tight results by identifying complete families of incompatible distributions.

The most important feature of quantum inflation is that it can be applied to study correlations in any causal structure. In fact, in the same way that we have seen it is capable of identifying distributions not compatible with the tripartite-line scenario, we applied it to detect distributions that are not compatible with the triangle scenario of Fig. 5.1c[6]. The simplest example is the fact that it is not possible to achieve a perfect three-partite correlation in a triangle structure. This impossibility was first proven in [SA15], and recovered in [WSF19, Ex. 1] using a non-fanout inflation. The three-way perfect correlation distribution, also known as GHZ distribution, is

$$P_{\mathrm{GHZ}}(a,b,c) = \begin{cases} \frac{1}{2} & \text{if } a=b=c \\ 0 & \text{otherwise} \end{cases}. \tag{5.29}$$

[6] Recall that scalar extension cannot be applied to this scenario, since no visible variable is independent of any other.

Recall from Section 5.2.2 that non-fanout inflations imposed only constraints related to the causal structure and not to the nature of the latent variables. Therefore, it is not possible to reproduce P_{GHZ} in the triangle scenario even when the latent variables are quantum. This result can be also obtained using quantum inflation. Indeed, P_{GHZ} does not admit a second-order quantum inflation, and this fact is detected at the first level of the associated NPO hierarchy, which requires a moment matrix Γ of size 13×13.

Unknown incompatible distributions

Not only quantum inflation recovers known results from various different quantum networks, but it can also offer new insights on distributions incompatible with a given causal scenario. An example of this is the W distribution, defined as

$$P_{\text{W}}(a,b,c) = \begin{cases} \frac{1}{3} & \text{if } a+b+c=1 \\ 0 & \text{otherwise} \end{cases} \tag{5.30}$$

This distribution describes the coordinated action of one party outputting 1 while the remaining two output 0. Until now the question of whether the W distribution could have a quantum realization in the triangle scenario remained open. However, it can be shown that P_{W} does not admit a second-order quantum inflation, and therefore it cannot be realized in the triangle scenario with quantum latent variables. This incompatibility is identified at the second level of the associated NPO hierarchy.

In the same way, the Mermin-GHZ distribution, which can be realized by performing appropriate measurements in a tripartite GHZ state and has the form

$$P_{\text{Mermin}}(a,b,c|x,y,z) = \begin{cases} 1/8 & \text{if } x+y+z=0 \mod 2 \\ \left[1+(-1)^{a+b+c}\right]/8 & \text{if } x+y+z=1 \\ \left[1-(-1)^{a+b+c}\right]/8 & \text{if } x+y+z=3 \end{cases}, \tag{5.31}$$

can be proven not to have a realization in the quantum triangle scenario with inputs. In fact, P_{Mermin} does not admit a second-order quantum inflation. This incompatibility is identified at the first local level of the associated NPO hierarchy, this is, solving the SDP associated to the coresponding monomial set $\mathfrak{L}_1$.

Importantly, the fact that quantum inflation is a general technique allows one to compute noise robustnesses. This is, it allows one to find the minimum visibilities, in the sense of Eq. (5.18), that distributions require to have in order

Distribution	$v_{\max}$	Inflation order	NPO level
GHZ	0.4318	3	$\mathfrak{L}_1$
W	0.8039	2	$\mathfrak{L}_2^*$
Mermin-GHZ	0.7685	2	$\mathcal{S}_2 \cup \mathfrak{L}_1$

Table 5.2: Noise resistance to realizations in the quantum triangle scenario. For $v \geq v_{\max}$, the corresponding noisy distribution, in the sense of Eq. (5.18), cannot be realized in the triangle scenario when the latent variables are quantum. For computing the value for the W distribution we used the set $\mathfrak{L}_2^*$, which is a subset of the local level of order two, restricted to products of length no larger than 4.

not to admit realizations in the quantum triangle scenario. The results for the distributions already discussed along the chapter, namely the GHZ distribution of Eq. (5.29), the W distribution of Eq. (5.30) and the Mermin-GHZ distribution of Eq. (5.31), are shown in Table 5.2.

For visibilities above those depicted, the distributions are certified not to admit the corresponding inflations, and therefore they are not realizable in the quantum triangle scenario. It is possible, however, that these values are not tight, in the sense that there may exist values $v \leq v_{\max}$ for which the respective distribution, albeit possibly admitting second- or third-order inflations, does not admit higher-order ones.

In fact, the modification of quantum inflation to classical compatibility allows one to make a similar table for the robustness against realizations in terms of classical latent variables. The corresponding results are shown in Table 5.3.

For the problem of compatibility of a distribution with a causal structure, there are some nonlinear constraints, related to factorizations of expectation values, that acquire a linear form in the problem variables when making use of the distribution given. These constraints can be imposed in the associated SDPs, enforcing more restrictions that in general lower the visibility bounds presented above. In Appendix D.2 we elaborate on this matter, and show the analogues of Tables 5.2 and 5.3 when these constraints are enforced.

There is an interesting feature to notice. From the results of [SA15, WSF19] it is known that the GHZ distribution is not realizable in the triangle causal scenario, even for latent variables that are only limited by the no-signaling principle. Therefore, enforcing stricter "causality constraints", in the form of compatibility with higher-order inflations, should provide tighter approximations than imposing stricter "quantum/classical constraints" in the form of increased NPO levels. We observe this phenomenon explicitly, as lower values of $v_{\max}$ for P_{GHZ} are obtained at inflation level 3 using the NPO set $\mathfrak{L}_1$ than that obtained at inflation level 2 using the NPO set $\mathcal{S}_2 \cup \mathfrak{L}_1$. For the W distribution the con-

Distribution	$v_{\max}$	Inflation order	NPO level
GHZ	0.4286	3	$\mathfrak{L}_1$
W	0.8005	2	$\mathfrak{L}_2$
Mermin-GHZ	0.5000	2	$\mathfrak{S}_2 \cup \mathfrak{L}_1$

Table 5.3: Noise resistance to realizations in the classical triangle scenario. For $v \geq v_{\max}$, the corresponding noisy distribution, in the sense of Eq. (5.18), cannot be realized in the triangle scenario when the latent variables are classical.

straints on the causal structure are not as relevant (in fact, to date the question of whether P_{W} is realizable in the triangle scenario with nonsignaling resources remains open), and in fact the opposite occurs: in this case it is more important to enforce quantum constraints than to enforce those imposing compliance with the triangle structure.

A feature of semidefinite programming that has been widely employed in quantum information theory is the fact that the solution certificates of SDPs can be interpreted as a Bell-like inequalities or witnesses that are capable of identifying correlations not obtainable using quantum resources. The certificates obtained when using quantum inflation can have a similar interpretation as witnesses of quantum causal incompatibility. For instance, the certificate that provides the value of $v_{\max} = 0.8039$ for the W distribution in the quantum triangle scenario gives rise to the inequality

W's certificate

$$
\begin{aligned}
&-2P_{\mathrm{inf}}(A^{1,1}{=}0) + P_{\mathrm{inf}}(A^{1,1}{=}0, A^{2,2}{=}0)\\
&-2P_{\mathrm{inf}}(A^{1,1}{=}0, B^{1,1}{=}0) + 4P_{\mathrm{inf}}(A^{1,1}{=}0, B^{2,2}{=}0)\\
&-4P_{\mathrm{inf}}(A^{1,1}{=}0, B^{1,1}{=}0, A^{2,2}{=}0, B^{2,2}{=}0)\\
&+2P_{\mathrm{inf}}(A^{1,1}{=}0, A^{2,2}{=}0, B^{1,1}{=}0) + 2P_{\mathrm{inf}}(A^{1,1}{=}0, B^{1,1}{=}0, B^{2,2}{=}0)\\
&-2P_{\mathrm{inf}}(A^{1,2}{=}0, B^{1,2}{=}0, C^{1,2}{=}0)\\
&+[A \to B \to C \to A] + [A \to C \to B \to A]\\
&\leq 0.
\end{aligned}
$$

This certificate is defined in terms of a quantum distribution defined on a second-order inflation. Now, by using the moment identification rules of Eq. (5.26), this certificate can be transformed into a nonlinear witness of tripartite distributions whose violation signals the distribution as being incompatible with a realization in the triangle scenario with quantum latent variables (note that we switch to

the expectation-value picture):

W's witness

$$\begin{aligned}&\langle A\rangle+\langle A\rangle^2-\langle A\rangle\langle AB\rangle-\langle A\rangle\langle AC\rangle-2\langle BC\rangle+\langle B\rangle\langle C\rangle-\langle BC\rangle^2-\langle A\rangle\langle B\rangle\langle C\rangle\\&+[A\to B\to C\to A]+[A\to C\to B\to A]\\&\leq 3.\end{aligned} \tag{5.32}$$

In fact, the W distribution of Eq. (5.30) attains a value of $3+{}^8\!/_9$ for this witness. Furthermore, having an explicit witness one can now see that the numerical value of 0.8039 in Table 5.2 in reality corresponds to $3(2-\sqrt{3})$.

Upper bounds to optimal values

The constraints (5.23), (5.24) and (5.25) characterize relaxations of the set of quantum correlations compatible with a given causal structure. This characterization can be employed to easily bound optimal values of polynomials of the measurement operators in the problem via NPO theory [PNA10]. We now give some examples of applications of this procedure, by finding upper bounds to the maximum value that certain Bell-like operators can achieve in the quantum triangle scenario. We also use commuting-operator quantum inflation to bound the maximum values achievable in the classical triangle scenario, and complement these with the values for the tripartite quantum and classical Bell scenarios. The results herein have been obtained by considering second-order inflations, solving the associated NPO problems with the set of monomials $\mathfrak{S}_2 \cup \mathfrak{L}_1$. Including the identity operator, we find that each party has 9 possible operators at this SDP level, such that the resulting moment matrix involved in the calculations has size 873×873.

Mermin's Inequality [Mer90]

$$\langle A_1B_0C_0\rangle + \langle A_0B_1C_0\rangle + \langle A_0B_0C_1\rangle - \langle A_1B_1C_1\rangle \;\leq \begin{cases} 2 & \mathcal{L}^{\triangle},\, \mathcal{L} \\ 3.085^* & \mathcal{Q}^{\triangle} \\ 4 & \mathcal{Q},\, \mathcal{NS}^{\triangle} \end{cases}, \tag{5.33}$$

Svetlichny's Inequality [Sve87]

$$\begin{array}{l}\langle A_1B_0C_0\rangle + \langle A_0B_1C_0\rangle + \langle A_0B_0C_1\rangle - \langle A_1B_1C_1\rangle \\ -\langle A_0B_1C_1\rangle - \langle A_1B_0C_1\rangle - \langle A_1B_1C_0\rangle + \langle A_0B_0C_0\rangle\end{array} \leq \begin{cases} 4 & \mathcal{L}^{\triangle},\, \mathcal{L} \\ 4.405^* & \mathcal{Q}^{\triangle} \\ 4\sqrt{2} & \mathcal{Q} \\ 8 & \mathcal{NS}^{\triangle} \end{cases}. \tag{5.34}$$

The asterisk means that the values given are only upper bounds; that is, quantum inflation shows that Mermin's and Svetlichny's inequalities cannot exceed 3.085 or 4.405, respectively, in the quantum triangle scenario, but those values may not be achievable either.

Interestingly, it is known that both the algebraic maximum of 4 for Mermin's inequality and the algebraic maximum of 8 for Svetlichny's inequality can be achieved even in the triangle structure, if one considers that the latent nodes distribute nonsignaling resources. For Mermin's inequality, in fact, the algebraic maximum can be obtained when parties A and B share a PR box while C acts deterministically. Similarly, the construction in the triangle scenario achieving the algebraic maximum for Svetlichny's inequality was demonstrated in [BLM$^+$05, Sec. III-C]. A similar phenomenon occurs in the classical case. It suffices that A and B output 0 irrespective of the values of the latent variables they receive, while C outputs 0 upon receiving, for instance, $\lambda_{\text{AC}} \oplus \lambda_{\text{BC}} = 0$ (this will be the measurement C_0), and a random value upon receiving $\lambda_{\text{AC}} \oplus \lambda_{\text{BC}} = 1$, to achieve the local bound of Mermin's inequality. If upon receiving $\lambda_{\text{AC}} \oplus \lambda_{\text{BC}} = 1$, C outputs 1, then the maximal local value of Svetlichny's inequality is achieved.

However, a difference exists between the correlations achievable in the triangle scenario and in the tripartite Bell scenario with quantum latent variables. In the triangle scenario it is not possible to achieve the quantum maximal bounds of the inequalities, so having access to tripartite resources allows for improved performance in information tasks. In fact, given that the values given for $\mathcal{Q}^{\triangle}$ represent upper bounds to the real maximum values achievable, it is an open question whether one can identify quantum correlations in the triangle scenario using these inequalities or if, when increasing the inflation and NPO hierarchies, one will find that the values for $\mathcal{Q}^{\triangle}$ will collapse to those of $\mathcal{L}$. However, a value larger than 3.085 of Mermin's inequality, or larger than 4.405 of Svetlichny's inequality, certify that the observed correlations are impossible to produce in a quantum triangle setup.

It is possible to optimize also nonlinear quantities, such as the witness derived in the previous section. Despite being a nonlinear object, it can be written as a linear objective subject to inflation constraints, and thus be optimized with

semidefinite programming. In fact, one has that

W's witness

$$\text{Equation (5.32)} \leq \begin{cases} 3^* & \mathcal{L}^{\triangle},\ \mathcal{Q}^{\triangle} \\ 3+\frac{8}{9} & \mathcal{L},\ \mathcal{Q},\ \mathcal{NS} \end{cases}. \tag{5.35}$$

The upper bounds for $\mathcal{L}$, $\mathcal{Q}$ and $\mathcal{NS}$ are tight, and can be obtained by cylindrical algebraic decomposition [Col75]. However, for scenarios with more parties or more complex polynomials where other techniques are not efficient, one could employ quantum inflation in Bell scenarios to upper-bound this value. Considering a sixth-order inflation of the tripartite Bell scenario, and the SDP associated to the subset of the corresponding $\mathfrak{L}_6$ composed of operator strings of length shorter than 4 (giving rise to a moment matrix Γ of size 988×988), one still obtains an upper bound of $4 + 3/5$, reasonably above the actual optimum.

5.5 Comparison: inflation vs. extension

Quantum inflation and scalar extension are two methods of very different nature. While quantum inflation can be used to study any kind of causal structures, scalar extension is designed to impose very specific types of constraints, which may make it a better choice when studying specific causal structures. We now contrast both techniques at identifying correlations incompatible with quantum and classical realizations in the tripartite-line scenario of Fig. 5.1d, where both techniques can be applied. We analyze both the tightness of the approximations to the set of compatible correlations that each method produces, and their computational requirements.

For performing the comparison, we study the noise robustness of the extremal points of the tripartite nonsignaling polytope, which were completely characterized in [PBS11]. In fact, the P^{22} distribution of Eq. (5.19) corresponds to point number 45.

We will not use the noise model of Eq. (5.18). Instead, we will insert the causal structure into the noise. We will assume that the probability distributions are constructed from noisy latent variables, which with probability $1-v$ send random, uncorrelated information to both parties they influence. This model can be written as

$$\begin{aligned} P_{v_{\text{AB}},v_{\text{BC}}} =& v_{\text{AB}}v_{\text{BC}}P + v_{\text{AB}}(1-v_{\text{BC}})P_{\text{A}}\mathbb{1}_{\text{BC}} \\ &+ (1-v_{\text{AB}})v_{\text{BC}}\mathbb{1}_{\text{AB}}P_{\text{C}} + (1-v_{\text{AB}})(1-v_{\text{BC}})\mathbb{1}_{\text{ABC}}, \end{aligned} \tag{5.36}$$

where v_{AB} and v_{BC} are the visibilities of both sources, P_p is the marginal probability distribution over the variable p, and $\mathbb{1}$ is the uniformly random distribution. For simplicity, we will assume that $v_{AB} = v_{BC} = v$.

5.5.1 Quantum compatibility

We begin the comparison by looking at the smallest visibility needed in order to identify that the corresponding distribution does not have a realization in terms of quantum latent variables. Let us begin by recalling the results for P^{22} in the presence of global noise [this is, that of Eq. (5.18)]. For this noise model, it was known that a quantum realization was not possible for $v > 1/2$. As shown in Sections 5.3 and 5.4, both methods were able to recover this result. Using scalar extension, the SDP that identified all incompatible distributions asks for the positivity of a moment matrix Γ built from the set of monomials $\mathcal{S}_3$, extended with the set of operators $\{\langle S_i\rangle\mathbb{1}\}_i$, where S_i is any product of operators of party A of length $2 \leq \ell \leq 5$. These monomials generate a moment matrix of size 71×71. On the other hand, quantum inflation is capable of recovering the result using a partial inflation where there are two copies of ρ_{AB} but only one of ρ_{BC}, and the corresponding monomial set $\mathcal{S}_2$, that generates a moment matrix of size 65×65. In fact, it is even enough with the 43×43 moment matrix generated by the analogue of the set $\mathcal{S}_{1+AB} \subsetneq \mathcal{S}_2$, which does not contain monomials of the type $H_a H_b$ [NPA08].

Now we compare the visibilities, per Eq. (5.36) with $v_{AB} = v_{BC}$, beyond which the methods can identify that the extremal points of the tripartite nonsignaling polytope do not have realizations in the tripartite-line scenario with classical latent variables. The results can be found in Fig. 5.10.

For scalar extension, we solve the NPO problem associated to the monomial set $\mathcal{S}_5$, extended with all monomials of the form $\langle S_i\rangle\mathbb{1}$ where S_i is any product of the operators of party A of length $2 \leq \ell \leq 9$. This problem involves finding the positivity of a moment matrix $\tilde{\Gamma}$ of size 247×247. For quantum inflation, the results are obtained by considering the compatibility with a third-order inflation, solving the NPO problem associated to the corresponding monomial set $\mathfrak{L}_1$, which involves a moment matrix of size 931×931. By looking at Fig. 5.10 it is apparent that there is no clear advantage when using one method instead of the other. Out of the 46 boxes, in 16 of them scalar extension identifies a larger range of visibilities for which the corresponding noisy distribution is not realizable in the tripartite-line scenario with quantum latent variables, in 13 of them it is the opposite, and quantum inflation provides stricter bounds to $v_{\max}$, and in the remaining 17 both methods provide the same bound. However, the computational requirements for obtaining the bounds using scalar extension

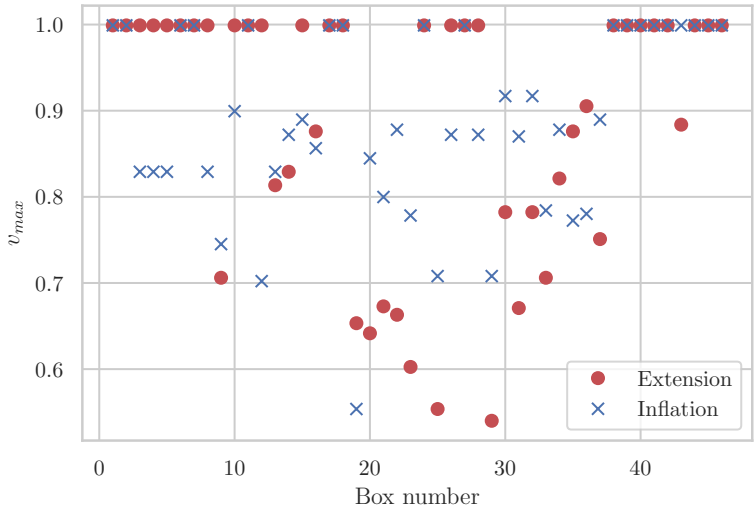

Figure 5.10: Values of the maximum visibility $v_{\max}$ beyond which quantum inflation (in blue crosses) and scalar extension (in red dots) identify the vertices of the tripartite nonsignaling polytope as incompatible with quantum realizations in the tripartite-line scenario. The values for scalar extension are obtained using the corresponding NPO level $\mathcal{S}_5$, extended with the monomial set $\{\langle\prod_{i=1}^{n}\Pi_i^{\mathrm{A}}\rangle\mathbb{1}\}_{n=2}^{9}$, where Π_i^{A} is any projector acting on A's Hilbert space. The values for quantum inflation are obtained when solving the problem of compatibility with a third-order inflation, using the associated NPO set $\mathfrak{L}_1$.

are notably lower than those needed for quantum inflation. As we shall see in the next section, the comparison will be much easier in the case of identifying nonclassical distributions.

It is important to recall that, for both scalar extension and quantum inflation, we deal with relaxations of the set of correlations compatible with the quantum tripartite-line scenario. Therefore, the values of $v_{\max}$ shown are just upper bounds to the real visibilities below which the correlations begin to admit bilocal hidden variable models of the form (5.16). It has to be stressed that this is the case even when both methods provide the same bound, although due to their intrinsically different nature, this fact constitutes a strong evidence that the value given is tight.

5.5.2 Classical compatibility

As it has been already discussed in the previous sections and in Appendix D.1, both scalar extension and quantum inflation can be used for characterizing

classical correlations in causal structures. It is thus possible to compare how both methods compare when solving problems in classical causal inference.

To begin with, we already saw that scalar extension allowed to discard every distribution of the family P_v^{22} not realizable in terms of classical latent variables. This follows from solving the SDP associated to the corresponding monomial set $\mathcal{S}_2 \cup \mathfrak{L}_1$, extended with the only monomial $\langle \Pi^{A}_{0|0}\Pi^{A}_{0|1}\rangle\mathbb{1}$, resulting in a moment matrix of size 31×31. In contrast, the commuting-variable variant of quantum inflation needs much higher levels in the hierarchy in order to produce tight results. In fact, for the largest problems in our computational reach, namely determining compatibility with a third-order inflation using the monomial set $\mathfrak{L}_1$ (which requires a 931×931 moment matrix), and determining compatibility with a second-order inflation using the subset of the monomial set $\mathfrak{L}_2$ composed of products of length smaller or equal to 3 (which requires a moment matrix of size 633×633), the family of distributions for $v \in [1/4, 0.328]$ was not identified as incompatible.

The same happens when comparing the visibilities beyond which the methods can identify that the extremal points of the tripartite nonsignaling polytope do not have realizations in the tripartite-line scenario with classical latent variables. The results obtained are shown in Fig. 5.11. For scalar extension, we solve the NPO problem associated to the monomial set $\mathcal{S}_6$, that characterizes completely the set of classical correlations (see Appendix D.1) in the tripartite Bell causal structure of Fig. 5.1b, extended with the only monomial $\langle \Pi^{A}_{0|0}\Pi^{A}_{0|1}\rangle\mathbb{1}$ to further restrict to the tripartite-line scenario of Fig. 5.1d. This problem involves finding the positivity of a moment matrix $\tilde{\Gamma}$ of size 65×65. For quantum inflation, the results are obtained by considering the compatibility with a third-order inflation, solving the NPO problem associated to the corresponding monomial set $\mathfrak{L}_1$, which involves a moment matrix of size 931×931.

Despite the great difference in problem sizes, scalar extension detects the incompatibility of a much larger range of distributions than quantum inflation. In fact, scalar extension is capable of identifying nonbilocal distributions that quantum inflation does not for all nonsignaling boxes, except for box number 1 where the $v_{\max}$ given by both methods coincide in the exact value. This is the trivial box representing the deterministic strategies, which have realizations in the tripartite-line scenario for any v. It is also notable that for 17 of the boxes, third-level inflations cannot identify incompatible correlations. In fact, when contrasting the values provided by quantum inflation in Figs. 5.10 and 5.11, one sees that inserting the constraints associated to classical scenarios is not imposing any new restrictions in the problems being solved. We must note that, as exemplified by the results of the P^{22} distribution and the distributions in the triangle scenario discussed in Section 5.4.3, in general classical constraints

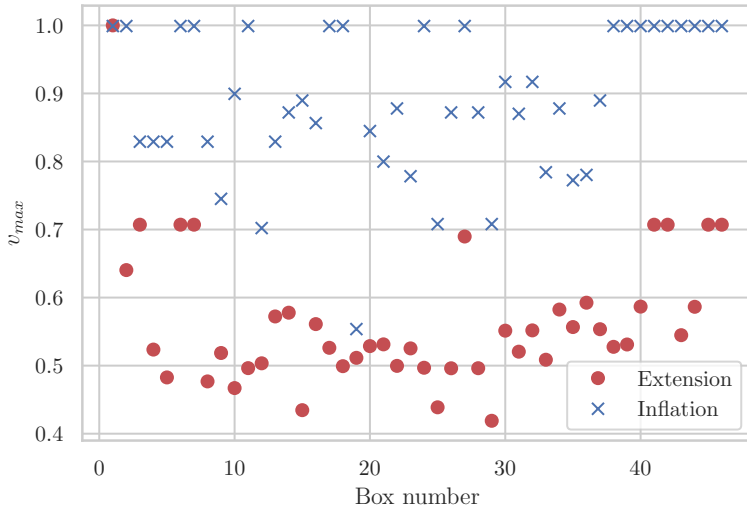

Figure 5.11: Values of the maximal visibility $v_{\max}$ beyond which quantum inflation (in blue crosses) and scalar extension (in red dots) identify the vertices of the tripartite nonsignaling polytope as incompatible with local realizations in the tripartite-line scenario. The values for scalar extension are obtained using the corresponding NPO level $\mathcal{S}_6$ (the highest possible, per Appendix D.1), extended with the monomial $\langle \Pi^{A}_{0|0}\Pi^{A}_{0|1}\rangle\mathbb{1}$. The values for quantum inflation are obtained when solving the problem of compatibility with a third-order inflation, using the associated NPO set $\mathfrak{L}_1$.

in quantum inflation are not redundant.

All in all, these results show that scalar extension enforces in a more effective way the constraints arising from causal independence while having lower computational demands, which allow one to analyze stricter relaxations of the sets of compatible correlations and thus obtain tighter bounds to optimal values.

5.6 Discussion

In this chapter we have introduced two techniques, *scalar extension* and *quantum inflation*, that allow for testing the hypothesis that an observed multipartite correlation has been generated in a particular quantum network.

On the one hand, scalar extension imposes relaxations of the independence constraints between causally independent nodes in a causal structure. We have applied this method to the study of correlations in the tripartite-line scenario, showing that quantum correlations not explainable by realizations in terms of

classical latent variables can be established between the parties even in the case of very low detection efficiencies—at least as low as 0.001%. The results imply that it is possible to activate the nonlocal properties of measurement devices by arranging them in a network geometry. This could, in principle, be used to establish nonlocal correlations among long distances: having the measurement devices of parties A and B in a controlled research facility with high-performance components (so party A's measurement device has a high detection efficiency), nonlocal correlations could be established between the research facility and a cheap, low-performance client apparatus C.

On the other hand, quantum inflation is a systematic method to discern whether an observable distribution is compatible with a causal explanation involving quantum degrees of freedom. With it, we have characterized the quantum correlations achievable in the triangle causal structure, proving that some well-known distributions cannot be generated by performing measurements on quantum states, and computing upper bounds to the maximum values achievable of multipartite Bell inequalities.

Finally, we have analyzed how both techniques compare in a causal structure where both are applicable. We observed that, while scalar extension is a technique applicable only to a certain family of causal structures, in general it enforces factorization constraints in a stronger way than quantum inflation, thereby leading to tighter results with less computational demand in those scenarios where it is applicable. This is the case specially when constraining correlations achievable with classical latent variables.

Our results constitute fundamental contributions to the characterization of quantum correlations in networks. Due to the fact that a small modification allows for the characterization of the sets of classical correlations achievable in networks, our results are also of importance for the field of classical causal inference. Nevertheless, some questions remain open.

The most important problem to be solved is whether the methods tightly describe the corresponding sets of correlations in the asymptotic limit. There is special motivation for the case of quantum inflation, since it is known that its commuting-variable version converges to the classical inflation method, and this one presents asymptotic convergence to the sets of correlations in classical causal structures. Moreover, it is also known that, for a fixed inflation level, the associated NPO hierarchy to determine compatibility with the inflation converges in the asymptotic limit.

From the point of view of applications, our results enable the construction and study of complex multipartite quantum information protocols. We expect that this will be a major field of research in the near future. For example, the use of scalar extension in line scenarios with more nodes could be of relevance

for the certification of nonclassical correlations in repeater networks.

Moreover, fundamental questions may also benefit of the methods described. They can be used for attempting to solve paradigmatic problems in quantum information [Gis17], and even to identify quantum phenomena in thermodynamic or biological processes. For this, it will be crucial to develop numerical tools to reduce the complexity of the computations involved [Ros18].

6 Conclusions and outlook

Quantum information science is a field that has experienced an enormous growth in the last two decades. In fact, it has been used to tackle both foundational problems in the theory of quantum mechanics, technological challenges when building commercial devices, and many other topics in between. This success has led to the fact that now, ideas from quantum information science are gradually being adopted in various other areas of science and engineering.

This thesis explored, precisely, the role of quantum information outside quantum information. This was done in two different ways. On the one hand, we experimented the application of methods and ideas natural to quantum information science in other areas such as thermodynamics in the microscopic regime or causal inference. On the other, we analyzed the role and impact of quantum information—understood as information stored in quantum systems—in fields as fundamental as quantum field theory, or as applied as machine learning.

Our results provide new insights on, and methods to study, the role of quantum effects in various areas of physics. In this chapter we briefly review the main conclusions drawn during this journey, and explore open directions and avenues for future work.

Quantum information in quantum field theory

In Chapter 2 we analyzed the accessibility of the entanglement in the vacuum state of quantum fields through its transfer to quantum systems that can be manipulated easily. In a general setting of entanglement harvesting, we have given necessary conditions for the extraction of entanglement from spacelike-separated spacetime regions when the quantum systems employed for the harvesting have degenerate energy levels. Also, focusing in the setting of harvesting from the electromagnetic vacuum, we have studied how much entanglement can be extracted by realistic atomic probes. We have analyzed how this extraction is affected by varying the different experimental parameters in the setup, such as the relative orientation and distance between the probes or the time delay between the interaction of each probe with the field. For doing so, we developed a

general model of relativistic light-matter interaction that is capable of including any arbitrary transition in the study of atomic entanglement harvesting from the electromagnetic vacuum.

On the ambitious, long-term goal of the measurement and exploitation of vacuum correlations, our results make significant contributions towards a realistic modeling of the phenomenon. However, a major concern is the low magnitude of the figures of merit involved. Therefore an important line of future research is that of finding similar phenomena in more experimentally friendly setups, for instance by harvesting timelike entanglement from squeezed or coherent states of fields in cavities.

A shorter-term interesting line of research is the assessment of the effect of the quantum vacuum in quantum information protocols. While there have already been some results, for instance in assessing how much information is leaking to the vacuum—and thus could be read by an adversary—during a process of randomness generation and how this leakage can be minimized, the effect of the quantum vacuum in protocols such as quantum key distribution is very much unexplored. Also, one can consider the vacuum as an active element of their setup, and assess how it can be harnessed for devising better or more efficient protocols. An example of this could be the exploitation of vacuum correlations in a repeated manner for generating entangled pairs between distant parties.

Quantum information theory for thermodynamics

In Chapter 3 we showed how concepts and formalisms borrowed from the theories of quantum information and quantum optics can provide insights in the analysis of thermodynamic processes. We used a popular formalism in quantum information theory to develop the first model of thermodynamical phenomena in the microscopic regime that does not rely in any of the approximations performed in macroscopic thermodynamics. We created a model that is able to circumvent the standard assumptions of weak coupling, slow driving, and infinite size of the baths, and which allows one to explore the individual constituents of the systems in play. This last feature is of special relevance, as it allows one to find explanations to observed global-scale phenomena in the dynamics at the scale of the individual constituents.

To demonstrate its power, we studied a harmonic oscillator undergoing an Otto cycle between two finite-sized baths. We were able to analyze the exchanges of work and heat between the different elements, and observe the exhaustion of the machine after a finite number of cyclic operations. By analyzing the buildup of correlations between the thermal baths and the working medium, we could relate the creation of bath-bath correlations with the decrease in performance

of the machine, and give a precise estimate of the amount of cycles that the engine can undergo before exhaustion in terms of the size of the baths.

The Otto cycle is one of the simplest, yet very important, thermodynamic cycle one can consider studying. A line of future research is thus applying the same formalism to understand the dynamics of more complex thermodynamic processes. In fact, an important problem that remains open is that of the characterization of adiabatic processes within the formalism of Gaussian quantum mechanics. Interesting questions that would be of particular relevance for this are demonstrating the existence of Hamiltonian evolutions that lead to adiabatic transformations, and the construction of methods for quantifying deviations from adiabaticity.

In the model we observe a very strong relation between the performance of the engine and how correlated the baths are to each other. It would be desirable to have an experimental confirmation of this relation, which would aid in devising protocols for extending the lifetime of quantum thermal engines. While this may require inserting more realistic forms of the different interactions that take place in the operation of the engine, the formalism of Gaussian quantum mechanics is still capable of modeling operations that occur in experimental scenarios.

Quantum information for machine learning

In Chapter 4 we explored a particular application in which encoding and processing information in quantum systems could provide a computational advantage over encoding and processing it in classical systems, namely, in the field of machine learning. We developed a quantum algorithm for training feedforward neural networks using Bayesian techniques. This type of training, desirable for multiple reasons, is hard to execute in classical computers for datasets of practical relevance. Our algorithm reduces the computational cost of the training by exploiting quantum entanglement and superposition. Furthermore, in order to estimate the feasibility of our approach, we implemented the core routine of the algorithm in state-of-the-art quantum processors.

These results fall within a field—quantum computing—that has recently obtained a boost in attention primarily due to its potential commercial applications and new developments in hardware. It is expected that in the near or mid-term future new devices will be built in which one will be able to execute in full the protocol developed, and many other, for solving complex problems of practical relevance. Until then, an important avenue of research will be how to efficiently exploit quantum properties in small, noisy and sparsely connected quantum chips. In this regard, hybrid algorithms where one optimizes the classical parameters of a quantum circuit represent a promising approach, as they

are able to obviate experimental imperfections.

In a broader scope, algorithms similar to the one developed in Chapter 4 could have a strong role in making unsupervised learning widely available. To date, unsupervised learning of the important features of a set of data is a very much desired task in many fields of research and industry. However, its high computational requirements—even higher than those for supervised learning—and difficulties for training prevent its use even in simple scenarios. Quantum computers would not only provide an advantage in terms of the computational resources needed for training, but it is also expected that using quantum systems to encode information enable models with improved abilities for learning data. Therefore, aside from the still-needed hardware developments, machine learning will be very much benefited from research in quantum algorithms for unsupervised learning.

Causality with quantum information

In Chapter 5 we studied whether it is possible to understand the cause-effect relations between measured quantities when these causes and effects may involve quantum systems. We developed two methods, based on relaxations of the original problem, to test the hypothesis that an observed multipartite correlation has been generated in any given setup where quantum states and/or local hidden variables are distributed between the parties. One of the methods can only be applied to a specific family of causal configurations, while the second can be applied to contrast against any causal structure. Nevertheless, we showed that the restricted-case method generally performs better than the general-case method when identifying correlations not compatible with scenarios where both techniques are applicable.

A field that will directly benefit from the techniques developed is quantum information science. In fact, we have already obtained a number of results, such as upper bounds to the maximum values of Bell inequalities achievable in different causal scenarios, that will be of great relevance in current and new protocols of quantum information processing. For instance, one can imagine situations where one would wish to certify that parties share resources according to some structure before revealing some information to all (or only some) of them. On a more fundamental view, it will be interesting to understand the implications in multipartite entanglement and nonlocality that derive from the techniques. In this sense, one can think of devising means of certifying the genuinely multipartite character of nature, for instance by experimentally observing a violation of a Bell-like causal inequality above the limit achievable with restricted correlations.

However, the extent of the results obtained goes beyond quantum informa-

tion science. Indeed, the fact that the techniques developed can also be used to characterize classical correlations lead to potential applications in the fields of statistical inference and machine learning. An interesting line of research that can be thought of is the hard-coding—or even the discovery—of causal constraints in machine learning algorithms for various applications, or more ambitiously in the field of cognitive robotics. Moreover, one of the methods—scalar extension—can be viewed as a way of relaxing nonlinear constraints in convex optimization problems. As such, a problem of great relevance is to understand how the technique can be used in general convex optimization problems, or to relax nonconvex problems into convex ones. A generalization of this type could have applications in disciplines like finance, structural optimization, or circuit design.

A current drawback of the methods is their high computational cost. Thus, an important line of future research is the development of efficient implementations. An example of promising avenues is the exploitation of the many symmetries the methods introduce, which may help reducing drastically the size of the computational problems to be solved, enabling the study of correlations in larger and more complex causal scenarios.

Appendices

Appendices of Chapter 2

A.1 Completeness relation of the polarization vectors

In this appendix we derive explicitly the completeness relation of Eq. (2.23), following [CTDRG87, page 36]. Recall that the polarization basis $\{\boldsymbol{\epsilon}(\boldsymbol{k},s)\}_{s=1}^{3}$ is an orthonormal basis of $\mathbb{R}^3$. We can choose $\boldsymbol{\epsilon}(\boldsymbol{k},3)\propto\hat{\boldsymbol{k}}$, the other two unit vectors being mutually orthogonal and perpendicular to this, which we denote $\boldsymbol{\epsilon}(\boldsymbol{k},\perp_1)$ and $\boldsymbol{\epsilon}(\boldsymbol{k},\perp_2)$. With this choice, the completeness relation (2.23), when written in components, reads

$$\sum_{s=\perp_1,\perp_2}\epsilon_j(\boldsymbol{k},s)\epsilon_l(\boldsymbol{k},s)=\epsilon_j(\boldsymbol{k},\perp_1)\epsilon_l(\boldsymbol{k},\perp_1)+\epsilon_j(\boldsymbol{k},\perp_2)\epsilon_l(\boldsymbol{k},\perp_2). \tag{A.1}$$

Now, using an auxiliary orthonormal basis of $\mathbb{R}^3$, $\{\mathbf{e}_i\}_{i=1}^{3}$, the sum can be recast as

$$\begin{aligned}\sum_{s=\perp_1,\perp_2}\epsilon_j(\boldsymbol{k},s)\epsilon_l(\boldsymbol{k},s)=&\Big(\boldsymbol{e}_j\cdot\boldsymbol{\epsilon}(\boldsymbol{k},\perp_1)\Big)\Big(\boldsymbol{e}_l\cdot\boldsymbol{\epsilon}(\boldsymbol{k},\perp_1)\Big)\\&+\Big(\boldsymbol{e}_j\cdot\boldsymbol{\epsilon}(\boldsymbol{k},\perp_2)\Big)\Big(\boldsymbol{e}_l\cdot\boldsymbol{\epsilon}(\boldsymbol{k},\perp_2)\Big)\\&+\Big(\boldsymbol{e}_j\cdot\hat{\boldsymbol{k}}\Big)\Big(\boldsymbol{e}_l\cdot\hat{\boldsymbol{k}}\Big)-\Big(\boldsymbol{e}_j\cdot\hat{\boldsymbol{k}}\Big)\Big(\boldsymbol{e}_l\cdot\hat{\boldsymbol{k}}\Big),\end{aligned} \tag{A.2}$$

where in the last line we have added and subtracted the same quantity. Written in this form, the first three terms of Eq. (A.2) represent the coefficients of the scalar product $\boldsymbol{e}_j\cdot\boldsymbol{e}_l$ expressed in the $\{\boldsymbol{\epsilon}(\boldsymbol{k},\perp_1),\boldsymbol{\epsilon}(\boldsymbol{k},\perp_2),\hat{\boldsymbol{k}}\}$ basis, while the last term represents the product of the components of $\hat{\boldsymbol{k}}$ in the directions of $\boldsymbol{e}_j$ and $\boldsymbol{e}_l$. Therefore one can write

$$\sum_{s=\perp_1,\perp_2}\epsilon_j(\boldsymbol{k},s)\epsilon_l(\boldsymbol{k},s)=\boldsymbol{e}_j\cdot\boldsymbol{e}_l-\frac{k_jk_l}{|\boldsymbol{k}|^2}=\delta_{jl}-\frac{k_jk_l}{|\boldsymbol{k}|^2}, \tag{A.3}$$

which is just the expression in components of Eq. (2.23).

A.2 Positivity of the density matrix

In this appendix we explicitly show how the time-evolved density matrix (2.28) satisfies the positivity conditions for an X state given in [MMST16]. Following [MMST16], a generic density matrix for an X state can be written as follows

$$\hat{\rho}_{\mathrm{X}} = \begin{pmatrix} r_{11} & 0 & 0 & r_{14}e^{-\mathrm{i}\xi} \\ 0 & r_{22} & r_{23}e^{-\mathrm{i}\varsigma} & 0 \\ 0 & r_{23}e^{\mathrm{i}\varsigma} & r_{33} & 0 \\ r_{14}e^{\mathrm{i}\xi} & 0 & 0 & r_{44} \end{pmatrix}, \tag{A.4}$$

where $\{r_{ij}\}, \xi, \varsigma \in \mathbb{R}$.

A necessary condition for this matrix to represent a quantum state is that it is positive-semidefinite, which amounts to its eigenvalues being nonnegative. This restriction (as noted in [MMST16]) imposes the following constraints on the elements of $\hat{\rho}_{\mathrm{X}}$:

$$r_{11}r_{44} \geq r_{14}^2, \tag{A.5}$$

$$r_{22}r_{33} \geq r_{23}^2. \tag{A.6}$$

In the following, we will show explicitly how the matrix (2.28) is positive-semidefinite (to leading order) by analyzing its eigenvalues, and we will relate the results to the above constraints (A.5) and (A.6).

The eigenvalues of Eq. (2.28) are

$$E_1 = \frac{1}{2}\left\{1 - (L_{\mathrm{AA}} + L_{\mathrm{BB}})e^2 + \sqrt{[1 - e^2(L_{\mathrm{AA}} + L_{\mathrm{BB}})^2]^2 + 4e^4|M|^2}\right\} + \mathcal{O}(e^4), \tag{A.7a}$$

$$E_2 = \frac{1}{2}\left\{1 - (L_{\mathrm{AA}} + L_{\mathrm{BB}})e^2 - \sqrt{[1 - e^2(L_{\mathrm{AA}} + L_{\mathrm{BB}})^2]^2 + 4e^4|M|^2}\right\} + \mathcal{O}(e^4), \tag{A.7b}$$

$$E_3 = \frac{1}{2}e^2\left[L_{\mathrm{AA}} + L_{\mathrm{BB}} + \sqrt{(L_{\mathrm{AA}} - L_{\mathrm{BB}})^2 + 4|L_{\mathrm{AB}}|^2}\right] + \mathcal{O}(e^4), \tag{A.7c}$$

$$E_4 = \frac{1}{2}e^2\left[L_{\mathrm{AA}} + L_{\mathrm{BB}} - \sqrt{(L_{\mathrm{AA}} - L_{\mathrm{BB}})^2 + 4|L_{\mathrm{AB}}|^2}\right] + \mathcal{O}(e^4), \tag{A.7d}$$

where $\mathcal{L}_{\mu\nu} = e^2 L_{\mu\nu}$ and $\mathcal{M} = e^2 M$, and we have assumed $e_{\mathrm{A}} = e_{\mathrm{B}} = e$ for simplicity.

It is already clear that $E_1 > 0$ and $E_3 > 0$. Expanding E_2 and E_4 in powers of the coupling strength we obtain

$$E_2 = -e^4|M|^2 + \mathcal{O}(e^6), \tag{A.8}$$

$$E_4 = \frac{1}{2}e^2\left(L_{\mathrm{AA}} + L_{\mathrm{BB}} - \sqrt{(L_{\mathrm{AA}} - L_{\mathrm{BB}})^2 + 4|L_{\mathrm{AB}}|^2}\right) + \mathcal{O}(e^4). \tag{A.9}$$

The eigenvalue E_2 seems to always be less than zero. Nevertheless, it must be noted that this contribution appears at fourth order, and therefore, to second order we have $E_2 = 0 + \mathcal{O}(e^4)$, which is compatible with the matrix being positive-semidefinite at this order. A very similar issue has been discussed when computing the eigenvalues of the partial transpose of Eq. (2.28), as well as in previous works [PKMM15]. This result relates to the first condition (A.5). Recalling Eq. (2.28), it is easy to see that

$$r_{11} = \mathcal{O}(1), \quad r_{14} = \mathcal{O}(e^2) > 0, \quad r_{44} = \mathcal{O}(e^4). \tag{A.10}$$

Therefore, to second order in the coupling strength, the inequality (A.5) is satisfied trivially.

Notice that in order to take into account the $\mathcal{O}(e^4)$ contribution to E_2 with a consistent perturbative analysis, the full fourth-order correction to the state (2.28) should be computed. If one analyzes E_2 consistently to fourth order in perturbation theory including the contributions missing in Eq. (A.7b), it can be proved that E_2 is strictly positive.

On the other hand, the requirement that $E_4 \geq 0$ implies

$$L_{\mathrm{AA}} L_{\mathrm{BB}} \geq |L_{\mathrm{AB}}|^2, \tag{A.11}$$

which is actually the second condition $r_{22} r_{33} \geq r_{23}^2$.

To check that this inequality is actually satisfied, recall the expression for $L_{\mu\nu}$, obtained from Eq. (2.29),

$$\begin{aligned} L_{\mu\nu}^{\mathrm{EM}} = &\int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2 \, e^{\mathrm{i}(\Omega_\mu t_1 - \Omega_\nu t_2)} \mathcal{X}_\mu(t_1) \mathcal{X}_\nu(t_2) \\ &\times \int \mathrm{d}^3\boldsymbol{x}_1 \int \mathrm{d}^3\boldsymbol{x}_2 \, \boldsymbol{F}_\nu^{*\mathbf{t}}(\boldsymbol{x}_2 - \boldsymbol{x}_\nu) \mathbf{W}(\boldsymbol{x}_2, \boldsymbol{x}_1, t_2, t_1) \boldsymbol{F}_\mu(\boldsymbol{x}_1 - \boldsymbol{x}_\mu). \end{aligned} \tag{A.12}$$

When we insert the expressions for the smearing vectors from Eq. (2.12) and the Wightman 2-tensor of the electric field from Eq. (2.31), this expression reads

$$\begin{aligned} L_{\mu\nu}^{\mathrm{EM}} = &\int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2 e^{\mathrm{i}(\Omega_\mu t_1 - \Omega_\nu t_2)} \mathcal{X}_\mu(t_1) \mathcal{X}_\nu(t_2) \\ &\times \int \mathrm{d}^3\boldsymbol{x}_1 \int \mathrm{d}^3\boldsymbol{x}_2 \, \psi_e(\boldsymbol{x}_2 - \boldsymbol{x}_\mu) \psi_g^*(\boldsymbol{x}_2 - \boldsymbol{x}_\mu) \psi_e^*(\boldsymbol{x}_1 - \boldsymbol{x}_\nu) \psi_g(\boldsymbol{x}_1 - \boldsymbol{x}_\nu) \\ &\times \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3} \frac{|\boldsymbol{k}|}{2} e^{-\mathrm{i}|\boldsymbol{k}|(t_2 - t_1)} e^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}_2} e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}_1} \, (\boldsymbol{x}_2 - \boldsymbol{x}_\mu)^{\mathbf{t}} \left(\mathbb{1} - \frac{\boldsymbol{k} \otimes \boldsymbol{k}}{|\boldsymbol{k}|^2} \right) (\boldsymbol{x}_1 - \boldsymbol{x}_\nu). \end{aligned} \tag{A.13}$$

Finally, we perform the change of variables $\boldsymbol{x} = \boldsymbol{x}_1 - \boldsymbol{x}_\nu$, $\boldsymbol{x}' = \boldsymbol{x}_2 - \boldsymbol{x}_\mu$ to get the final expression

$$\begin{aligned} L_{\mu\nu}^{\text{EM}} &= \int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{\infty}\mathrm{d}t_2\, e^{\mathrm{i}(\Omega_\mu t_1 - \Omega_\nu t_2)}\mathcal{X}_\mu(t_1)\mathcal{X}_\nu(t_2)\int\mathrm{d}^3\boldsymbol{x}\int\mathrm{d}^3\boldsymbol{x}'\,\psi_e(\boldsymbol{x}')\psi_g^*(\boldsymbol{x}')\psi_e^*(\boldsymbol{x})\psi_g(\boldsymbol{x}) \\ &\quad\times\int\frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\frac{|\boldsymbol{k}|}{2}\,e^{-\mathrm{i}|\boldsymbol{k}|(t_2-t_1)}e^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}'}e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\,\left(\boldsymbol{x}'\right)^{\mathbf{t}}\left(\mathbb{1}-\frac{\boldsymbol{k}\otimes\boldsymbol{k}}{|\boldsymbol{k}|^2}\right)\boldsymbol{x}\,e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_\mu-\boldsymbol{x}_\nu)} \\ &= \sum_i\int\frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\frac{|\boldsymbol{k}|}{2}\left|\int_{-\infty}^{\infty}\mathrm{d}t\,e^{\mathrm{i}(\Omega+|\boldsymbol{k}|)t}\mathcal{X}(t)\int\mathrm{d}^3\boldsymbol{y}\,\psi_e^*(\boldsymbol{y})\psi_g(\boldsymbol{y})e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{y}}\,y_i\right|^2 e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_\mu-\boldsymbol{x}_\nu)} \\ &\quad-\int\frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\frac{|\boldsymbol{k}|}{2}\left|\int_{-\infty}^{\infty}\mathrm{d}t\,e^{\mathrm{i}(\Omega+|\boldsymbol{k}|)t}\mathcal{X}(t)\int\mathrm{d}^3\boldsymbol{y}\,\psi_e^*(\boldsymbol{y})\psi_g(\boldsymbol{y})e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{y}}\,\frac{\boldsymbol{y}\cdot\boldsymbol{k}}{|\boldsymbol{k}|}\right|^2 e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_\mu-\boldsymbol{x}_\nu)}, \end{aligned} \tag{A.14}$$

where we have already assumed that the two atoms are equal, and y_i denote the components of the vector $\boldsymbol{y}$.

The only difference in this case between L_{AB} and both L_{AA} and L_{BB} is the last phase factor $e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_\mu-\boldsymbol{x}_\nu)}$, which is $e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_\text{A}-\boldsymbol{x}_\text{B})}$ for L_{AB} and 1 for both L_{AA} and L_{BB}. Now, given that for a nonnegative function $f(t) \geq 0\,\forall t$,

$$\int_{-\infty}^{\infty}\mathrm{d}t\,f(t) \geq \left|\int_{-\infty}^{\infty}\mathrm{d}t\,e^{\mathrm{i}\omega t}f(t)\right| \tag{A.15}$$

it is easy to see that $L_{\text{AA}} \geq |L_{\text{AB}}|$ and $L_{\text{BB}} \geq |L_{\text{AB}}|$, thereby satisfying Eq. (A.11). Performing analogous calculations, the same conclusion can be reached for the scalar cases and arbitrary smearing functions.

A.3 Explicit calculation of $\mathcal{L}_{\mu\mu}^{\text{EM}}$ and $\mathcal{M}^{\text{EM}}$

In this appendix we perform a step-by-step derivation of Eqs. (2.42) and (2.43) starting from Eqs. (2.29) and (2.30). We begin without fixing the ground nor the excited states, and we will perform the calculations as general as possible, with the only assumption that both atoms have the same atomic structure (same ground and excited states). We will particularize to the $1s \to 2p_z$ transition only at the very end of each section of this appendix.

A.3.1 Local term

Let us begin with the complete expression of the local term

$$\begin{aligned}\mathcal{L}^{\text{EM}}_{\mu\mu} &= e_\mu^2 \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2\, e^{\mathrm{i}\Omega_\mu(t_1-t_2)} \mathcal{X}_\mu(t_1)\mathcal{X}_\mu(t_2) \\ &\quad \times \int \mathrm{d}^3\boldsymbol{x}_1' \int \mathrm{d}^3\boldsymbol{x}_2'\, \boldsymbol{F}_\mu^{*\mathbf{t}}(\boldsymbol{x}_2' - \boldsymbol{x}_\mu)\mathbf{W}(\boldsymbol{x}_2', \boldsymbol{x}_1', t_2, t_1)\boldsymbol{F}_\mu(\boldsymbol{x}_1' - \boldsymbol{x}_\mu) \\ &= e_\mu^2 \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2\, e^{\mathrm{i}\Omega_\mu(t_1-t_2)} \mathcal{X}_\mu(t_1)\mathcal{X}_\mu(t_2) \\ &\quad \times \int \mathrm{d}^3\boldsymbol{x}_1 \int \mathrm{d}^3\boldsymbol{x}_2\, \psi_e(\boldsymbol{x}_2)\psi_g^*(\boldsymbol{x}_2)\psi_e^*(\boldsymbol{x}_1)\psi_g(\boldsymbol{x}_1) \\ &\quad \times \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3} \frac{|\boldsymbol{k}|}{2} e^{-\mathrm{i}|\boldsymbol{k}|(t_2-t_1)} e^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}_2} e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}_1}\, \boldsymbol{x}_2^{\mathbf{t}} \left(\mathbb{1} - \frac{\boldsymbol{k}\otimes\boldsymbol{k}}{|\boldsymbol{k}|^2}\right) \boldsymbol{x}_1, \end{aligned} \tag{A.16}$$

where we have already written the smearing vector as $\boldsymbol{F}(\boldsymbol{x}) = \psi_e^*(\boldsymbol{x})\,\boldsymbol{x}\,\psi_g(\boldsymbol{x})$ as per Eq. (2.12), and the Wightman 2-tensor of the electric field is [note this is just Eq. (2.31)]

$$\begin{aligned}\mathbf{W}(\boldsymbol{x}_2, \boldsymbol{x}_1, t_2, t_1) &= \langle 0|\hat{\boldsymbol{E}}(\boldsymbol{x}_2, t_2) \otimes \hat{\boldsymbol{E}}(\boldsymbol{x}_1, t_1)|0\rangle \\ &= \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3} \frac{|\boldsymbol{k}|}{2} e^{-\mathrm{i}|\boldsymbol{k}|(t_2-t_1)} e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_2-\boldsymbol{x}_1)} \left(\mathbb{1} - \frac{\boldsymbol{k}\otimes\boldsymbol{k}}{|\boldsymbol{k}|^2}\right). \end{aligned} \tag{A.17}$$

First of all, we perform the change of variables $\boldsymbol{x}_1 = \boldsymbol{x}_1' - \boldsymbol{x}_\mu$, $\boldsymbol{x}_2 = \boldsymbol{x}_2' - \boldsymbol{x}_\mu$ to eliminate the explicit dependence of the smearing vector on the atomic position $\boldsymbol{x}_\mu$. After that, we choose spherical coordinates to perform the integrations and use the following decompositions involving spherical harmonics:

$$\psi_{nlm}(\boldsymbol{x}) = R_{nl}(|\boldsymbol{x}|)Y_{lm}(\hat{\boldsymbol{x}}), \tag{A.18}$$

$$\begin{aligned} e^{\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{y}} &= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} 4\pi \mathrm{i}^l j_l(|\boldsymbol{x}||\boldsymbol{y}|)Y_{lm}(\hat{\boldsymbol{x}})Y_{lm}^*(\hat{\boldsymbol{y}}) \\ &= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} 4\pi \mathrm{i}^l j_l(|\boldsymbol{x}||\boldsymbol{y}|)Y_{lm}^*(\hat{\boldsymbol{x}})Y_{lm}(\hat{\boldsymbol{y}}), \end{aligned} \tag{A.19}$$

$$\boldsymbol{x}\cdot\boldsymbol{y} = \frac{4\pi}{3}|\boldsymbol{x}||\boldsymbol{y}| \left[Y_{10}(\hat{\boldsymbol{x}})Y_{10}(\hat{\boldsymbol{y}}) - Y_{11}(\hat{\boldsymbol{x}})Y_{1-1}(\hat{\boldsymbol{y}}) - Y_{1-1}(\hat{\boldsymbol{x}})Y_{11}(\hat{\boldsymbol{y}})\right], \tag{A.20}$$

where the arguments of the spherical harmonics $\hat{\boldsymbol{x}} = (\theta_{\boldsymbol{x}}, \phi_{\boldsymbol{x}})$ are the azimuthal and polar coordinates of the unit vector $\hat{\boldsymbol{x}}$ and $R_{nl}(|\boldsymbol{x}|)$ are the radial hydrogenoid wave functions [GP92].

To make calculations less cumbersome, we separate Eq. (A.16) into two parts, one that contains the term with the identity matrix $\mathbb{1}$ and the other that contains the term with the momentum dyadic $\boldsymbol{k}\otimes\boldsymbol{k}$, and compute each of them separately.

Let us begin with the term containing the identity. Substituting Eqs. (A.18), (A.19) and (A.20) we get

$$\begin{aligned}
\mathcal{L}^{\mathrm{EM}}_{\mu\mu}\Big|_{\mathbb{1}} = e_\mu^2 \int_0^\infty \frac{\mathrm{d}|\boldsymbol{k}|}{(2\pi)^3}\frac{|\boldsymbol{k}|^3}{2}\sum_{l=0}^{\infty}\sum_{m=-l}^{l} 4\pi \mathrm{i}^l \sum_{l'=0}^{\infty}\sum_{m'=-l'}^{l'} 4\pi \mathrm{i}^{l'}(-1)^{l'}\frac{4\pi}{3} \\
\times \int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{\infty}\mathrm{d}t_2\, e^{\mathrm{i}\Omega_\mu(t_1-t_2)}\mathcal{X}_\mu(t_1)\mathcal{X}_\mu(t_2)e^{-\mathrm{i}|\boldsymbol{k}|(t_2-t_1)} \\
\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_2|\,|\boldsymbol{x}_2|^3 R_{n_e,l_e}(|\boldsymbol{x}_2|)R^*_{n_g,l_g}(|\boldsymbol{x}_2|)j_l(|\boldsymbol{k}||\boldsymbol{x}_2|) \\
\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_1|\,|\boldsymbol{x}_1|^3 R^*_{n_e,l_e}(|\boldsymbol{x}_1|)R_{n_g,l_g}(|\boldsymbol{x}_1|)j_{l'}(|\boldsymbol{k}||\boldsymbol{x}_1|) \\
\times \int \mathrm{d}\Omega_k\, Y_{lm}(\hat{\boldsymbol{k}})Y_{l'm'}(\hat{\boldsymbol{k}}) \\
\times \int \mathrm{d}\Omega_1 Y^*_{l_e,m_e}(\hat{\boldsymbol{x}}_1)Y_{l_g,m_g}(\hat{\boldsymbol{x}}_1)Y^*_{l'm'}(\hat{\boldsymbol{x}}_1)\int \mathrm{d}\Omega_2 Y_{l_e,m_e}(\hat{\boldsymbol{x}}_2)Y^*_{l_g,m_g}(\hat{\boldsymbol{x}}_2)Y^*_{lm}(\hat{\boldsymbol{x}}_2) \\
\times \left[Y_{10}(\hat{\boldsymbol{x}}_1)Y_{10}(\hat{\boldsymbol{x}}_2) - Y_{11}(\hat{\boldsymbol{x}}_1)Y_{1-1}(\hat{\boldsymbol{x}}_2) - Y_{1-1}(\hat{\boldsymbol{x}}_1)Y_{11}(\hat{\boldsymbol{x}}_2)\right], \qquad \text{(A.21)}
\end{aligned}$$

where the $(-1)^{l'}$ factor comes from the identity $Y_{lm}(-\hat{\boldsymbol{r}}) = (-1)^l Y_{lm}(\hat{\boldsymbol{r}})$ and $j_l(x)$ are the spherical Bessel functions.

Written in this form, almost each line of Eq. (A.21) can be computed separately. For instance, using the Gaussian switching function (2.36) the time integrals in the second line yield

$$\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{\infty}\mathrm{d}t_2\, e^{\mathrm{i}\Omega_\mu(t_1-t_2)}\mathcal{X}_\mu(t_1)\mathcal{X}_\mu(t_2)e^{-\mathrm{i}|\boldsymbol{k}|(t_2-t_1)} = \pi T^2 e^{-\frac{1}{2}T^2(\Omega_\mu+|\boldsymbol{k}|)^2}. \qquad \text{(A.22)}$$

The fourth line can be readily evaluated:

$$\int \mathrm{d}\Omega_k\, Y_{lm}(\hat{\boldsymbol{k}})Y_{l'm'}(\hat{\boldsymbol{k}}) = (-1)^{m'}\delta_{l,l'}\delta_{m,-m'}, \qquad \text{(A.23)}$$

directly from the orthogonality relations of spherical harmonics [the $(-1)^{m'}$ factor comes from the fact that $Y^*_{lm} = (-1)^m Y_{l-m}$]. The simple form that this integral on solid angle takes allows us to easily compute the sums in m' and l'. The two last lines can be computed using the following identity involving the

integral of four spherical harmonics over the unit sphere S^2,

$$\begin{aligned}
&\int \mathrm{d}\Omega\, Y^*_{l_1m_1}(\theta,\phi)Y_{l_2m_2}(\theta,\phi)Y^*_{l_3m_3}(\theta,\phi)Y_{l_4m_4}(\theta,\phi)\\
&\quad=\sum_{\lambda=0}^{\infty}\sum_{\mu=-\lambda}^{\lambda}\frac{2\lambda+1}{4\pi}\sqrt{(2l_1+1)(2l_2+1)(2l_3+1)(2l_4+1)}\\
&\qquad\times\begin{pmatrix}l_1 & l_3 & \lambda\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}l_1 & l_3 & \lambda\\ -m_1 & -m_3 & -\mu\end{pmatrix}\begin{pmatrix}l_2 & l_4 & \lambda\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}l_2 & l_4 & \lambda\\ m_2 & m_4 & \mu\end{pmatrix}, \qquad \text{(A.24)}
\end{aligned}$$

where $\mathrm{d}\Omega = \mathrm{d}(\cos\theta)\mathrm{d}\phi$ and $\begin{pmatrix}l_1 & l_2 & l_3\\ m_1 & m_2 & m_3\end{pmatrix}$ represents the Wigner $3j$-symbols.

Using Eq. (A.24), the sums over l', m and m' of all the integrals over solid angles—the last three lines of Eq. (A.21)—yield

$$\begin{aligned}
&\sum_{l'}\sum_{m}\sum_{m'}\mathrm{i}^{l+l'}(-1)^{l'}j_{l'}(|\boldsymbol{k}||\boldsymbol{x}_1|)\int \mathrm{d}\Omega_k\, Y_{lm}(\hat{\boldsymbol{k}})Y_{l'm'}(\hat{\boldsymbol{k}})\\
&\quad\times\int \mathrm{d}\Omega_1 Y^*_{l_e,m_e}(\hat{\boldsymbol{x}}_1)Y_{l_g,m_g}(\hat{\boldsymbol{x}}_1)Y^*_{l'm'}(\hat{\boldsymbol{x}}_1)\int \mathrm{d}\Omega_2 Y_{l_e,m_e}(\hat{\boldsymbol{x}}_2)Y^*_{l_g,m_g}(\hat{\boldsymbol{x}}_2)Y^*_{lm}(\hat{\boldsymbol{x}}_2)\\
&\quad\times[Y_{10}(\hat{\boldsymbol{x}}_1)Y_{10}(\hat{\boldsymbol{x}}_2)-Y_{11}(\hat{\boldsymbol{x}}_1)Y_{1-1}(\hat{\boldsymbol{x}}_2)-Y_{1-1}(\hat{\boldsymbol{x}}_1)Y_{11}(\hat{\boldsymbol{x}}_2)]\\
&=\frac{3(-1)^{m_g-m_e}\mathrm{i}^{2l}(-1)^l}{(4\pi)^2}(2l+1)(2l_e+1)(2l_g+1)j_l(|\boldsymbol{k}||\boldsymbol{x}_1|)\\
&\quad\times\sum_{\lambda,\lambda'}(2\lambda+1)(2\lambda'+1)\begin{pmatrix}l & l_g & \lambda\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}l_e & 1 & \lambda\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}l & l_e & \lambda'\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}l_g & 1 & \lambda'\\ 0 & 0 & 0\end{pmatrix}\\
&\quad\times\Bigg[\begin{pmatrix}l_e & 1 & \lambda\\ m_e & 0 & -m_e\end{pmatrix}\begin{pmatrix}l & l_g & \lambda\\ m_g-m_e & -m_g & m_e\end{pmatrix}\\
&\qquad\times\begin{pmatrix}l_g & 1 & \lambda'\\ m_g & 0 & -m_g\end{pmatrix}\begin{pmatrix}l & l_e & \lambda'\\ m_e-m_g & -m_e & m_g\end{pmatrix}\\
&\qquad+\begin{pmatrix}l_e & 1 & \lambda\\ m_e & -1 & 1-m_e\end{pmatrix}\begin{pmatrix}l & l_g & \lambda\\ 1+m_g-m_e & -m_g & m_e-1\end{pmatrix}\\
&\qquad\times\begin{pmatrix}l_g & 1 & \lambda'\\ m_g & 1 & -1-m_g\end{pmatrix}\begin{pmatrix}l & l_e & \lambda'\\ m_e-m_g-1 & -m_e & 1+m_g\end{pmatrix}\\
&\qquad+\begin{pmatrix}l_e & 1 & \lambda\\ m_e & 1 & -1-m_e\end{pmatrix}\begin{pmatrix}l & l_g & \lambda\\ m_g-m_e-1 & -m_g & m_e+1\end{pmatrix}\\
&\qquad\times\begin{pmatrix}l_g & 1 & \lambda'\\ m_g & -1 & 1-m_g\end{pmatrix}\begin{pmatrix}l & l_e & \lambda'\\ 1+m_e-m_g & -m_e & m_g-1\end{pmatrix}\Bigg]. \qquad \text{(A.25)}
\end{aligned}$$

After substituting all these in Eq. (A.21) we obtain

$$
\begin{aligned}
\mathcal{L}^{\text{EM}}_{\mu\mu}\Big|_{\mathbb{1}} = e_\mu^2 \int_0^\infty \frac{\mathrm{d}|\boldsymbol{k}|}{(2\pi)^3}\frac{|\boldsymbol{k}|^3}{2}\sum_{l=0}^{\infty}(4\pi)^2 \mathrm{i}^{2l}(-1)^l \frac{4\pi}{3}\pi T^2 e^{-\frac{1}{2}T^2(\Omega_\mu+|\boldsymbol{k}|)^2} \\
\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_2|\,|\boldsymbol{x}_2|^3 R_{n_e,l_e}(|\boldsymbol{x}_2|)R^*_{n_g,l_g}(|\boldsymbol{x}_2|)j_l(|\boldsymbol{k}||\boldsymbol{x}_2|) \\
\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_1|\,|\boldsymbol{x}_1|^3 R^*_{n_e,l_e}(|\boldsymbol{x}_1|)R_{n_g,l_g}(|\boldsymbol{x}_1|)j_l(|\boldsymbol{k}||\boldsymbol{x}_1|) \\
\times \frac{3(-1)^{m_g-m_e}}{(4\pi)^2}(2l+1)(2l_e+1)(2l_g+1) \\
\times \sum_{\lambda,\lambda'}(2\lambda+1)(2\lambda'+1)\begin{pmatrix} l & l_g & \lambda \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l_e & 1 & \lambda \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & l_e & \lambda' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l_g & 1 & \lambda' \\ 0 & 0 & 0\end{pmatrix} \\
\times \Bigg[\begin{pmatrix} l_e & 1 & \lambda \\ m_e & 0 & -m_e\end{pmatrix}\begin{pmatrix} l & l_g & \lambda \\ m_g-m_e & -m_g & m_e\end{pmatrix} \\
\quad \times \begin{pmatrix} l_g & 1 & \lambda' \\ m_g & 0 & -m_g\end{pmatrix}\begin{pmatrix} l & l_e & \lambda' \\ m_e-m_g & -m_e & m_g\end{pmatrix} \\
+ \begin{pmatrix} l_e & 1 & \lambda \\ m_e & -1 & 1-m_e\end{pmatrix}\begin{pmatrix} l & l_g & \lambda \\ 1+m_g-m_e & -m_g & m_e-1\end{pmatrix} \\
\quad \times \begin{pmatrix} l_g & 1 & \lambda' \\ m_g & 1 & -1-m_g\end{pmatrix}\begin{pmatrix} l & l_e & \lambda' \\ m_e-m_g-1 & -m_e & 1+m_g\end{pmatrix} \\
+ \begin{pmatrix} l_e & 1 & \lambda \\ m_e & 1 & -1-m_e\end{pmatrix}\begin{pmatrix} l & l_g & \lambda \\ m_g-m_e-1 & -m_g & m_e+1\end{pmatrix} \\
\quad \times \begin{pmatrix} l_g & 1 & \lambda' \\ m_g & -1 & 1-m_g\end{pmatrix}\begin{pmatrix} l & l_e & \lambda' \\ 1+m_e-m_g & -m_e & m_g-1\end{pmatrix}\Bigg]. \qquad \text{(A.26)}
\end{aligned}
$$

This expression is fully general, for any two arbitrary levels of the hydrogenoid atom coupled dipolarly to the vacuum state of the electromagnetic field. No more integrations can be performed unless we specify which particular atomic electron states are the ground and excited states.

Before doing that, let us compute the contribution to the local term (A.16) that is proportional to the momentum dyadic $\boldsymbol{k}\otimes\boldsymbol{k}$. The strategy we will follow will be the same as in the previous case. After substituting in Eqs. (A.18),

(A.19) and (A.20) the contribution reads

$$
\begin{aligned}
\mathcal{L}^{\text{EM}}_{\mu\mu}\Big|_{\boldsymbol{k}\otimes\boldsymbol{k}} = e_\mu^2 \int_0^\infty \frac{\mathrm{d}|\boldsymbol{k}|}{(2\pi)^3}\frac{|\boldsymbol{k}|^3}{2}\sum_{l=0}^{\infty}\sum_{m=-l}^{l} 4\pi \mathrm{i}^l \sum_{l'=0}^{\infty}\sum_{m'=-l'}^{l'} 4\pi \mathrm{i}^{l'}(-1)^{l'}\left(\frac{4\pi}{3}\right)^2 \\
\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_2|\,|\boldsymbol{x}_2|^3 R_{n_e,l_e}(|\boldsymbol{x}_2|)R^*_{n_g,l_g}(|\boldsymbol{x}_2|)j_l(|\boldsymbol{k}||\boldsymbol{x}_2|) \\
\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_1|\,|\boldsymbol{x}_1|^3 R^*_{n_e,l_e}(|\boldsymbol{x}_1|)R_{n_g,l_g}(|\boldsymbol{x}_1|)j_{l'}(|\boldsymbol{k}||\boldsymbol{x}_1|) \\
\times \pi T^2 e^{-\frac{1}{2}T^2(\Omega_\mu+|\boldsymbol{k}|)^2}\int \mathrm{d}\Omega_1 Y^*_{l_e,m_e}(\hat{\boldsymbol{x}}_1)Y_{l_g,m_g}(\hat{\boldsymbol{x}}_1)Y^*_{l'm'}(\hat{\boldsymbol{x}}_1) \\
\times \int \mathrm{d}\Omega_2 Y_{l_e,m_e}(\hat{\boldsymbol{x}}_2)Y^*_{l_g,m_g}(\hat{\boldsymbol{x}}_2)Y^*_{lm}(\hat{\boldsymbol{x}}_2)\int \mathrm{d}\Omega_k\, Y_{lm}(\hat{\boldsymbol{k}})Y_{l'm'}(\hat{\boldsymbol{k}}) \\
\times \Big[Y_{10}(\hat{\boldsymbol{x}}_1)Y_{10}(\hat{\boldsymbol{k}}) - Y_{11}(\hat{\boldsymbol{x}}_1)Y_{1-1}(\hat{\boldsymbol{k}}) - Y_{1-1}(\hat{\boldsymbol{x}}_1)Y_{11}(\hat{\boldsymbol{k}})\Big] \\
\times \Big[Y_{10}(\hat{\boldsymbol{k}})Y_{10}(\hat{\boldsymbol{x}}_2) - Y_{11}(\hat{\boldsymbol{k}})Y_{1-1}(\hat{\boldsymbol{x}}_2) - Y_{1-1}(\hat{\boldsymbol{k}})Y_{11}(\hat{\boldsymbol{x}}_2)\Big]\,, \qquad \text{(A.27)}
\end{aligned}
$$

where we have already performed the integrals in time.

The novelty introduced by the dyadic is that the integral in the solid angle of $\boldsymbol{k}$ is now no longer trivial, although it can be readily performed using Eq. (A.24). Integrating over all solid angles we obtain

$$
\begin{aligned}
&\sum_{m,m'}\int \mathrm{d}\Omega_1 Y^*_{l_e,m_e}(\hat{\boldsymbol{x}}_1)Y_{l_g,m_g}(\hat{\boldsymbol{x}}_1)Y^*_{l'm'}(\hat{\boldsymbol{x}}_1) \\
&\quad\times \int \mathrm{d}\Omega_2 Y_{l_e,m_e}(\hat{\boldsymbol{x}}_2)Y^*_{l_g,m_g}(\hat{\boldsymbol{x}}_2)Y^*_{lm}(\hat{\boldsymbol{x}}_2)\int \mathrm{d}\Omega_k\, Y_{lm}(\hat{\boldsymbol{k}})Y_{l'm'}(\hat{\boldsymbol{k}}) \\
&\quad\times \Big[Y_{10}(\hat{\boldsymbol{x}}_1)Y_{10}(\hat{\boldsymbol{k}}) - Y_{11}(\hat{\boldsymbol{x}}_1)Y_{1-1}(\hat{\boldsymbol{k}}) - Y_{1-1}(\hat{\boldsymbol{x}}_1)Y_{11}(\hat{\boldsymbol{k}})\Big] \\
&\quad\times \Big[Y_{10}(\hat{\boldsymbol{k}})Y_{10}(\hat{\boldsymbol{x}}_2) - Y_{11}(\hat{\boldsymbol{k}})Y_{1-1}(\hat{\boldsymbol{x}}_2) - Y_{1-1}(\hat{\boldsymbol{k}})Y_{11}(\hat{\boldsymbol{x}}_2)\Big] \\
&= 9(2l+1)(2l'+1)(2l_g+1)(2l_e+1) \\
&\quad\times\sum_{\lambda'\lambda''}\frac{(2\lambda'+1)(2\lambda''+1)}{(4\pi)^3}\begin{pmatrix} l_e & 1 & \lambda' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l_g & 1 & \lambda'' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l' & l_e & \lambda'' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & l_g & \lambda' \\ 0 & 0 & 0\end{pmatrix} \\
&\quad\times (A_{\mathcal{L}} + B_{\mathcal{L}})\,, \qquad \text{(A.28)}
\end{aligned}
$$

where the two quantities $A_{\mathcal{L}}$ and $B_{\mathcal{L}}$ read

$$
\begin{aligned}
A_{\mathcal{L}} =& \sum_{\lambda} (2\lambda+1) \begin{pmatrix} 1 & 1 & \lambda \\ 0 & 0 & 0 \end{pmatrix}^2 \begin{pmatrix} l & l' & \lambda \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & l_g & \lambda' \\ m_g - m_e & -m_g & m_e \end{pmatrix} \begin{pmatrix} l_e & 1 & \lambda' \\ m_e & 0 & -m_e \end{pmatrix} \\
& \times \begin{pmatrix} l & l' & \lambda \\ m_e - m_g & m_g - m_e & 0 \end{pmatrix} \begin{pmatrix} l' & l_e & \lambda'' \\ m_e - m_g & -m_e & m_g \end{pmatrix} \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & 0 & -m_g \end{pmatrix} \\
& + \sum_{\lambda} (2\lambda+1) \begin{pmatrix} 1 & 1 & \lambda \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & \lambda \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} l & l' & \lambda \\ 0 & 0 & 0 \end{pmatrix} \\
& \times \Bigg[\begin{pmatrix} l & l' & \lambda \\ m_e - m_g - 1 & m_g - m_e & 1 \end{pmatrix} \begin{pmatrix} l' & l_e & \lambda'' \\ m_e - m_g & -m_e & m_g \end{pmatrix} \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & 0 & -m_g \end{pmatrix} \\
& \quad \times \begin{pmatrix} l & l_g & \lambda' \\ m_g - m_e + 1 & -m_g & m_e - 1 \end{pmatrix} \begin{pmatrix} l_e & 1 & \lambda' \\ m_e & -1 & 1 - m_e \end{pmatrix} \\
& \quad + \begin{pmatrix} l & l' & \lambda \\ m_e - m_g & m_g - m_e - 1 & 1 \end{pmatrix} \begin{pmatrix} l' & l_e & \lambda'' \\ m_e - m_g + 1 & -m_e & m_g - 1 \end{pmatrix} \\
& \quad \times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & -1 & 1 - m_g \end{pmatrix} \begin{pmatrix} l & l_g & \lambda' \\ m_g - m_e & -m_g & m_e \end{pmatrix} \begin{pmatrix} l_e & 1 & \lambda' \\ m_e & 0 & -m_e \end{pmatrix} \Bigg] \\
& + \sum_{\lambda} (2\lambda+1) \begin{pmatrix} 1 & 1 & \lambda \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & \lambda \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} l & l' & \lambda \\ 0 & 0 & 0 \end{pmatrix} \\
& \times \Bigg[\begin{pmatrix} l & l' & \lambda \\ m_e - m_g + 1 & m_g - m_e & -1 \end{pmatrix} \begin{pmatrix} l' & l_e & \lambda'' \\ m_e - m_g & -m_e & m_g \end{pmatrix} \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & 0 & -m_g \end{pmatrix} \\
& \quad \times \begin{pmatrix} l & l_g & \lambda' \\ m_g - m_e - 1 & -m_g & m_e + 1 \end{pmatrix} \begin{pmatrix} l_e & 1 & \lambda' \\ m_e & 1 & -1 - m_e \end{pmatrix} \\
& \quad + \begin{pmatrix} l & l' & \lambda \\ m_e - m_g & m_g - m_e + 1 & -1 \end{pmatrix} \begin{pmatrix} l' & l_e & \lambda'' \\ m_e - m_g - 1 & -m_e & m_g + 1 \end{pmatrix} \\
& \quad \times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & 1 & -1 - m_g \end{pmatrix} \begin{pmatrix} l & l_g & \lambda' \\ m_g - m_e & -m_g & m_e \end{pmatrix} \begin{pmatrix} l_e & 1 & \lambda' \\ m_e & 0 & -m_e \end{pmatrix} \Bigg] \\
& + \sum_{\lambda} (2\lambda+1) \begin{pmatrix} 1 & 1 & \lambda \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & \lambda \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} l & l' & \lambda \\ 0 & 0 & 0 \end{pmatrix} \\
& \times \Bigg[\begin{pmatrix} l & l' & \lambda \\ m_e - m_g - 1 & m_g - m_e + 1 & 0 \end{pmatrix} \begin{pmatrix} l' & l_e & \lambda'' \\ m_e - m_g - 1 & -m_e & m_g + 1 \end{pmatrix} \\
& \quad \times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & 1 & -1 - m_g \end{pmatrix} \begin{pmatrix} l & l_g & \lambda' \\ m_g - m_e + 1 & -m_g & m_e - 1 \end{pmatrix} \begin{pmatrix} l_e & 1 & \lambda' \\ m_e & -1 & 1 - m_e \end{pmatrix} \\
& \quad + \begin{pmatrix} l & l' & \lambda \\ m_e - m_g + 1 & m_g - m_e - 1 & 0 \end{pmatrix} \begin{pmatrix} l' & l_e & \lambda'' \\ m_e - m_g + 1 & -m_e & m_g - 1 \end{pmatrix} \\
& \quad \times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & -1 & 1 - m_g \end{pmatrix} \begin{pmatrix} l & l_g & \lambda' \\ m_g - m_e - 1 & -m_g & m_e + 1 \end{pmatrix} \begin{pmatrix} l_e & 1 & \lambda' \\ m_e & 1 & -1 - m_e \end{pmatrix} \Bigg],
\end{aligned}
\tag{A.29}
$$

$$
\begin{aligned}
B_{\mathcal{L}} = &\sqrt{\frac{2}{3}}\begin{pmatrix} l & l' & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l & l' & 2 \\ m_e-m_g-1 & m_g-m_e-1 & 2 \end{pmatrix}\begin{pmatrix} l' & l_e & \lambda'' \\ 1+m_e-m_g & -m_e & m_g-1 \end{pmatrix} \\
&\times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & -1 & 1-m_g \end{pmatrix}\begin{pmatrix} l_e & 1 & \lambda' \\ m_e & -1 & 1-m_e \end{pmatrix}\begin{pmatrix} l & l_g & \lambda' \\ 1-m_e+m_g & -m_g & m_e-1 \end{pmatrix} \\
&+\sqrt{\frac{2}{3}}\begin{pmatrix} l & l' & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l & l' & 2 \\ m_e-m_g+1 & m_g-m_e+1 & -2 \end{pmatrix}\begin{pmatrix} l' & l_e & \lambda'' \\ m_e-m_g-1 & -m_e & m_g+1 \end{pmatrix} \\
&\times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & 1 & -1-m_g \end{pmatrix}\begin{pmatrix} l & l_g & \lambda' \\ m_g-m_e-1 & -m_g & m_e+1 \end{pmatrix}\begin{pmatrix} l_e & 1 & \lambda' \\ m_e & 1 & -1-m_e \end{pmatrix}.
\end{aligned} \tag{A.30}
$$

Now we particularize the local term $\mathcal{L}_{\mu\mu}$ to the atomic transition studied in the main text. Namely, we consider the ground state of both detectors to be the hydrogenoid $1s$ orbital, and the excited state a hydrogenoid $2p_z$ orbital. Therefore, we have $l_e = 1$, $l_g = 0$, $m_e = 0$ and $m_g = 0$. In this scenario, the identity term (A.21) reads

$$
\begin{aligned}
\mathcal{L}^{\text{EM}}_{\mu\mu}\Big|_{\mathbb{1}} = &\, e_\mu^2 \int_0^\infty \frac{\mathrm{d}|\boldsymbol{k}|}{(2\pi)^3}\frac{|\boldsymbol{k}|^3}{2}\sum_{l=0}^{\infty}(4\pi)^2\frac{4\pi}{3}\pi T^2 e^{-\frac{1}{2}T^2(\Omega_\mu+|\boldsymbol{k}|)^2} \\
&\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_2|\,|\boldsymbol{x}_2|^3 R_{2,1}(|\boldsymbol{x}_2|)R^*_{1,0}(|\boldsymbol{x}_2|)j_l(|\boldsymbol{k}||\boldsymbol{x}_2|) \\
&\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_1|\,|\boldsymbol{x}_1|^3 R^*_{2,1}(|\boldsymbol{x}_1|)R_{1,0}(|\boldsymbol{x}_1|)j_l(|\boldsymbol{k}||\boldsymbol{x}_1|) \\
&\times \frac{3}{(4\pi)^2}(2l+1)(2+1)(0+1) \\
&\times \sum_{\lambda,\lambda'}(2\lambda+1)(2\lambda'+1)\begin{pmatrix} l & 0 & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l & 1 & \lambda' \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & \lambda' \\ 0 & 0 & 0 \end{pmatrix} \\
&\times \left[\begin{pmatrix} 1 & 1 & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l & 0 & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & \lambda' \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l & 1 & \lambda' \\ 0 & 0 & 0 \end{pmatrix}\right. \\
&\quad + \begin{pmatrix} 1 & 1 & \lambda \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} l & 0 & \lambda \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 & \lambda' \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} l & 1 & \lambda' \\ -1 & 0 & 1 \end{pmatrix} \\
&\quad \left. + \begin{pmatrix} 1 & 1 & \lambda \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} l & 0 & \lambda \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & \lambda' \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} l & 1 & \lambda' \\ 1 & 0 & -1 \end{pmatrix}\right].
\end{aligned} \tag{A.31}
$$

Using the properties of the Wigner $3j$-symbols it is clear that the only nonzero contributions to the sum are $\lambda = 0, 1, 2$ and $\lambda' = 1$. Additionally, the first $3j$-symbol enforces that the nonzero terms in the sum are only those satisfying

$l=\lambda$. Performing the sums, the expression is simplified to

$$\begin{aligned}\mathcal{L}^{\text{EM}}_{\mu\mu}\Big|_{\mathbb{1}} =&\frac{e_\mu^2}{12\pi}T^2\int_0^\infty \mathrm{d}|\boldsymbol{k}|\,|\boldsymbol{k}|^3 e^{-\frac{1}{2}T^2(\Omega_\mu+|\boldsymbol{k}|)^2}\\ &\times\int_0^\infty \mathrm{d}|\boldsymbol{x}_2|\,|\boldsymbol{x}_2|^3 R_{2,1}(|\boldsymbol{x}_2|)R^*_{1,0}(|\boldsymbol{x}_2|)\int_0^\infty \mathrm{d}|\boldsymbol{x}_1|\,|\boldsymbol{x}_1|^3 R^*_{2,1}(|\boldsymbol{x}_1|)R_{1,0}(|\boldsymbol{x}_1|)\\ &\times\left[j_0(|\boldsymbol{k}||\boldsymbol{x}_1|)j_0(|\boldsymbol{k}||\boldsymbol{x}_2|)+2j_2(|\boldsymbol{k}||\boldsymbol{x}_1|)j_2(|\boldsymbol{k}||\boldsymbol{x}_2|)\right].\end{aligned} \tag{A.32}$$

The integrals over $\mathrm{d}|\boldsymbol{x}_1|$ and $\mathrm{d}|\boldsymbol{x}_2|$ can now be readily evaluated using the identity

$$\begin{aligned}&\int_0^\infty \mathrm{d}r\, r^3 R_{2,1}(r)R_{1,0}(r)j_l(|\boldsymbol{k}|r)\\ &\qquad=8\sqrt{2\pi}3^{-l-\frac{11}{2}}a_0^{l+1}|\boldsymbol{k}|^l\Gamma(l+5)\,{}_2\tilde{F}_1\left(\frac{l+5}{2},\frac{l+6}{2};l+\frac{3}{2};-\frac{4}{9}a_0^2|\boldsymbol{k}|^2\right),\end{aligned} \tag{A.33}$$

where ${}_2\tilde{F}_1\,(a,b;c;z)={}_2F_1(a,b;c;z)/\Gamma(z)$ is the regularized hypergeometric function.

After particularizing Eq. (A.33) to the cases appearing in Eq. (A.32), we arrive to the final expression for the identity contribution to the local term

$$\mathcal{L}^{\text{EM}}_{\mu\mu}\Big|_{\mathbb{1}} =e_\mu^2\frac{663552}{\pi}a_0^2T^2\int_0^\infty \mathrm{d}|\boldsymbol{k}|\,|\boldsymbol{k}|^3 e^{-\frac{1}{2}T^2(\Omega_\mu+|\boldsymbol{k}|)^2}\frac{16a_0^4|\boldsymbol{k}|^4-8a_0^2|\boldsymbol{k}|^2+9}{\left(4a_0^2|\boldsymbol{k}|^2+9\right)^8}. \tag{A.34}$$

Next, we apply the same procedure and techniques to the term with the

dyadic contribution in momenta, which yields

$$
\begin{aligned}
\mathcal{L}^{\mathrm{EM}}_{\mu\mu}\big|_{\boldsymbol{k}\otimes\boldsymbol{k}} &= e_\mu^2 \int_0^\infty \frac{\mathrm{d}|\boldsymbol{k}|}{(2\pi)^3}\,\frac{|\boldsymbol{k}|^3}{2}\sum_{l=0}^{\infty} 4\pi \mathrm{i}^l \sum_{l'=0}^{\infty} 4\pi \mathrm{i}^{l'}(-1)^{l'}\left(\frac{4\pi}{3}\right)^2 \pi T^2 e^{-\frac{1}{2}T^2(\Omega_\mu+|\boldsymbol{k}|)^2}\,\frac{9(2l+1)(2l'+1)\cdot 1\cdot 3}{(4\pi)^3} \\
&\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_2|\,|\boldsymbol{x}_2|^3 R_{2,1}(|\boldsymbol{x}_2|)R^*_{1,0}(|\boldsymbol{x}_2|)j_l(|\boldsymbol{k}||\boldsymbol{x}_2|)\int_0^\infty \mathrm{d}|\boldsymbol{x}_1|\,|\boldsymbol{x}_1|^3 R^*_{2,1}(|\boldsymbol{x}_1|)R_{1,0}(|\boldsymbol{x}_1|)j_{l'}(|\boldsymbol{k}||\boldsymbol{x}_1|) \\
&\times \Bigg\{ (2l+1)(2+1)\sum_\lambda (2\lambda+1)\begin{pmatrix}1&1&\lambda\\0&0&0\end{pmatrix}^2\begin{pmatrix}l&l'&\lambda\\0&0&0\end{pmatrix}^2\begin{pmatrix}l&0&l\\0&0&0\end{pmatrix}^2\begin{pmatrix}1&1&l\\0&0&0\end{pmatrix}^2\begin{pmatrix}l'&1&1\\0&0&0\end{pmatrix}^2\begin{pmatrix}0&1&1\\0&0&0\end{pmatrix}^2 \\
&\quad + \sqrt{\frac{2}{3}}(2+1)(2l+1)\begin{pmatrix}l&l'&2\\0&0&0\end{pmatrix}\begin{pmatrix}l&l'&2\\-1&-1&2\end{pmatrix}\begin{pmatrix}l'&1&1\\0&0&0\end{pmatrix}\begin{pmatrix}l'&1&1\\1&0&-1\end{pmatrix}\begin{pmatrix}0&1&1\\0&0&0\end{pmatrix} \\
&\qquad \times \begin{pmatrix}0&1&1\\0&-1&1\end{pmatrix}\begin{pmatrix}l&0&l\\0&0&0\end{pmatrix}\begin{pmatrix}l&0&l\\1&0&-1\end{pmatrix}\begin{pmatrix}1&1&l\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&l\\0&-1&1\end{pmatrix} \\
&\quad + \sqrt{\frac{2}{3}}(2+1)(2l+1)\begin{pmatrix}l&l'&2\\0&0&0\end{pmatrix}\begin{pmatrix}l&l'&2\\1&1&-2\end{pmatrix}\begin{pmatrix}l'&1&1\\0&0&0\end{pmatrix}\begin{pmatrix}l'&1&1\\-1&0&1\end{pmatrix}\begin{pmatrix}0&1&1\\0&0&0\end{pmatrix} \\
&\qquad \times \begin{pmatrix}0&1&1\\0&1&-1\end{pmatrix}\begin{pmatrix}l&0&l\\0&0&0\end{pmatrix}\begin{pmatrix}l&0&l\\-1&0&1\end{pmatrix}\begin{pmatrix}1&1&l\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&l\\0&1&-1\end{pmatrix} \\
&\quad + \sum_\lambda (2\lambda+1)(2l+1)(2+1)\begin{pmatrix}1&1&\lambda\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&\lambda\\0&1&-1\end{pmatrix}\begin{pmatrix}l&l'&\lambda\\0&0&0\end{pmatrix} \\
&\qquad \times \begin{pmatrix}l&0&l\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&l\\0&0&0\end{pmatrix}\begin{pmatrix}0&1&1\\0&0&0\end{pmatrix}\begin{pmatrix}l'&1&1\\0&0&0\end{pmatrix} \\
&\qquad\quad \times \Bigg[\begin{pmatrix}l&l'&\lambda\\-1&0&1\end{pmatrix}\begin{pmatrix}l'&1&1\\0&0&0\end{pmatrix}\begin{pmatrix}0&1&1\\0&0&0\end{pmatrix}\begin{pmatrix}l&0&l\\1&0&-1\end{pmatrix}\begin{pmatrix}1&1&l\\0&-1&1\end{pmatrix} \\
&\qquad\qquad + \begin{pmatrix}l&l'&\lambda\\0&-1&1\end{pmatrix}\begin{pmatrix}l'&1&1\\1&0&-1\end{pmatrix}\begin{pmatrix}0&1&1\\0&-1&1\end{pmatrix}\begin{pmatrix}l&0&l\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&l\\0&0&0\end{pmatrix}\Bigg] \\
&\quad + \sum_\lambda (2\lambda+1)(2l+1)(2+1)\begin{pmatrix}1&1&\lambda\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&\lambda\\0&-1&1\end{pmatrix}\begin{pmatrix}l&l'&\lambda\\0&0&0\end{pmatrix} \\
&\qquad \times \begin{pmatrix}l&0&l\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&l\\0&0&0\end{pmatrix}\begin{pmatrix}0&1&1\\0&0&0\end{pmatrix}\begin{pmatrix}l'&1&1\\0&0&0\end{pmatrix} \\
&\qquad\quad \times \Bigg[\begin{pmatrix}l&l'&\lambda\\1&0&-1\end{pmatrix}\begin{pmatrix}l'&1&1\\0&0&0\end{pmatrix}\begin{pmatrix}0&1&1\\0&0&0\end{pmatrix}\begin{pmatrix}l&0&l\\-1&0&1\end{pmatrix}\begin{pmatrix}1&1&l\\0&1&-1\end{pmatrix} \\
&\qquad\qquad + \begin{pmatrix}l&l'&\lambda\\0&1&-1\end{pmatrix}\begin{pmatrix}l'&1&1\\-1&0&1\end{pmatrix}\begin{pmatrix}0&1&1\\0&1&-1\end{pmatrix}\begin{pmatrix}l&0&l\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&l\\0&0&0\end{pmatrix}\Bigg] \\
&\quad + \sum_\lambda (2\lambda+1)(2l+1)(2+1)\begin{pmatrix}1&1&\lambda\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&\lambda\\1&-1&0\end{pmatrix}\begin{pmatrix}l&l'&\lambda\\0&0&0\end{pmatrix} \\
&\qquad \times \begin{pmatrix}l&0&l\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&l\\0&0&0\end{pmatrix}\begin{pmatrix}0&1&1\\0&0&0\end{pmatrix}\begin{pmatrix}l'&1&1\\0&0&0\end{pmatrix} \\
&\qquad\quad \times \Bigg[\begin{pmatrix}l&l'&\lambda\\-1&1&0\end{pmatrix}\begin{pmatrix}l'&1&1\\-1&0&1\end{pmatrix}\begin{pmatrix}0&1&1\\0&1&-1\end{pmatrix}\begin{pmatrix}l&0&l\\1&0&-1\end{pmatrix}\begin{pmatrix}1&1&l\\0&-1&1\end{pmatrix} \\
&\qquad\qquad + \begin{pmatrix}l&l'&\lambda\\1&-1&0\end{pmatrix}\begin{pmatrix}l'&1&1\\1&0&-1\end{pmatrix}\begin{pmatrix}0&1&1\\0&-1&1\end{pmatrix}\begin{pmatrix}l&0&l\\-1&0&1\end{pmatrix}\begin{pmatrix}1&1&l\\0&1&-1\end{pmatrix}\Bigg]\Bigg\}.
\end{aligned}
\tag{A.35}
$$

The properties of the Wigner $3j$-symbols can be used to cancel many of the summands. In particular, we can restrict the sums to just $\lambda, l, l' = 0, 1, 2$, leading to the expression

$$\begin{aligned}\mathcal{L}^{\mathrm{EM}}_{\mu\mu}\Big|_{\boldsymbol{k}\otimes\boldsymbol{k}} &= e_\mu^2 \frac{T^2}{36\pi} \int_0^\infty \mathrm{d}|\boldsymbol{k}|\, |\boldsymbol{k}|^3 e^{-\frac{1}{2}T^2(\Omega_\mu+|\boldsymbol{k}|)^2} \\ &\quad\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_1|\, |\boldsymbol{x}_1|^3 R^*_{2,1}(|\boldsymbol{x}_1|) R_{1,0}(|\boldsymbol{x}_1|) \left[j_0(|\boldsymbol{k}||\boldsymbol{x}_1|) - 2j_2(|\boldsymbol{k}||\boldsymbol{x}_1|)\right] \\ &\quad\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_2|\, |\boldsymbol{x}_2|^3 R_{2,1}(|\boldsymbol{x}_2|) R^*_{1,0}(|\boldsymbol{x}_2|) \left[j_0(|\boldsymbol{k}||\boldsymbol{x}_2|) - 2j_2(|\boldsymbol{k}||\boldsymbol{x}_2|)\right] \\ &= e_\mu^2 \frac{24576}{\pi} T^2 a_0^2 \int_0^\infty \mathrm{d}|\boldsymbol{k}|\, |\boldsymbol{k}|^3 e^{-\frac{1}{2}T^2(\Omega_\mu+|\boldsymbol{k}|)^2} \frac{\left(20a_0^2|\boldsymbol{k}|^2 - 9\right)^2}{\left(4a_0^2|\boldsymbol{k}|^2 + 9\right)^8}. \end{aligned} \tag{A.36}$$

Finally, subtracting Eq. (A.36) from Eq. (A.34)—recall Eq. (A.16)—one arrives to the expression of the local term Eq. (2.42),

$$\mathcal{L}_{\mu\mu} = e^2 \frac{49152}{\pi} T^2 a_0^2 \int_0^\infty \mathrm{d}|\boldsymbol{k}| \frac{|\boldsymbol{k}|^3 e^{-\frac{1}{2}T^2(\Omega_\mu+|\boldsymbol{k}|)^2}}{\left(4a_0^2|\boldsymbol{k}|^2 + 9\right)^6}. \tag{A.37}$$

A.3.2 Nonlocal term

The nonlocal term (2.43) contains two different summands, which differ on the order of the subindices A and B. For each of them, as in the case of the local term, one can perform a separation into one part containing the identity and another one containing the momentum dyadic. In this part of the appendix we will compute explicitly the first summand of the term, which we will call $\mathcal{M}^{\mathrm{AB}}$, and derive the other from this one using symmetry arguments. The explicit expression for $\mathcal{M}^{\mathrm{AB}}$ is

$$\begin{aligned}\mathcal{M}^{\mathrm{AB}} &= -\, e_{\mathrm{A}} e_{\mathrm{B}} \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{t_1} \mathrm{d}t_2 e^{\mathrm{i}(\Omega_{\mathrm{A}} t_1 + \Omega_{\mathrm{B}} t_2)} \mathcal{X}_{\mathrm{A}}(t_1) \mathcal{X}_{\mathrm{B}}(t_2) \\ &\quad\times \int \mathrm{d}^3\boldsymbol{x}_1 \int \mathrm{d}^3\boldsymbol{x}_2\, \boldsymbol{F}^{\mathbf{t}}_{\mathrm{A}}(\boldsymbol{x}_1) \mathbf{W}(\boldsymbol{x}_1 + \boldsymbol{x}_{\mathrm{A}}, \boldsymbol{x}_2 + \boldsymbol{x}_{\mathrm{B}}, t_1, t_2) \boldsymbol{F}_{\mathrm{B}}(\boldsymbol{x}_2) \\ &= -\, e_{\mathrm{A}} e_{\mathrm{B}} \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{t_1} \mathrm{d}t_2 \int \mathrm{d}^3\boldsymbol{x}_1 \int \mathrm{d}^3\boldsymbol{x}_2\, e^{\mathrm{i}(\Omega_{\mathrm{A}} t_1 + \Omega_{\mathrm{B}} t_2)} \mathcal{X}_{\mathrm{A}}(t_1) \mathcal{X}_{\mathrm{B}}(t_2) \\ &\quad\times \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3} \frac{|\boldsymbol{k}|}{2} e^{-\mathrm{i}|\boldsymbol{k}|(t_1 - t_2)} e^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}_1} e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}_2} e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_{\mathrm{A}} - \boldsymbol{x}_{\mathrm{B}})} \\ &\quad\times \boldsymbol{x}_1^{\mathbf{t}} \left(\mathbb{1} - \frac{\boldsymbol{k}\otimes\boldsymbol{k}}{|\boldsymbol{k}|^2}\right) \boldsymbol{x}_2 \psi^*_{e_{\mathrm{A}}}(\boldsymbol{x}_1) \psi_{g_{\mathrm{A}}}(\boldsymbol{x}_1) \psi^*_{e_{\mathrm{B}}}(\boldsymbol{x}_2) \psi_{g_{\mathrm{B}}}(\boldsymbol{x}_2), \end{aligned} \tag{A.38}$$

where we have already performed the translations $\boldsymbol{x}_1 = \boldsymbol{x}'_1 - \boldsymbol{x}_{\mathrm{A}}$, $\boldsymbol{x}_2 = \boldsymbol{x}'_2 - \boldsymbol{x}_{\mathrm{B}}$ to eliminate the explicit dependence on $\boldsymbol{x}_{\mathrm{A}}$ and $\boldsymbol{x}_{\mathrm{B}}$ from the smearing vectors.

Note that the correlation term (A.38) depends on the relative spatial orientation of the two atoms. With the aim of defining a common reference frame for the two atoms, we will refer the orientation of atom B to the reference frame of atom A. This means that if the spherical harmonics used to describe atom A's orbitals are $Y^{\mathrm{A}}_{lm}(\theta_{\mathrm{A}}, \phi_{\mathrm{A}}) = Y_{lm}(\theta_{\mathrm{A}}, \phi_{\mathrm{A}})$, atom B's angular wave functions will be given by [MP87]

$$Y^{\mathrm{B}}_{lm}(\theta_{\mathrm{B}}, \phi_{\mathrm{B}}) = \sum_{\mu=-l}^{l} Y_{l\mu}(\theta_{\mathrm{B}}, \phi_{\mathrm{B}}) \mathcal{D}^{l}_{\mu,m}(\psi, \vartheta, \varphi), \tag{A.39}$$

where $(\psi, \vartheta, \varphi)$ are the Euler angles characterizing the rotation between the reference frames of B and A (see Fig. 2.1), and $\mathcal{D}^{l}_{\mu,m}$ represents the Wigner D-matrix coefficients, which characterize the rotation of the angular momentum operators and are given by

$$\begin{aligned} \mathcal{D}^{l}_{\mu,m}(\psi, \vartheta, \varphi) = e^{-\mathrm{i}(\mu\psi + m\varphi)} \sqrt{\frac{(l-m)!(l+\mu)!}{(l+m)!(l-\mu)!}} \frac{\cos(\vartheta/2)^{2l+m-\mu} [-\sin(\vartheta/2)]^{\mu-m}}{(\mu-m)!} \\ \times {}_2F_1\left(\mu - l, -m - l; \mu - m + 1; \tan^2 \frac{\vartheta}{2}\right). \end{aligned} \tag{A.40}$$

Taking this subtlety into account, and in the same way as we did with the local term $\mathcal{L}_{\mu\mu}$, we separate the term $\mathcal{M}^{\mathrm{AB}}$ into two parts, one containing the identity matrix and the other one containing the momentum dyadic $\boldsymbol{k} \otimes \boldsymbol{k}$. The

contribution to $\mathcal{M}^{\mathrm{AB}}$ containing the identity matrix is

$$
\begin{aligned}
\mathcal{M}^{\mathrm{AB}}\Big|_{\mathbb{1}} = &- e_{\mathrm{A}} e_{\mathrm{B}} \int_0^\infty \frac{\mathrm{d}|\boldsymbol{k}|}{(2\pi)^3} \frac{|\boldsymbol{k}|^3}{2} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} 4\pi \mathrm{i}^l \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} 4\pi \mathrm{i}^{l'} (-1)^{l'} \\
&\times \sum_{l''=0}^{\infty} \sum_{m''=-l''}^{l''} 4\pi \mathrm{i}^{l''} j_{l''}(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}} - \boldsymbol{x}_{\mathrm{B}}|) \frac{4\pi}{3} \\
&\times \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{t_1} \mathrm{d}t_2 \, e^{\mathrm{i}(\Omega_{\mathrm{A}} t_1 + \Omega_{\mathrm{B}} t_2)} \mathcal{X}_{\mathrm{A}}(t_1) \mathcal{X}_{\mathrm{B}}(t_2) e^{-\mathrm{i}|\boldsymbol{k}|(t_1 - t_2)} \\
&\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_1| \, |\boldsymbol{x}_1|^3 R^*_{n_e, l_e}(|\boldsymbol{x}_1|) R_{n_g, l_g}(|\boldsymbol{x}_1|) j_l(|\boldsymbol{k}||\boldsymbol{x}_1|) \\
&\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_2| \, |\boldsymbol{x}_2|^3 R^*_{n_e, l_e}(|\boldsymbol{x}_2|) R_{n_g, l_g}(|\boldsymbol{x}_2|) j_{l'}(|\boldsymbol{k}||\boldsymbol{x}_2|) \\
&\times \int \mathrm{d}\Omega_k \, Y_{lm}(\hat{\boldsymbol{k}}) Y_{l'm'}(\hat{\boldsymbol{k}}) Y_{l''m''}(\hat{\boldsymbol{k}}) Y^*_{l''m''}(\widehat{\Delta \boldsymbol{x}}) \\
&\times \int \mathrm{d}\Omega_1 (Y^{\mathrm{A}}_{l_e, m_e})^*(\hat{\boldsymbol{x}}_1) Y^{\mathrm{A}}_{l_g, m_g}(\hat{\boldsymbol{x}}_1) Y^*_{lm}(\hat{\boldsymbol{x}}_1) \\
&\times \int \mathrm{d}\Omega_2 (Y^{\mathrm{B}}_{l_e, m_e})^*(\hat{\boldsymbol{x}}_2) Y^{\mathrm{B}}_{l_g, m_g}(\hat{\boldsymbol{x}}_2) Y^*_{l'm'}(\hat{\boldsymbol{x}}_2) \\
&\times \left[Y_{10}(\hat{\boldsymbol{x}}_1) Y_{10}(\hat{\boldsymbol{x}}_2) - Y_{11}(\hat{\boldsymbol{x}}_1) Y_{1-1}(\hat{\boldsymbol{x}}_2) - Y_{1-1}(\hat{\boldsymbol{x}}_1) Y_{11}(\hat{\boldsymbol{x}}_2) \right], \quad \text{(A.41)}
\end{aligned}
$$

where $\widehat{\Delta \boldsymbol{x}}$ is the unit vector pointing in the direction of $\boldsymbol{x}_{\mathrm{A}} - \boldsymbol{x}_{\mathrm{B}}$.

This term is actually very similar to its local counterpart (A.21). Nevertheless, now one has an extra term $e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_{\mathrm{A}} - \boldsymbol{x}_{\mathrm{B}})}$, which needs to be decomposed into spherical harmonics as well using Eq. (A.19). This decomposition appears in the second and sixth lines of Eq. (A.41). Therefore there are three spherical harmonics in the integral over $\mathrm{d}\Omega_{\boldsymbol{k}}$ and five in each integral over $\mathrm{d}\Omega_1$ and $\mathrm{d}\Omega_2$. These integrals can be readily evaluated using the identities (although implicit

in the expressions, all the spherical harmonics depend on the same angles θ, ϕ)

$$\int \mathrm{d}\Omega\, Y_{l_1 m_1} Y_{l_2 m_2} Y_{l_3 m_3}$$
$$= \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}, \tag{A.42}$$

$$\int \mathrm{d}\Omega\, Y_{l_1 m_1} Y_{l_2 m_2} Y_{l_3 m_3} Y_{l_4 m_4} Y_{l_5 m_5}$$
$$= \sum_{\lambda,\lambda'=0}^{\infty} \frac{(2\lambda+1)(2\lambda'+1)}{4\pi} \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)(2l_4+1)(2l_5+1)}{4\pi}}$$
$$\times \begin{pmatrix} l_1 & l_2 & \lambda \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & \lambda \\ m_1 & m_2 & -m_1-m_2 \end{pmatrix} \begin{pmatrix} l_3 & \lambda' & \lambda \\ 0 & 0 & 0 \end{pmatrix}$$
$$\times \begin{pmatrix} l_3 & \lambda' & \lambda \\ m_3 & m_4+m_5 & m_1+m_2 \end{pmatrix} \begin{pmatrix} l_4 & l_5 & \lambda' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_4 & l_5 & \lambda' \\ m_4 & m_5 & -m_4-m_5 \end{pmatrix}. \tag{A.43}$$

We can orient the z axis in the integral over $\boldsymbol{k}$ along the vector $\boldsymbol{x}_\mathrm{A} - \boldsymbol{x}_\mathrm{B}$. In that way we can write $Y^*_{l''m''}(\widehat{\Delta \boldsymbol{x}}) = \delta_{m''0}\sqrt{(2l''+1)/(4\pi)}$, and thus easily perform the sum over m''.

The next step is to perform the integrals over the solid angles and the sums

over m and m':

$$
\begin{aligned}
&\sum_{m,m'} \int \mathrm{d}\Omega_k \, Y_{lm}(\hat{\boldsymbol{k}}) Y_{l'm'}(\hat{\boldsymbol{k}}) Y_{l''0}(\hat{\boldsymbol{k}}) \int \mathrm{d}\Omega_1 \, Y^*_{l_e,m_e}(\hat{\boldsymbol{x}}_1) Y_{l_g,m_g}(\hat{\boldsymbol{x}}_1) Y^*_{lm}(\hat{\boldsymbol{x}}_1) \\
&\quad \times \int \mathrm{d}\Omega_2 \sum_{\epsilon=-l_e}^{l_e} (\mathcal{D}^{l_e}_{\epsilon,m_e})^* Y^*_{l_e,\epsilon}(\hat{\boldsymbol{x}}_2) \sum_{\eta=-l_g}^{l_g} \mathcal{D}^{l_g}_{\eta,m_g} Y_{l_g,\eta}(\hat{\boldsymbol{x}}_2) Y^*_{l'm'}(\hat{\boldsymbol{x}}_2) \\
&\quad \times \left[Y_{10}(\hat{\boldsymbol{x}}_1) Y_{10}(\hat{\boldsymbol{x}}_2) - Y_{11}(\hat{\boldsymbol{x}}_1) Y_{1-1}(\hat{\boldsymbol{x}}_2) - Y_{1-1}(\hat{\boldsymbol{x}}_1) Y_{11}(\hat{\boldsymbol{x}}_2) \right] Y^*_{l''0}(\widehat{\Delta \boldsymbol{x}}) \\
&= \frac{3}{(4\pi)^3} (2l_e+1)(2l_g+1)(2l+1)(2l'+1)(2l''+1) \sum_{\eta} \mathcal{D}^{l_g}_{\eta,m_g} (\mathcal{D}^{l_e}_{\eta+m_g-m_e,m_e})^* \\
&\quad \times \sum_{\lambda,\lambda'} (2\lambda+1)(2\lambda'+1) \begin{pmatrix} l & l_e & \lambda \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_g & 1 & \lambda \\ 0 & 0 & 0 \end{pmatrix} \\
&\quad \times \begin{pmatrix} l' & l_e & \lambda' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_g & 1 & \lambda' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix} \\
&\quad \times \Bigg[\begin{pmatrix} l_g & 1 & \lambda \\ m_g & 0 & -m_g \end{pmatrix} \begin{pmatrix} l & l_e & \lambda \\ m_e - m_g & -m_e & m_g \end{pmatrix} \begin{pmatrix} l_g & 1 & \lambda' \\ \eta & 0 & -\eta \end{pmatrix} \\
&\qquad \times \begin{pmatrix} l' & l_e & \lambda' \\ m_g - m_e & m_e - m_g - \eta & \eta \end{pmatrix} \begin{pmatrix} l & l' & l'' \\ m_g - m_e & m_e - m_g & 0 \end{pmatrix} \\
&\qquad - \begin{pmatrix} l_g & 1 & \lambda \\ m_g & 1 & -1-m_g \end{pmatrix} \begin{pmatrix} l & l_e & \lambda \\ m_e - m_g - 1 & -m_e & m_g + 1 \end{pmatrix} \begin{pmatrix} l_g & 1 & \lambda' \\ \eta & -1 & 1-\eta \end{pmatrix} \\
&\qquad \times \begin{pmatrix} l' & l_e & \lambda' \\ m_g - m_e + 1 & m_e - m_g - \eta & \eta - 1 \end{pmatrix} \begin{pmatrix} l & l' & l'' \\ m_g - m_e + 1 & m_e - m_g - 1 & 0 \end{pmatrix} \\
&\qquad - \begin{pmatrix} l_g & 1 & \lambda \\ m_g & -1 & 1-m_g \end{pmatrix} \begin{pmatrix} l & l_e & \lambda \\ 1 + m_e - m_g & -m_e & m_g - 1 \end{pmatrix} \begin{pmatrix} l_g & 1 & \lambda' \\ \eta & 1 & -1-\eta \end{pmatrix} \\
&\qquad \times \begin{pmatrix} l' & l_e & \lambda' \\ m_g - m_e - 1 & m_e - m_g - \eta & \eta + 1 \end{pmatrix} \begin{pmatrix} l & l' & l'' \\ m_g - m_e - 1 & 1 + m_e - m_g & 0 \end{pmatrix} \Bigg] .
\end{aligned}
\tag{A.44}
$$

As it was the case for the local term, this term cannot be further simplified without specifying the specific atomic levels of the ground and excited states.

We address now the contribution of the dyadic $\boldsymbol{k}\otimes\boldsymbol{k}$, which reads

$$
\begin{aligned}
\mathcal{M}^{\mathrm{AB}}\Big|_{\boldsymbol{k}\otimes\boldsymbol{k}} = &- e_{\mathrm{A}}e_{\mathrm{B}}\int_0^\infty \frac{\mathrm{d}|\boldsymbol{k}|}{(2\pi)^3}\frac{|\boldsymbol{k}|^3}{2}\sum_{l=0}^{\infty}\sum_{m=-l}^{l} 4\pi\mathrm{i}^l \sum_{l'=0}^{\infty}\sum_{m'=-l'}^{l'} 4\pi\mathrm{i}^{l'}(-1)^{l'} \\
&\times \sum_{l''=0}^{\infty} 4\pi\mathrm{i}^{l''} j_{l''}(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)\sqrt{\frac{2l''+1}{4\pi}}\left(\frac{4\pi}{3}\right)^2 \\
&\times \int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\, e^{\mathrm{i}(\Omega_{\mathrm{A}}t_1+\Omega_{\mathrm{B}}t_2)}\mathcal{X}_{\mathrm{A}}(t_1)\mathcal{X}_{\mathrm{B}}(t_2)e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)} \\
&\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_1|\,|\boldsymbol{x}_1|^3 R^*_{n_e,l_e}(|\boldsymbol{x}_1|)R_{n_g,l_g}(|\boldsymbol{x}_1|)j_l(|\boldsymbol{k}||\boldsymbol{x}_1|) \\
&\times \int_0^\infty \mathrm{d}|\boldsymbol{x}_2|\,|\boldsymbol{x}_2|^3 R^*_{n_e,l_e}(|\boldsymbol{x}_2|)R_{n_g,l_g}(|\boldsymbol{x}_2|)j_{l'}(|\boldsymbol{k}||\boldsymbol{x}_2|) \\
&\times \int \mathrm{d}\Omega_k\, Y_{lm}(\hat{\boldsymbol{k}})Y_{l'm'}(\hat{\boldsymbol{k}})Y_{l''0}(\hat{\boldsymbol{k}})\int \mathrm{d}\Omega_1 (Y^{\mathrm{A}}_{l_e,m_e})^*(\hat{\boldsymbol{x}}_1)Y^{\mathrm{A}}_{l_g,m_g}(\hat{\boldsymbol{x}}_1)Y^*_{lm}(\hat{\boldsymbol{x}}_1) \\
&\times \int \mathrm{d}\Omega_2 (Y^{\mathrm{B}}_{l_e,m_e})^*(\hat{\boldsymbol{x}}_2)Y^{\mathrm{B}}_{l_g,m_g}(\hat{\boldsymbol{x}}_2)Y^*_{l'm'}(\hat{\boldsymbol{x}}_2) \\
&\times \Big[Y_{10}(\hat{\boldsymbol{x}}_1)Y_{10}(\hat{\boldsymbol{k}}) - Y_{11}(\hat{\boldsymbol{x}}_1)Y_{1-1}(\hat{\boldsymbol{k}}) - Y_{1-1}(\hat{\boldsymbol{x}}_1)Y_{11}(\hat{\boldsymbol{k}})\Big] \\
&\times \Big[Y_{10}(\hat{\boldsymbol{k}})Y_{10}(\hat{\boldsymbol{x}}_2) - Y_{11}(\hat{\boldsymbol{k}})Y_{1-1}(\hat{\boldsymbol{x}}_2) - Y_{1-1}(\hat{\boldsymbol{k}})Y_{11}(\hat{\boldsymbol{x}}_2)\Big]. \qquad \text{(A.45)}
\end{aligned}
$$

In this case, the sums over m and m' and the integrals over solid angles in

the two last lines yield

$$
\begin{aligned}
&\sum_{m,m'} \int \mathrm{d}\Omega_k \, Y_{lm}(\hat{\boldsymbol{k}}) Y_{l'm'}(\hat{\boldsymbol{k}}) Y_{l''0}(\hat{\boldsymbol{k}}) \int \mathrm{d}\Omega_1 Y^*_{l_e,m_e}(\hat{\boldsymbol{x}}_1) Y_{l_g,m_g}(\hat{\boldsymbol{x}}_1) Y^*_{lm}(\hat{\boldsymbol{x}}_1) \\
&\qquad \times \int \mathrm{d}\Omega_2 \sum_{\epsilon} (\mathcal{D}^{l_e}_{\epsilon,m_e})^* Y^*_{l_e,\epsilon}(\hat{\boldsymbol{x}}_2) \sum_{\eta} \mathcal{D}^{l_g}_{\eta,m_g} Y_{l_g,\eta}(\hat{\boldsymbol{x}}_2) Y^*_{l'm'}(\hat{\boldsymbol{x}}_2) \\
&\qquad \times \Big[Y_{10}(\hat{\boldsymbol{x}}_1) Y_{10}(\hat{\boldsymbol{k}}) - Y_{11}(\hat{\boldsymbol{x}}_1) Y_{1-1}(\hat{\boldsymbol{k}}) - Y_{1-1}(\hat{\boldsymbol{x}}_1) Y_{11}(\hat{\boldsymbol{k}}) \Big] \\
&\qquad \times \Big[Y_{10}(\hat{\boldsymbol{k}}) Y_{10}(\hat{\boldsymbol{x}}_2) - Y_{11}(\hat{\boldsymbol{k}}) Y_{1-1}(\hat{\boldsymbol{x}}_2) - Y_{1-1}(\hat{\boldsymbol{k}}) Y_{11}(\hat{\boldsymbol{x}}_2) \Big] \\
&= \sum_{\lambda'\lambda''\lambda'''} \frac{(2\lambda'+1)(2\lambda''+1)(2\lambda'''+1)}{(4\pi)^3} (2l+1)(2l'+1)(2l_g+1)(2l_e+1) \\
&\qquad \times 9 \sqrt{\frac{2l''+1}{4\pi}} \sum_{\eta=-l_g}^{l_g} (\mathcal{D}^{l_e}_{\eta+m_g-m_e,m_e})^* \mathcal{D}^{l_g}_{\eta,m_g} \begin{pmatrix} l & l' & \lambda' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_e & l & \lambda'' \\ 0 & 0 & 0 \end{pmatrix} \\
&\qquad \times \begin{pmatrix} l_g & 1 & \lambda'' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_g & 1 & \lambda''' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_e & l' & \lambda''' \\ 0 & 0 & 0 \end{pmatrix} (A_{\mathcal{M}} + B_{\mathcal{M}}) \,, \qquad \text{(A.46)}
\end{aligned}
$$

where now the quantities $B_{\mathcal{M}}$ and $A_{\mathcal{M}}$ are

$$
\begin{aligned}
B_{\mathcal{M}} &= \sqrt{\frac{2}{3}} \begin{pmatrix} l'' & \lambda' & 2 \\ 0 & 0 & 0 \end{pmatrix} \\
&\times \Bigg[\begin{pmatrix} l'' & \lambda' & 2 \\ 0 & -2 & 2 \end{pmatrix} \begin{pmatrix} l & l' & \lambda' \\ m_g - m_e - 1 & m_e - m_g - 1 & 2 \end{pmatrix} \begin{pmatrix} l_e & l & \lambda'' \\ -m_e & 1 + m_e - m_g & m_g - 1 \end{pmatrix} \\
&\quad \times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & -1 & 1 - m_g \end{pmatrix} \begin{pmatrix} l_e & l' & \lambda''' \\ m_e - m_g - \eta & m_g - m_e + 1 & \eta - 1 \end{pmatrix} \begin{pmatrix} l_g & 1 & \lambda''' \\ \eta & -1 & 1 - \eta \end{pmatrix} \\
&+ \begin{pmatrix} l'' & \lambda' & 2 \\ 0 & 2 & -2 \end{pmatrix} \begin{pmatrix} l & l' & \lambda' \\ m_g - m_e + 1 & 1 + m_e - m_g & -2 \end{pmatrix} \begin{pmatrix} l_e & l & \lambda'' \\ -m_e & m_e - m_g - 1 & m_g + 1 \end{pmatrix} \\
&\quad \times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & 1 & -1 - m_g \end{pmatrix} \begin{pmatrix} l_e & l' & \lambda''' \\ m_e - m_g - \eta & m_g - m_e - 1 & \eta + 1 \end{pmatrix} \begin{pmatrix} l_g & 1 & \lambda''' \\ \eta & 1 & -1 - \eta \end{pmatrix} \Bigg] \,, \qquad \text{(A.47)}
\end{aligned}
$$

$$
\begin{aligned}
A_{\mathcal{M}} = &\sum_{\lambda}(2\lambda+1)\begin{pmatrix} l'' & \lambda' & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & \lambda \\ 0 & 0 & 0 \end{pmatrix}^2 \\
&\times \begin{pmatrix} l & l' & \lambda' \\ m_g - m_e & m_e - m_g & 0 \end{pmatrix}\begin{pmatrix} l_e & l & \lambda'' \\ -m_e & m_e - m_g & m_g \end{pmatrix} \\
&\times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & 0 & -m_g \end{pmatrix}\begin{pmatrix} l_e & l' & \lambda''' \\ m_e - m_g - \eta & m_g - m_e & \eta \end{pmatrix}\begin{pmatrix} l_g & 1 & \lambda''' \\ \eta & 0 & -\eta \end{pmatrix} \\
&- \sum_{\lambda}(2\lambda+1)\begin{pmatrix} 1 & 1 & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & \lambda \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda \\ 0 & -1 & 1 \end{pmatrix} \\
&\times \Bigg[\begin{pmatrix} l & l' & \lambda' \\ m_g - m_e & m_e - m_g - 1 & 1 \end{pmatrix}\begin{pmatrix} l_e & l & \lambda'' \\ -m_e & m_e - m_g & m_g \end{pmatrix}\begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & 0 & -m_g \end{pmatrix} \\
&\quad \times \begin{pmatrix} l_e & l' & \lambda''' \\ m_e - m_g - \eta & m_g - m_e + 1 & \eta - 1 \end{pmatrix}\begin{pmatrix} l_g & 1 & \lambda''' \\ \eta & -1 & 1 - \eta \end{pmatrix} \\
&\quad + \begin{pmatrix} l & l' & \lambda' \\ m_g - m_e - 1 & m_e - m_g & 1 \end{pmatrix}\begin{pmatrix} l_e & l & \lambda'' \\ -m_e & 1 + m_e - m_g & m_g - 1 \end{pmatrix} \\
&\quad \times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & -1 & 1 - m_g \end{pmatrix}\begin{pmatrix} l_e & l' & \lambda''' \\ m_e - m_g - \eta & m_g - m_e & \eta \end{pmatrix}\begin{pmatrix} l_g & 1 & \lambda''' \\ \eta & 0 & -\eta \end{pmatrix}\Bigg] \\
&- \sum_{\lambda}(2\lambda+1)\begin{pmatrix} 1 & 1 & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & \lambda \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda \\ 0 & 1 & -1 \end{pmatrix} \\
&\times \Bigg[\begin{pmatrix} l & l' & \lambda' \\ m_g - m_e & 1 + m_e - m_g & -1 \end{pmatrix}\begin{pmatrix} l_e & l & \lambda'' \\ -m_e & m_e - m_g & m_g \end{pmatrix}\begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & 0 & -m_g \end{pmatrix} \\
&\quad \times \begin{pmatrix} l_e & l' & \lambda''' \\ m_e - m_g - \eta & m_g - m_e - 1 & \eta + 1 \end{pmatrix}\begin{pmatrix} l_g & 1 & \lambda''' \\ \eta & 1 & -1 - \eta \end{pmatrix} \\
&\quad + \begin{pmatrix} l & l' & \lambda' \\ m_g - m_e + 1 & m_e - m_g & -1 \end{pmatrix}\begin{pmatrix} l_e & l & \lambda'' \\ -m_e & m_e - m_g - 1 & m_g + 1 \end{pmatrix} \\
&\quad \times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & 1 & -1 - m_g \end{pmatrix}\begin{pmatrix} l_e & l' & \lambda''' \\ m_e - m_g - \eta & m_g - m_e & \eta \end{pmatrix}\begin{pmatrix} l_g & 1 & \lambda''' \\ \eta & 0 & -\eta \end{pmatrix}\Bigg] \\
&+ \sum_{\lambda}(2\lambda+1)\begin{pmatrix} 1 & 1 & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & \lambda \\ 1 & -1 & 0 \end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda \\ 0 & 0 & 0 \end{pmatrix} \\
&\times \Bigg[\begin{pmatrix} l & l' & \lambda' \\ m_g - m_e + 1 & m_e - m_g - 1 & 0 \end{pmatrix}\begin{pmatrix} l_e & l & \lambda'' \\ -m_e & m_e - m_g - 1 & 1 + m_g \end{pmatrix} \\
&\quad \times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & 1 & -1 - m_g \end{pmatrix}\begin{pmatrix} l_e & l' & \lambda''' \\ m_e - m_g - \eta & m_g - m_e + 1 & \eta - 1 \end{pmatrix}\begin{pmatrix} l_g & 1 & \lambda''' \\ \eta & -1 & 1 - \eta \end{pmatrix} \\
&\quad + \begin{pmatrix} l & l' & \lambda' \\ m_g - m_e - 1 & 1 + m_e - m_g & 0 \end{pmatrix}\begin{pmatrix} l_e & l & \lambda'' \\ -m_e & 1 + m_e - m_g & m_g - 1 \end{pmatrix} \\
&\quad \times \begin{pmatrix} l_g & 1 & \lambda'' \\ m_g & -1 & 1 - m_g \end{pmatrix}\begin{pmatrix} l_e & l' & \lambda''' \\ m_e - m_g - \eta & m_g - m_e - 1 & \eta + 1 \end{pmatrix}\begin{pmatrix} l_g & 1 & \lambda''' \\ \eta & 1 & -1 - \eta \end{pmatrix}\Bigg].
\end{aligned}
\tag{A.48}
$$

We particularize now to the scenario described in the paper, recall, the ground states being hydrogenoid $1s$ states and the excited states being hydrogenoid $2p_z$ states. In this particular situation, the contribution of the nonlocal term proportional to the identity reads

$$
\begin{aligned}
\mathcal{M}^{\mathrm{AB}}\Big|_{\mathbb{1}} = & -e_{\mathrm{A}}e_{\mathrm{B}}\int_0^\infty \frac{\mathrm{d}|\boldsymbol{k}|}{(2\pi)^3}\frac{|\boldsymbol{k}|^3}{2}\sum_{l=0}^{\infty}4\pi\mathrm{i}^l\sum_{l'=0}^{\infty}4\pi\mathrm{i}^{l'}(-1)^{l'}\sum_{l''=0}^{\infty}4\pi\mathrm{i}^{l''}j_{l''}(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)\frac{4\pi}{3}\\
&\times\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\, e^{\mathrm{i}(\Omega_{\mathrm{A}}t_1+\Omega_{\mathrm{B}}t_2)}\mathcal{X}_{\mathrm{A}}(t_1)\mathcal{X}_{\mathrm{B}}(t_2)e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\\
&\times\int_0^\infty \mathrm{d}|\boldsymbol{x}_1|\,|\boldsymbol{x}_1|^3R^*_{2,1}(|\boldsymbol{x}_1|)R_{1,0}(|\boldsymbol{x}_1|)j_l(|\boldsymbol{k}||\boldsymbol{x}_1|)\\
&\times\int_0^\infty \mathrm{d}|\boldsymbol{x}_2|\,|\boldsymbol{x}_2|^3R^*_{2,1}(|\boldsymbol{x}_2|)R_{1,0}(|\boldsymbol{x}_2|)j_{l'}(|\boldsymbol{k}||\boldsymbol{x}_2|)\\
&\times\frac{3}{(4\pi)^3}(2+1)(0+1)(2l+1)(2l'+1)(2l''+1)\sum_{\eta}\mathcal{D}^0_{\eta,0}(\mathcal{D}^1_{\eta,0})^*\\
&\times\sum_{\lambda,\lambda'}(2\lambda+1)(2\lambda'+1)\begin{pmatrix}l&1&\lambda\\0&0&0\end{pmatrix}\begin{pmatrix}0&1&\lambda\\0&0&0\end{pmatrix}\begin{pmatrix}l'&1&\lambda'\\0&0&0\end{pmatrix}\begin{pmatrix}0&1&\lambda'\\0&0&0\end{pmatrix}\begin{pmatrix}l&l'&l''\\0&0&0\end{pmatrix}\\
&\times\Bigg[\begin{pmatrix}0&1&\lambda\\0&0&0\end{pmatrix}\begin{pmatrix}l&1&\lambda\\0&0&0\end{pmatrix}\begin{pmatrix}0&1&\lambda'\\\eta&0&-\eta\end{pmatrix}\begin{pmatrix}l'&1&\lambda'\\0&-\eta&\eta\end{pmatrix}\begin{pmatrix}l&l'&l''\\0&0&0\end{pmatrix}\\
&\quad-\begin{pmatrix}0&1&\lambda\\0&1&-1\end{pmatrix}\begin{pmatrix}l&1&\lambda\\-1&0&1\end{pmatrix}\begin{pmatrix}0&1&\lambda'\\\eta&-1&1-\eta\end{pmatrix}\begin{pmatrix}l'&1&\lambda'\\1&-\eta&\eta-1\end{pmatrix}\begin{pmatrix}l&l'&l''\\1&-1&0\end{pmatrix}\\
&\quad-\begin{pmatrix}0&1&\lambda\\0&-1&1\end{pmatrix}\begin{pmatrix}l&1&\lambda\\1&0&-1\end{pmatrix}\begin{pmatrix}0&1&\lambda'\\\eta&1&-1-\eta\end{pmatrix}\begin{pmatrix}l'&1&\lambda'\\-1&-\eta&\eta+1\end{pmatrix}\begin{pmatrix}l&l'&l''\\-1&1&0\end{pmatrix}\Bigg].
\end{aligned}
\tag{A.49}
$$

The Wigner D-matrix coefficients are nonzero only for $\eta=0$, for which $\mathcal{D}^0_{0,0}=1$ and $\mathcal{D}^1_{0,0}(\psi,\vartheta,\varphi)=\cos\vartheta$. Using the properties of the $3j$-symbols we see that the sums have nonzero terms only for $\lambda=1$ and $\lambda'=1$. Additionally, we also obtain the restrictions $l=0,1,2$, $l'=0,1,2$ and, as a consequence of these

last two, $l''=0,1,2,3,4$. Therefore, the computation of the sums yields

$$\begin{aligned}
\mathcal{M}^{\mathrm{AB}}\Big|_{\mathbb{1}} =& -e_{\mathrm{A}}e_{\mathrm{B}}\frac{\cos\vartheta}{12\pi^2}\int_0^\infty \mathrm{d}|\boldsymbol{k}|\,|\boldsymbol{k}|^3\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\, e^{\mathrm{i}(\Omega_{\mathrm{A}}t_1+\Omega_{\mathrm{B}}t_2)}\mathcal{X}_{\mathrm{A}}(t_1)\mathcal{X}_{\mathrm{B}}(t_2)e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\\
&\times\int_0^\infty \mathrm{d}|\boldsymbol{x}_1|\,|\boldsymbol{x}_1|^3 R^*_{2,1}(|\boldsymbol{x}_1|)R_{1,0}(|\boldsymbol{x}_1|)\int_0^\infty \mathrm{d}|\boldsymbol{x}_2|\,|\boldsymbol{x}_2|^3 R^*_{2,1}(|\boldsymbol{x}_2|)R_{1,0}(|\boldsymbol{x}_2|)\\
&\times\Big\{j_0(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)\left[j_0(|\boldsymbol{k}||\boldsymbol{x}_1|)j_0(|\boldsymbol{k}||\boldsymbol{x}_2|)+2j_2(|\boldsymbol{k}||\boldsymbol{x}_1|)j_2(|\boldsymbol{k}||\boldsymbol{x}_2|)\right]\\
&+2j_2(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)\left[j_2(|\boldsymbol{k}||\boldsymbol{x}_1|)j_0(|\boldsymbol{k}||\boldsymbol{x}_2|)+j_0(|\boldsymbol{k}||\boldsymbol{x}_1|)j_2(|\boldsymbol{k}||\boldsymbol{x}_2|)\right.\\
&\left.-j_2(|\boldsymbol{k}||\boldsymbol{x}_1|)j_2(|\boldsymbol{k}||\boldsymbol{x}_2|)\right]\Big\}\\
=& -e_{\mathrm{A}}e_{\mathrm{B}}\frac{\cos\vartheta}{12\pi^2}\int_0^\infty \mathrm{d}|\boldsymbol{k}|\,|\boldsymbol{k}|^3\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\, e^{\mathrm{i}(\Omega_{\mathrm{A}}t_1+\Omega_{\mathrm{B}}t_2)}\mathcal{X}_{\mathrm{A}}(t_1)\mathcal{X}_{\mathrm{B}}(t_2)e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\\
&\times\left[j_0(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)\frac{7962624a_0^2\left(16a_0^4|\boldsymbol{k}|^4-8a_0^2|\boldsymbol{k}|^2+9\right)}{\left(4a_0^2|\boldsymbol{k}|^2+9\right)^8}\right.\\
&\left.-j_2(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)\frac{28311552a_0^4|\boldsymbol{k}|^2\left(8a_0^2|\boldsymbol{k}|^2-9\right)}{\left(4a_0^2|\boldsymbol{k}|^2+9\right)^8}\right]. \qquad \text{(A.50)}
\end{aligned}$$

Operating in an analogous way, the contribution containing the $\boldsymbol{k}\otimes\boldsymbol{k}$ dyadic

takes the following form

$$
\begin{aligned}
\mathcal{M}^{\mathrm{AB}}\big|_{\boldsymbol{k}\otimes\boldsymbol{k}} &= -e_{\mathrm{A}}e_{\mathrm{B}}(\mathcal{D}^{1}_{0,0})^{*}\mathcal{D}^{0}_{0,0}\int_0^\infty \frac{\mathrm{d}|\boldsymbol{k}|}{(2\pi)^3}\frac{|\boldsymbol{k}|^3}{2}\sum_{l=0}^{\infty}4\pi\mathrm{i}^{l}\sum_{l'=0}^{\infty}4\pi\mathrm{i}^{l'}(-1)^{l'}\sum_{l''=0}^{\infty}4\pi\mathrm{i}^{l''}j_{l''}(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)\left(\frac{4\pi}{3}\right)^2\\
&\times\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\, e^{\mathrm{i}(\Omega_{\mathrm{A}}t_1+\Omega_{\mathrm{B}}t_2)}\mathcal{X}_{\mathrm{A}}(t_1)\mathcal{X}_{\mathrm{B}}(t_2)e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\sum_{\lambda'}\frac{243(2\lambda'+1)(2l+1)(2l'+1)(2l''+1)}{(4\pi)^4}\\
&\times\int_0^\infty \mathrm{d}|\boldsymbol{x}_1|\,|\boldsymbol{x}_1|^3R^{*}_{2,1}(|\boldsymbol{x}_1|)R_{1,0}(|\boldsymbol{x}_1|)j_l(|\boldsymbol{k}||\boldsymbol{x}_1|)\int_0^\infty \mathrm{d}|\boldsymbol{x}_2|\,|\boldsymbol{x}_2|^3R^{*}_{2,1}(|\boldsymbol{x}_2|)R_{1,0}(|\boldsymbol{x}_2|)j_{l'}(|\boldsymbol{k}||\boldsymbol{x}_2|)\\
&\times\begin{pmatrix} l & l' & \lambda'\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 1 & l & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 1 & l' & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}\\
&\times\Bigg\{\sqrt{\frac{2}{3}}\begin{pmatrix} l'' & \lambda' & 2\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l'' & \lambda' & 2\\ 0 & -2 & 2\end{pmatrix}\begin{pmatrix} l & l' & \lambda'\\ -1 & -1 & 2\end{pmatrix}\\
&\quad\times\begin{pmatrix} 1 & l & 1\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} 1 & l' & 1\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & -1 & 1\end{pmatrix}\\
&\quad+\sqrt{\frac{2}{3}}\begin{pmatrix} l'' & \lambda' & 2\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l'' & \lambda' & 2\\ 0 & 2 & -2\end{pmatrix}\begin{pmatrix} l & l' & \lambda'\\ 1 & 1 & -2\end{pmatrix}\\
&\quad\times\begin{pmatrix} 1 & l & 1\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix} 1 & l' & 1\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 1 & -1\end{pmatrix}\\
&\quad+\sum_{\lambda}(2\lambda+1)\begin{pmatrix} l'' & \lambda' & \lambda\\ 0 & 0 & 0\end{pmatrix}^2\begin{pmatrix} 1 & 1 & \lambda\\ 0 & 0 & 0\end{pmatrix}^2\begin{pmatrix} l & l' & \lambda'\\ 0 & 0 & 0\end{pmatrix}\\
&\quad\times\begin{pmatrix} 1 & l & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 1 & l' & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}\\
&\quad-\sum_{\lambda}(2\lambda+1)\begin{pmatrix} 1 & 1 & \lambda\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 1 & 1 & \lambda\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda\\ 0 & -1 & 1\end{pmatrix}\\
&\quad\times\Bigg[\begin{pmatrix} l & l' & \lambda'\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} 1 & l & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 1 & l' & 1\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & -1 & 1\end{pmatrix}\\
&\qquad+\begin{pmatrix} l & l' & \lambda'\\ -1 & 0 & 1\end{pmatrix}\begin{pmatrix} 1 & l & 1\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} 1 & l' & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}\Bigg]\\
&\quad-\sum_{\lambda}(2\lambda+1)\begin{pmatrix} 1 & 1 & \lambda\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 1 & 1 & \lambda\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda\\ 0 & 1 & -1\end{pmatrix}\\
&\quad\times\Bigg[\begin{pmatrix} l & l' & \lambda'\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix} 1 & l & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 1 & l' & 1\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 1 & -1\end{pmatrix}\\
&\qquad+\begin{pmatrix} l & l' & \lambda'\\ 1 & 0 & -1\end{pmatrix}\begin{pmatrix} 1 & l & 1\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix} 1 & l' & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}\Bigg]\\
&\quad+\sum_{\lambda}(2\lambda+1)\begin{pmatrix} 1 & 1 & \lambda\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} 1 & 1 & \lambda\\ 1 & -1 & 0\end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l'' & \lambda' & \lambda\\ 0 & 0 & 0\end{pmatrix}\\
&\quad\times\Bigg[\begin{pmatrix} l & l' & \lambda'\\ 1 & -1 & 0\end{pmatrix}\begin{pmatrix} 1 & l & 1\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix} 1 & l' & 1\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & -1 & 1\end{pmatrix}\\
&\qquad+\begin{pmatrix} l & l' & \lambda'\\ -1 & 1 & 0\end{pmatrix}\begin{pmatrix} 1 & l & 1\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} 1 & l' & 1\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix} 0 & 1 & 1\\ 0 & 1 & -1\end{pmatrix}\Bigg]\Bigg\}. \qquad \text{(A.51)}
\end{aligned}
$$

After evaluation of all the sums, the expression one reaches is

$$\begin{aligned}\mathcal{M}^{\mathrm{AB}}\Big|_{\boldsymbol{k}\otimes\boldsymbol{k}} =& -e_{\mathrm{A}}e_{\mathrm{B}}\frac{\cos\vartheta}{36\pi^2}\int \mathrm{d}|\boldsymbol{k}||\boldsymbol{k}|^3\left[j_0(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)-2j_2(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)\right]\\ &\times\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\, e^{\mathrm{i}(\Omega_{\mathrm{A}}t_1+\Omega_{\mathrm{B}}t_2)}\mathcal{X}_{\mathrm{A}}(t_1)\mathcal{X}_{\mathrm{B}}(t_2)e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\\ &\times\int \mathrm{d}|\boldsymbol{x}_1||\boldsymbol{x}_1|^3R^*_{2,1}(|\boldsymbol{x}_1|)R_{1,0}(|\boldsymbol{x}_1|)\left[j_0(|\boldsymbol{k}||\boldsymbol{x}_1|)-2j_2(|\boldsymbol{k}||\boldsymbol{x}_1|)\right]\\ &\times\int \mathrm{d}|\boldsymbol{x}_2||\boldsymbol{x}_2|^3R^*_{2,1}(|\boldsymbol{x}_2|)R_{1,0}(|\boldsymbol{x}_2|)\left[j_0(|\boldsymbol{k}||\boldsymbol{x}_2|)-2j_2(|\boldsymbol{k}||\boldsymbol{x}_2|)\right].\end{aligned} \tag{A.52}$$

Finally, performing the integrals over $\mathrm{d}|\boldsymbol{x}_1|$ and $\mathrm{d}|\boldsymbol{x}_2|$ yields

$$\begin{aligned}\mathcal{M}^{\mathrm{AB}}\Big|_{\boldsymbol{k}\otimes\boldsymbol{k}} =& -e_{\mathrm{A}}e_{\mathrm{B}}\frac{24576\cos\vartheta}{\pi^2}a_0^2\int_0^{\infty}\mathrm{d}|\boldsymbol{k}|\,|\boldsymbol{k}|^3\left[j_0(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)-2j_2(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)\right]\\ &\times\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\, e^{\mathrm{i}(\Omega_{\mathrm{A}}t_1+\Omega_{\mathrm{B}}t_2)}\mathcal{X}_{\mathrm{A}}(t_1)\mathcal{X}_{\mathrm{B}}(t_2)e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\frac{(9-20a_0^2|\boldsymbol{k}|^2)^2}{(4a_0^2|\boldsymbol{k}|^2+9)^8},\end{aligned} \tag{A.53}$$

and therefore, subtracting $\mathcal{M}^{\mathrm{AB}}|_{\boldsymbol{k}\otimes\boldsymbol{k}}$ from $\mathcal{M}^{\mathrm{AB}}|_{\mathbb{1}}$, the complete contribution $\mathcal{M}^{\mathrm{AB}}$ is

$$\begin{aligned}\mathcal{M}^{\mathrm{AB}} =& -e_{\mathrm{A}}e_{\mathrm{B}}\frac{49152\cos\vartheta}{\pi^2}a_0^2\int_0^{\infty}\mathrm{d}|\boldsymbol{k}|\,|\boldsymbol{k}|^3\frac{j_0(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)+j_2(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)}{(4a_0^2|\boldsymbol{k}|^2+9)^6}\\ &\times\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\, e^{\mathrm{i}(\Omega_{\mathrm{A}}t_1+\Omega_{\mathrm{B}}t_2)}\mathcal{X}_{\mathrm{A}}(t_1)\mathcal{X}_{\mathrm{B}}(t_2)e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}.\end{aligned} \tag{A.54}$$

One could be tempted to think that computing the term $\mathcal{M}^{\mathrm{BA}}$ just amounts to switching the labels $\mathrm{A}\leftrightarrow\mathrm{B}$ in the expression above. Nevertheless, this is not quite the case, since when we performed the integrals over solid angles we wrote the angular wave functions of atom B with respect the reference frame of atom A. We need to implement the following additional substitutions of Euler angles to obtain $\mathcal{M}_{\mathrm{BA}}$ from $\mathcal{M}_{\mathrm{AB}}$:

$$\psi_{\mathrm{B}\to\mathrm{A}}=-\varphi_{\mathrm{A}\to\mathrm{B}},\qquad \vartheta_{\mathrm{B}\to\mathrm{A}}=-\vartheta_{\mathrm{A}\to\mathrm{B}},\qquad \varphi_{\mathrm{B}\to\mathrm{A}}=-\psi_{\mathrm{A}\to\mathrm{B}}. \tag{A.55}$$

In the case under study here, this just amounts (in addition to changing the labels $\mathrm{A}\leftrightarrow\mathrm{B}$) to changing $\vartheta\to-\vartheta$ in Eq. (A.54). Therefore the complete

nonlocal term $\mathcal{M}$ is

$$\begin{aligned}\mathcal{M} =& -e_{\mathrm{A}}e_{\mathrm{B}}\frac{49152\cos\vartheta}{\pi^2}a_0^2\int_0^\infty \mathrm{d}|\boldsymbol{k}|\,|\boldsymbol{k}|^3\frac{j_0(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)+j_2(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)}{(4a_0^2|\boldsymbol{k}|^2+9)^6}\\ &\times\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\left[e^{\mathrm{i}(\Omega_{\mathrm{A}}t_1+\Omega_{\mathrm{B}}t_2)}\mathcal{X}_{\mathrm{A}}(t_1)\mathcal{X}_{\mathrm{B}}(t_2)e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\right.\\ &\left.+e^{\mathrm{i}(\Omega_{\mathrm{B}}t_1+\Omega_{\mathrm{A}}t_2)}\mathcal{X}_{\mathrm{B}}(t_1)\mathcal{X}_{\mathrm{A}}(t_2)e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\right].\end{aligned}\tag{A.56}$$

Finally, let us perform the integrals in time for Gaussian switching functions, which admit closed-form expressions. Recall $\mathcal{X}_\mu(t)=e^{-(t-t_\mu)^2/T^2}$. Therefore the time integrals take the form

$$\begin{aligned}&\int_{-\infty}^{\infty}\mathrm{d}t_1\int_{-\infty}^{t_1}\mathrm{d}t_2\left[e^{\mathrm{i}(\Omega_{\mathrm{A}}t_1+\Omega_{\mathrm{B}}t_2)}e^{-\frac{(t_1-t_{\mathrm{A}})^2}{T^2}}e^{-\frac{(t_2-t_{\mathrm{B}})^2}{T^2}}e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\right.\\ &\qquad\left.+e^{\mathrm{i}(\Omega_{\mathrm{B}}t_1+\Omega_{\mathrm{A}}t_2)}e^{-\frac{(t_1-t_{\mathrm{B}})^2}{T^2}}e^{-\frac{(t_2-t_{\mathrm{A}})^2}{T^2}}e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\right]\\ =&\sqrt{\pi}\frac{T}{2}\int_{-\infty}^{\infty}\mathrm{d}t_1\\ &\times\left\{e^{\mathrm{i}(\Omega_{\mathrm{A}}+\Omega_{\mathrm{B}})t_{\mathrm{B}}}e^{-\frac{1}{4}(|\boldsymbol{k}|+\Omega_{\mathrm{B}})^2T^2}e^{\mathrm{i}(\Omega_{\mathrm{A}}-|\boldsymbol{k}|)t_1}e^{-\frac{(t_1+t_{\mathrm{BA}})^2}{T^2}}\left[1+\mathrm{erf}\left(\frac{t_1}{T}-\mathrm{i}\frac{T}{2}(|\boldsymbol{k}|+\Omega_{\mathrm{B}})\right)\right]\right.\\ &\left.+e^{\mathrm{i}(\Omega_{\mathrm{A}}+\Omega_{\mathrm{B}})t_{\mathrm{A}}}e^{-\frac{1}{4}(|\boldsymbol{k}|+\Omega_{\mathrm{A}})^2T^2}e^{\mathrm{i}(\Omega_{\mathrm{B}}-|\boldsymbol{k}|)t_1}e^{-\frac{(t_1-t_{\mathrm{BA}})^2}{T^2}}\left[1+\mathrm{erf}\left(\frac{t_1}{T}-\mathrm{i}\frac{T}{2}(|\boldsymbol{k}|+\Omega_{\mathrm{A}})\right)\right]\right\}\\ =&\sqrt{\pi}\frac{T^2}{2}e^{\mathrm{i}(\Omega_{\mathrm{A}}+\Omega_{\mathrm{B}})t_{\mathrm{B}}}\left[\sqrt{\pi}e^{-\frac{1}{4}(\Omega_{\mathrm{A}}^2+\Omega_{\mathrm{B}}^2+2|\boldsymbol{k}|^2)T^2}e^{\frac{1}{2}|\boldsymbol{k}|(\Omega_{\mathrm{A}}-\Omega_{\mathrm{B}})T^2}e^{\mathrm{i}(|\boldsymbol{k}|-\Omega_{\mathrm{A}})t_{\mathrm{BA}}}\right.\\ &\left.+e^{-\frac{t_{\mathrm{BA}}^2}{T^2}}I\left(\frac{T}{2}(|\boldsymbol{k}|+\Omega_{\mathrm{B}}),T(|\boldsymbol{k}|-\Omega_{\mathrm{A}})+2\mathrm{i}\frac{t_{\mathrm{BA}}}{T}\right)\right]\\ &+\sqrt{\pi}\frac{T^2}{2}\left[\sqrt{\pi}e^{\mathrm{i}(\Omega_{\mathrm{A}}+\Omega_{\mathrm{B}})t_{\mathrm{B}}}e^{-\frac{1}{4}(\Omega_{\mathrm{A}}^2+\Omega_{\mathrm{B}}^2+2|\boldsymbol{k}|^2)T^2}e^{\frac{1}{2}|\boldsymbol{k}|(\Omega_{\mathrm{B}}-\Omega_{\mathrm{A}})T^2}e^{-\mathrm{i}(\Omega_{\mathrm{A}}+|\boldsymbol{k}|)t_{\mathrm{BA}}}\right.\\ &\left.+e^{\mathrm{i}(\Omega_{\mathrm{A}}+\Omega_{\mathrm{B}})t_{\mathrm{A}}}e^{-\frac{t_{\mathrm{BA}}^2}{T^2}}I\left(\frac{T}{2}(|\boldsymbol{k}|+\Omega_{\mathrm{A}}),T(|\boldsymbol{k}|+\Omega_{\mathrm{A}})+\mathrm{i}\frac{t_{\mathrm{BA}}}{T}\right)\right]\\ =&\frac{1}{2}\pi T^2e^{\frac{1}{4}\left(-2|\boldsymbol{k}|^2T^2+2|\boldsymbol{k}|\left[T^2(\Omega_{\mathrm{A}}-\Omega_{\mathrm{B}})+2\mathrm{i}t_{\mathrm{BA}}\right]-T^2\left(\Omega_{\mathrm{A}}^2+\Omega_{\mathrm{B}}^2\right)+4\mathrm{i}t_{\mathrm{B}}(\Omega_{\mathrm{A}}+\Omega_{\mathrm{B}})-4\mathrm{i}t_{\mathrm{BA}}\Omega_{\mathrm{A}}\right)}\\ &\times\left[\mathrm{erfc}\left(\frac{2t_{\mathrm{BA}}+\mathrm{i}T^2(2|\boldsymbol{k}|-\Omega_{\mathrm{A}}+\Omega_{\mathrm{B}})}{2\sqrt{2}T}\right)\right.\\ &\left.+e^{-|\boldsymbol{k}|\left(T^2(\Omega_{\mathrm{A}}-\Omega_{\mathrm{B}})+2\mathrm{i}t_{\mathrm{BA}}\right)}\mathrm{erfc}\left(\frac{-2t_{\mathrm{BA}}+\mathrm{i}T^2(2|\boldsymbol{k}|+\Omega_{\mathrm{A}}-\Omega_{\mathrm{B}})}{2\sqrt{2}T}\right)\right],\end{aligned}\tag{A.57}$$

where

$$I(a,b) = \int_{-\infty}^{\infty} \mathrm{d}x\, e^{-a^2 - \mathrm{i}bx - x^2} \mathrm{erf}(x - \mathrm{i}a) = -\mathrm{i}\sqrt{\pi} e^{-a^2 - \frac{b^2}{4}} \mathrm{erf}\left(\frac{a + \frac{b}{2}}{\sqrt{2}}\right) \quad \text{(A.58)}$$

is explicitly computed in [PKMM15].

A.4 Integration over angular variables of $\mathcal{M}$

In this appendix we perform the integrations in the generalized solid angle variables of the vector $\boldsymbol{k}$ that appear in the nonlocal term $\mathcal{M}$ in Eq. (2.62), namely

$$\int \mathrm{d}\Omega_{n-1}\, e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_\mathrm{A} - \boldsymbol{x}_\mathrm{B})}, \quad \text{(A.59)}$$

to compare the result to the corresponding integrals in the local term $\mathcal{L}$, which evaluate to the area of the $(n-1)$-sphere, i.e.,

$$\int \mathrm{d}\Omega_{n-1} = A_{n-1} = \frac{2\pi^{\frac{n}{2}}}{\Gamma\left(n/2\right)}. \quad \text{(A.60)}$$

In n dimensions there are $n-1$ angular variables, one of which (the polar angle ϕ_{n-1}) has the range $[0, 2\pi)$ and the rest (the azimuthal angles $\phi_1, \ldots \phi_{n-2}$) have range $[0, \pi]$. The solid angle element is therefore

$$\mathrm{d}\Omega_{n-1} = \sin^{n-2}(\phi_1) \sin^{n-3}(\phi_2) \ldots \sin(\phi_{n-2}) \mathrm{d}\phi_1 \mathrm{d}\phi_2 \ldots \mathrm{d}\phi_{n-1}. \quad \text{(A.61)}$$

Let us begin with the particularly simple case of $n = 2$ for illustration. Choosing the x axis of the integration frame to align with $\boldsymbol{x}_\mathrm{A} - \boldsymbol{x}_\mathrm{B}$, the integral easily evaluates to (see 10.9.4 and 10.16.9 in [OOL$^+$])

$$\begin{aligned} \int_0^{2\pi} \mathrm{d}\phi_1\, e^{\mathrm{i}|\boldsymbol{k}||\boldsymbol{x}_\mathrm{A} - \boldsymbol{x}_\mathrm{B}|\cos\phi_1} &= 2\pi J_0(|\boldsymbol{k}||\boldsymbol{x}_\mathrm{A} - \boldsymbol{x}_\mathrm{B}|) \\ &= 2\pi\, {}_0F_1\left(1; -\frac{(|\boldsymbol{k}||\boldsymbol{x}_\mathrm{A} - \boldsymbol{x}_\mathrm{B}|)^2}{4}\right), \end{aligned} \quad \text{(A.62)}$$

where ${}_0F_1$ is the confluent hypergeometric limit function.

In fact, the general case is not too difficult to compute either. In n spatial dimensions, one can choose to place one of the axes of the integration frame aligned with $\boldsymbol{x}_\mathrm{A} - \boldsymbol{x}_\mathrm{B}$, which simplifies the scalar product in the exponential to,

for instance, $|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|\cos(\phi_1)$. With this choice, the integrals evaluate to

$$\begin{aligned}\int \mathrm{d}\Omega_{n-1}\, e^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}})} &= 2\pi \prod_{m=2}^{n-2} \frac{\sqrt{\pi}\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m+1}{2}\right)} \int_0^{\pi} \mathrm{d}\phi_1\, \sin^{n-2}(\phi_1) e^{\mathrm{i}|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|\cos(\phi_1)} \\ &= 2\pi \left(\prod_{m=2}^{n-2} \frac{\sqrt{\pi}\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m+1}{2}\right)}\right) \sqrt{\pi}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} {}_0F_1\left(\frac{n}{2}; -\frac{(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)^2}{4}\right) \\ &= \frac{2\pi^{\frac{n}{2}}}{\Gamma\left(n/2\right)} {}_0F_1\left(\frac{n}{2}; -\frac{(|\boldsymbol{k}||\boldsymbol{x}_{\mathrm{A}}-\boldsymbol{x}_{\mathrm{B}}|)^2}{4}\right). \end{aligned} \tag{A.63}$$

For obtaining this expression, we have used again 10.9.4 and 10.16.9 in [OOL$^+$], noting that

$$\prod_{i=k}^{l} f_i \coloneqq 1 \qquad \text{for } l < k, \tag{A.64}$$

for any set of objects $\{f_i\}_i$, and that

$$\prod_{m=2}^{n-2} \frac{\sqrt{\pi}\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m+1}{2}\right)} = \begin{cases} 1 & n \leq 3 \\ \frac{\pi^{\frac{n-3}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} & n \geq 3 \end{cases}. \tag{A.65}$$

A.5 Time integrals in the overlapping case

In this appendix we examine the time integrals in the local term Eq. (2.53), given by

$$\begin{aligned}\mathcal{T}_{\mathcal{L}}(|\boldsymbol{k}|, T) &= \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2\, e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)} \mathcal{X}\left(t_1\right) \mathcal{X}\left(t_2\right) \\ &= 2\pi |\tilde{\mathcal{X}}(|\boldsymbol{k}|)|^2, \end{aligned} \tag{A.66}$$

where the tilde denotes the Fourier transform in the notation of Eq. (2.60). We show that

$$\mathcal{T}_{\mathcal{L}}(|\boldsymbol{k}|, T) = \mathrm{Re}\left[\mathcal{T}_0(|\boldsymbol{k}|, T)\right], \tag{A.67}$$

where $\mathcal{T}_{t_{\mathrm{BA}}}(|\boldsymbol{k}|, T)$ is given by Eq. (2.65).

To begin with, we see that for $t_{\mathrm{BA}} = 0$ the two summands of Eq. (2.65) coincide, leading to

$$\mathcal{T}_0(|\boldsymbol{k}|, T) = 2\int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2\, e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)} \mathcal{X}\left(t_1\right) \mathcal{X}\left(t_2\right) \theta(t_1 - t_2), \tag{A.68}$$

where $\theta(x)$ is the Heaviside step function. Using the identity $1=\theta(x)+\theta(-x)$ and performing the change of variables $t_1 \leftrightarrow t_2$ in the second summand the result follows:

$$\begin{aligned}\mathcal{T}_{\mathcal{L}}(|\boldsymbol{k}|,T) &= \int_{-\infty}^{\infty}\mathrm{d}t_1 \int_{-\infty}^{\infty}\mathrm{d}t_2\, e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\mathcal{X}\,(t_1)\,\mathcal{X}\,(t_2)\,[\theta(t_1-t_2)+\theta(t_2-t_1)] \\ &= \int_{-\infty}^{\infty}\mathrm{d}t_1 \int_{-\infty}^{\infty}\mathrm{d}t_2\,\mathcal{X}\,(t_1)\,\mathcal{X}\,(t_2)\,\theta(t_1-t_2)\left(e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}+e^{\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\right) \\ &= \int_{-\infty}^{\infty}\mathrm{d}t_1 \int_{-\infty}^{\infty}\mathrm{d}t_2\,\mathcal{X}\,(t_1)\,\mathcal{X}\,(t_2)\,\theta(t_1-t_2)2\mathrm{Re}\left(e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)}\right) \\ &= \mathrm{Re}\left[\mathcal{T}_0(|\boldsymbol{k}|,T)\right]. \end{aligned} \tag{A.69}$$

A.6 Evaluation of $\mathrm{Re}\,\mathcal{T}_{t_{\mathrm{BA}}}$

In this appendix we show that Eq. (2.65) leads to Eq. (2.68). Starting from Eq. (2.65) and changing variables by $t_1=t_2+s$ gives

$$\mathcal{T}_{t_{\mathrm{BA}}}(|\boldsymbol{k}|,T)=\int_{-\infty}^{\infty}\mathrm{d}t_2 \int_{0}^{\infty}\mathrm{d}s\, e^{-\mathrm{i}|\boldsymbol{k}|s}\left[\mathcal{X}\,(t_2+s)\,\mathcal{X}\,(t_2-t_{\mathrm{BA}})+\mathcal{X}\,(t_2+s-t_{\mathrm{BA}})\,\mathcal{X}\,(t_2)\right]. \tag{A.70}$$

Changing variables in the first summand by $\mu=t_2+s$ and renaming $\mu=t_2$ in the second summand, we obtain

$$\mathcal{T}_{t_{\mathrm{BA}}}(|\boldsymbol{k}|,T) = \int_{-\infty}^{\infty}\mathrm{d}\mu\,\mathcal{X}\,(\mu)\int_{0}^{\infty}\mathrm{d}s\, e^{-\mathrm{i}|\boldsymbol{k}|s}\left[\mathcal{X}\,(\mu-s-t_{\mathrm{BA}})+\mathcal{X}\,(\mu+s-t_{\mathrm{BA}})\right]. \tag{A.71}$$

And now, the real part of this expression is

$$\begin{aligned}\mathrm{Re}\left[\mathcal{T}_{t_{\mathrm{BA}}}(|\boldsymbol{k}|,T)\right] &= \int_{-\infty}^{\infty}\mathrm{d}\mu\,\mathcal{X}\,(\mu)\int_{0}^{\infty}\mathrm{d}s\cos(|\boldsymbol{k}|s)\left[\mathcal{X}\,(\mu-s-t_{\mathrm{BA}})+\mathcal{X}\,(\mu+s-t_{\mathrm{BA}})\right] \\ &= \frac{1}{2}\int_{-\infty}^{\infty}\mathrm{d}\mu\,\mathcal{X}\,(\mu)\int_{-\infty}^{\infty}\mathrm{d}s\, e^{\mathrm{i}|\boldsymbol{k}|s}\left[\mathcal{X}\,(\mu-s-t_{\mathrm{BA}})+\mathcal{X}\,(\mu+s-t_{\mathrm{BA}})\right] \\ &= \frac{\sqrt{2\pi}}{2}\int_{-\infty}^{\infty}\mathrm{d}\mu\,\mathcal{X}\,(\mu)\left[e^{\mathrm{i}|\boldsymbol{k}|(\mu-t_{\mathrm{BA}})}[\tilde{\mathcal{X}}\,(|\boldsymbol{k}|)]^{*}+e^{\mathrm{i}|\boldsymbol{k}|(-\mu+t_{\mathrm{BA}})}\tilde{\mathcal{X}}\,(|\boldsymbol{k}|)\right] \\ &= \pi\left[|\tilde{\mathcal{X}}(|\boldsymbol{k}|)|^2 e^{-\mathrm{i}|\boldsymbol{k}|t_{\mathrm{BA}}}+|\tilde{\mathcal{X}}(|\boldsymbol{k}|)|^2 e^{\mathrm{i}|\boldsymbol{k}|t_{\mathrm{BA}}}\right] \\ &= 2\pi|\tilde{\mathcal{X}}(|\boldsymbol{k}|)|^2\cos(|\boldsymbol{k}|t_{\mathrm{BA}}), \end{aligned} \tag{A.72}$$

where the second equality uses the evenness of $\mathcal{X}\,(\mu-s-t_{\mathrm{BA}})+\mathcal{X}\,(\mu+s-t_{\mathrm{BA}})$ in s.

A.7 Regularizations of instantaneous switching

In this appendix we present two regularizations of the Dirac delta that are "kink" limits of switchings largely employed in past literature [PKMM15] that lead to Eq. (2.79). For notational simplicity, we set $t_{\mathrm{A}} = 0$ and consider the formal expression

$$\mathcal{T}_0(|\boldsymbol{k}|) = 2\int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{t_1} \mathrm{d}t_2\, e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)} e^{\mathrm{i}\Omega(t_1+t_2)} \delta(t_1)\delta(t_2), \tag{A.73}$$

showing that each of the regularizations gives for $\mathcal{T}_0(|\boldsymbol{k}|)$ the value of unity.

Top-hat regularization

First, we regard the Dirac delta as a limit of the top-hat function,

$$\delta(t) = \lim_{\epsilon\to 0^+} \frac{1}{\epsilon} \begin{cases} 1 & \text{if } t \in \left[-\frac{\epsilon}{2}, \frac{\epsilon}{2}\right] \\ 0 & \text{otherwise} \end{cases}. \tag{A.74}$$

Then

$$\begin{aligned}
\mathcal{T}_0(|\boldsymbol{k}|) &= 2\int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{t_1} \mathrm{d}t_2\, e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)} e^{\mathrm{i}\Omega(t_1+t_2)} \delta(t_1)\delta(t_2) \\
&= 2\lim_{\epsilon\to 0^+} \lim_{\epsilon'\to 0^+} \frac{1}{\epsilon\epsilon'} \int_{-\epsilon/2}^{\epsilon/2} \mathrm{d}t_1 \int_{-\epsilon'/2}^{t_1} \mathrm{d}t_2\, e^{-\mathrm{i}|\boldsymbol{k}|(t_1-t_2)} e^{\mathrm{i}\Omega(t_1+t_2)} \\
&= 2\lim_{\epsilon\to 0^+} \lim_{\epsilon'\to 0^+} \frac{\mathrm{i}e^{-\frac{1}{2}\mathrm{i}\epsilon'(|\boldsymbol{k}|+\Omega)}}{\epsilon\epsilon'(|\boldsymbol{k}|+\Omega)} \int_{-\epsilon/2}^{\epsilon/2} \mathrm{d}t_1\, e^{-\mathrm{i}(|\boldsymbol{k}|-\Omega)t_1} \left[1 - e^{\frac{1}{2}\mathrm{i}(|\boldsymbol{k}|+\Omega)(2t_1+\epsilon')}\right] \\
&= \lim_{\epsilon\to 0^+} \lim_{\epsilon'\to 0^+} \frac{2\mathrm{i}}{\epsilon\epsilon'} \left[\frac{2e^{-\frac{1}{2}\mathrm{i}\epsilon'(|\boldsymbol{k}|+\Omega)} \sin\left[\frac{1}{2}\epsilon(|\boldsymbol{k}|-\Omega)\right]}{|\boldsymbol{k}|^2-\Omega^2} - \frac{\sin(\Omega\epsilon)}{|\boldsymbol{k}|\Omega+\Omega^2}\right] \\
&= \lim_{\epsilon\to 0^+} \lim_{\epsilon'\to 0^+} \frac{2\mathrm{i}}{\epsilon\epsilon'} \frac{(-\mathrm{i}\epsilon\epsilon')}{2} \\
&= 1.
\end{aligned} \tag{A.75}$$

Gaussian regularization

Second, we regard the Dirac delta as a limit of the Gaussian function,

$$\delta(t) = \lim_{\epsilon\to 0^+} \frac{1}{2\epsilon\sqrt{\pi}} e^{-\frac{t^2}{4\epsilon^2}}. \tag{A.76}$$

Then

$$\begin{aligned}\mathcal{T}_0(|\boldsymbol{k}|) &= \lim_{\epsilon\to 0^+}\lim_{\epsilon'\to 0^+}\frac{1}{2\pi\epsilon\epsilon'}\int_{-\infty}^{\infty}\mathrm{d}t_1\, e^{-\mathrm{i}(|\boldsymbol{k}|-\Omega)t_1}e^{-\frac{t_1^2}{4\epsilon^2}}\int_{-\infty}^{t_1}\mathrm{d}t_2\, e^{\mathrm{i}(|\boldsymbol{k}|+\Omega)t_2}e^{-\frac{t_2^2}{4\epsilon'^2}}\\ &= \lim_{\epsilon\to 0^+}\lim_{\epsilon'\to 0^+}\frac{e^{-\epsilon'^2(|\boldsymbol{k}|+\Omega)^2}}{2\sqrt{\pi}\epsilon}\\ &\quad\times\int_{-\infty}^{\infty}\mathrm{d}t_1\, e^{-\mathrm{i}(|\boldsymbol{k}|-\Omega)t_1}e^{-\frac{t_1^2}{4\epsilon^2}}\left[1+\operatorname{erf}\left(\frac{t_1}{2\epsilon'}-\mathrm{i}\epsilon'(|\boldsymbol{k}|+\Omega)\right)\right]. \quad \text{(A.77)}\end{aligned}$$

This is exactly Eq. (A2) in Appendix A of [PKMM15]. As shown there, the remaining integral has a closed-form expression, which yields

$$\begin{aligned}\mathcal{T}_0(|\boldsymbol{k}|) &= \lim_{\epsilon\to 0^+}\lim_{\epsilon'\to 0^+} e^{-\epsilon^2(|\boldsymbol{k}|+\Omega)^2}e^{-\epsilon'^2(|\boldsymbol{k}|-\Omega)^2}\left[1+\operatorname{erf}\left(\mathrm{i}\frac{\epsilon(|\boldsymbol{k}|-\Omega)+\epsilon'(|\boldsymbol{k}|+\Omega)}{\sqrt{2}}\right)\right]\\ &= 1. \quad \text{(A.78)}\end{aligned}$$

Appendices of Chapter 3

B.1 Adiabat without mass change

Instead of performing the Hamiltonian swap in Eq. (3.29), which is equivalent to simultaneously quenching both the frequency and the mass of the WM, in this appendix we explore the possibility of changing only the frequency of the WM. This is, doing the transformation

$$H_m = \frac{P_m^2}{2\mu} + \frac{\mu\omega_m^2 Q_m^2}{2} \longleftrightarrow H'_m = \frac{P_m^2}{2\mu} + \frac{\mu(\omega'_m)^2 Q_m^2}{2}, \tag{B.1}$$

where Q_m and P_m are the canonical position and momentum of the WM, and μ is its mass. In this case, the quadratures and the creation-annihilation operators change as well:

$$q'_m = Q_m\sqrt{\mu\omega'_m},\;\; a'_m = \sqrt{\frac{\mu\omega'_m}{2}}\left(Q_m + \frac{\mathrm{i}}{\mu\omega'_m}P_m\right). \tag{B.2}$$

In these terms, the change in Eq. (B.1) takes the form

$$H_m = \omega_m a_m^\dagger a_m \longleftrightarrow H'_m = \omega'_m (a'_m)^\dagger a'_m. \tag{B.3}$$

Now, if this change is performed instantaneously, the initial thermal state of the WM, $\rho \propto e^{-H_m/T_m}$, will remain unchanged. Since $[H'_m, H_m] \neq 0$, this will result in the state having coherences in the new energy eigenbasis [KR17], significantly reducing the efficiency of the engine (this can be straightforwardly deduced from the analysis that will be presented in Appendix B.2).

If, on the other hand, as a result of Eq. (B.1), the state were also changed to

$$\rho' \propto \exp\left(-\omega_m (a'_m)^\dagger a'_m / T_m\right), \tag{B.4}$$

its covariance matrix, as defined by Eq. (3.1) with $\boldsymbol{x}'_m$ instead of $\boldsymbol{x}_m$, would remain unchanged. If now the interaction Hamiltonian, H_{int} [see Eq. (3.23)], would couple to the bath degrees of freedom with the new quadrature, q'_m, instead of q_m, the dynamics of the overall covariance matrix would be the same as that presented in the main text.

Put in other words, were one to change to the new quadratures in all formulas, the dynamics on the level of covariance matrices and Hamiltonian matrices would remain unchanged. Hence, the results for such an engine would be exactly the same that those shown in Chapter 3.

For such a modification to work, one needs to have ρ evolving into ρ' as a result of Eq. (B.3). Since the instantaneous change in Hamiltonian cannot produce any changes in the state, let us allow the change to take some nonzero time τ_{ad}. The problem is now the following: is it possible to devise a quadratic Hamiltonian path connecting H_m to H'_m that is capable of evolving the state in time τ_{ad} so that the covariance matrix remains the same?

To answer the above question, let us notice that the covariance matrix remains unchanged as a result of quantum adiabatic evolution. The latter is defined by the unitary evolution operator

$$U = \sum_n |n\rangle_{H'_m} \langle n|_{H_m}, \tag{B.5}$$

where $|n\rangle_{H_m}$ and $|n\rangle_{H'_m}$ are the n^{th} eigenvalues of, respectively, H_m and H'_m (see, for instance, [MCI$^+$10]). Indeed, we have that

$$\begin{aligned}\sigma'_{ab} &= \mathrm{Tr}[U\rho U^\dagger (x'_a x'_b + x'_b x'_a)] \\ &= \mathrm{Tr}[\rho (x_a x_b + x_b x_a)],\end{aligned} \tag{B.6}$$

where we used the fact that

$$\begin{aligned}U^\dagger \boldsymbol{x}' U &= \sum_{n,k} |n\rangle_{H_m} \langle k|_{H_m} \cdot \langle n|_{H'_m} \boldsymbol{x}' |k\rangle_{H'_m} \\ &= \sum_{n,k} |n\rangle_{H_m} \langle k|_{H_m} \cdot \langle n|_{H_m} \boldsymbol{x} |k\rangle_{H_m} \\ &= \boldsymbol{x}.\end{aligned} \tag{B.7}$$

Now, as is shown in [MCI$^+$10], for a single oscillator, a shortcut to adiabaticity can be constructed for implementing the unitary (B.5) by adding the time-dependent term

$$H_I(t) = -\frac{\dot{\omega}_m(t)}{4\omega_m(t)}(QP + PQ) \tag{B.8}$$

to the Hamiltonian of the oscillator for the period of the adiabat τ_{ad}. Here, the function $\omega_m(t)$ is arbitrary provided that it satisfies $\omega_m(0) = \omega_m$ and $\omega_m(\tau_{ad}) = \omega'_m$.

In order for this new cycle to coincide with the one in the main text, we need τ_{ad} to approach to zero. This would require very fast generation of the term (B.8)

and very quick driving of the frequency. Although we leave the question of the experimental accessibility of such quick driving open, we note that the programs provided in the computational appendix [PKHB17] straightforwardly allow for simulating an Otto cycle with any nonzero τ_{ad}.

B.2 Full-cycle energetics

In this appendix we perform a detailed analysis of the work and heat involved in the perfect cycles of operation of the WM. As in the main text, we take as starting point the moment of the cycle where the WM is still in a state $\rho \propto e^{-\omega_c a_m^\dagger a_m / T_c^{(0)}}$ but its Hamiltonian has already been changed to $\omega_h a_m^\dagger a_m$, and the isochoric interaction with the hot bath is about to start. We label the initial temperature of the bath with superscript $^{(0)}$ to indicate that it was its temperature before the first cycle started. At this moment, the total Hamiltonian is just the sum of the individual Hamiltonians, and the energy before starting the isochore is thus

$$E^{(in)} = E_B^{(in)} + E_{WM}^{(in)} = E_B^{(in)} + \omega_h \, n\left(\frac{\omega_c}{T_c^{(0)}}\right), \tag{B.9}$$

where we have defined

$$n(x) = \frac{1}{e^x - 1}. \tag{B.10}$$

After the isochore, the total Hamiltonian returns to being the sum of the individual Hamiltonians and the state of the WM is again thermal (albeit now correlated with the bath). Hence, the energy of the system right after the isochore is

$$E^{(fin)} = E_B^{(fin)} + \omega_h \, n\left(\frac{\omega_h}{T_h^{(1)}}\right), \tag{B.11}$$

where $T_h^{(1)}$ is the temperature of the WM after the interaction with the hot bath. Recall that, as discussed in Section 3.4.1, although the parameters can be chosen so that $T_h^{(1)} = T_h$ exactly, this temperature does not need to be exactly the temperature of the bath T_h. In the example in the main text, namely, when $N_c = N_h = 30$, $\omega_c = 1$, $\omega_h = 2$, $T_c = 0.5$, $T_h = 4$ and $\tau = 100$, $T_h^{(1)}$ is slightly greater than T_h: $T_h^{(1)} - T_h \approx 1.3 \times 10^{-4}$. This small difference does not affect our analysis because such deviations, if not appearing in the first cycle, do appear in the subsequent ones. The work extracted during the hot isochore is then

$$W_{ih} = E^{(in)} - E^{(fin)} = Q + \omega_h \left[n\left(\frac{\omega_c}{T_c^{(0)}}\right) - n\left(\frac{\omega_h}{T_h^{(1)}}\right) \right], \tag{B.12}$$

where $Q = E_B^{(in)} - E_B^{(fin)}$.

Rearranging and adding the superscript $^{(1)}$ to indicate that the labeled quantities correspond to the end of the first cycle, the following expression for the heat exchanged during the cycle is obtained:

$$Q^{(1)} = \omega_h \left[n\left(\frac{\omega_h}{T_h^{(1)}}\right) - n\left(\frac{\omega_c}{T_c^{(0)}}\right)\right] + W_{ih}^{(1)}. \tag{B.13}$$

The following step in the cycle is the adiabatic expansion. The work extracted from this process is

$$W_{h\to c}^{(1)} = (\omega_h - \omega_c)\, n\left(\frac{\omega_h}{T_h^{(1)}}\right). \tag{B.14}$$

Next, during the cold isochoric interaction, an amount of work $W_{ic}^{(1)}$ is extracted, leaving the system at temperature $T_c^{(1)}$ (which, again, is slightly different from $T_c^{(0)} = T_c$. For the case studied in Section 3.4, $T_c^{(1)} - T_c^{(0)} \approx 7.7 \times 10^{-3}$).

Finally, during the adiabatic compression, one extracts the negative amount of work

$$W_{c\to h}^{(1)} = -(\omega_h - \omega_c)\, n\left(\frac{\omega_c}{T_c^{(1)}}\right). \tag{B.15}$$

The final state of the WM will be $\rho \propto e^{-\omega_c a_m^\dagger a_m / T_c^{(1)}}$, which explicitly shows that the WM is not completely cyclic. However, the deviations from cyclicity, in the case discussed in the main text, are of $\mathcal{O}(e^{-\omega_c/T_c^{(0)}} - e^{-\omega_c/T_c^{(1)}}) = \mathcal{O}(10^{-3})$, the same order of magnitude of the degradation during the perfect cycles.

The total work output of the cycle is the sum of the outputs in every step, that is

$$W^{(1)} = (\omega_h - \omega_c)\left[n\left(\frac{\omega_h}{T_h^{(1)}}\right) - n\left(\frac{\omega_c}{T_c^{(1)}}\right)\right] + W_{ih}^{(1)} + W_{ic}^{(1)}, \tag{B.16}$$

and thus the efficiency, $\eta^{(1)} = W^{(1)}/Q^{(1)}$, will amount to

$$\eta^{(1)} = \eta_O + \frac{\omega_c W_{ih}^{(1)} + \omega_h W_{ic}^{(1)}}{\omega_h Q^{(1)}} - \frac{\omega_h - \omega_c}{Q^{(1)}}\left[n\left(\frac{\omega_c}{T_c^{(1)}}\right) - n\left(\frac{\omega_c}{T_c^{(0)}}\right)\right], \tag{B.17}$$

where

$$\eta_O = 1 - \frac{\omega_c}{\omega_h} \tag{B.18}$$

is the maximal theoretical efficiency for the Otto engine in which the WM couples negligibly weakly to infinite, Markovian baths [RK06, KR17]. Since $T_c^{(1)} - T_c^{(0)} \ll 1$, whenever η_O is away from the Carnot value, namely, when

$$\frac{\omega_h}{T_h^{(1)}} < \frac{\omega_c}{T_c^{(0)}} \tag{B.19}$$

so that

$$Q^{(1)} \gg \max\left\{W_{ih}^{(1)}, W_{ic}^{(1)}\right\}, \tag{B.20}$$

the efficiency $\eta^{(1)}$ will be very close to η_O. In the examples studied in the chapter, both the hot and cold isochoric works are of the order of 10^{-3}, while $Q^{(1)} \approx 3$.

By the moment the second hot isochore (and hence the second cycle) is about to start, the perturbations created in the hot bath during the first isochore will have traveled away from the interaction node in the form of a wavepacket. However, the propagation of this wavepacket is not ideal in that it leaves a trace in the form of residual perturbations. In particular, before the beginning of the second isochore, the state of the interaction node will be slightly different from that at equilibrium. This means that $T_h^{(2)}$ will be even further from T_h than $T_h^{(1)}$. As our numerical analysis shows for perfect cycles, and as it is to be expected from the fact that the hot isochore extracts heat from the bath,

$$T_h \equiv T_h^{(0)} \approx T_h^{(1)} \gtrsim T_h^{(2)} \gtrsim \cdots . \tag{B.21}$$

With a similar reasoning, another result that we observe numerically is that

$$T_c \equiv T_c^{(0)} \approx T_c^{(1)} \lesssim T_c^{(2)} \lesssim \cdots . \tag{B.22}$$

Moreover, since (while within "perfect" cycles) the WM is thermal after each interaction with the baths, for the k^{th} perfect cycle one has that

$$Q^{(k)} = \omega_h \left[n\left(\frac{\omega_h}{T_h^{(k)}}\right) - n\left(\frac{\omega_c}{T_c^{(k-1)}}\right)\right] + W_{ih}^{(k)} \tag{B.23}$$

and

$$W^{(k)} = (\omega_h - \omega_c)\left[n\left(\frac{\omega_h}{T_h^{(k)}}\right) - n\left(\frac{\omega_c}{T_c^{(k)}}\right)\right] + W_{ih}^{(k)} + W_{ic}^{(k)}. \tag{B.24}$$

Therefore, the efficiency of the k^{th} "perfect" cycle is given by

$$\eta^{(k)} = \eta_O + \frac{\omega_c W_{ih}^{(k)} + \omega_h W_{ic}^{(k)}}{\omega_h Q^{(k)}} - \frac{\omega_h - \omega_c}{Q^{(k)}}\left[n\left(\frac{\omega_c}{T_c^{(k)}}\right) - n\left(\frac{\omega_c}{T_c^{(k-1)}}\right)\right]. \tag{B.25}$$

Along with Eqs. (B.21) and (B.22), Eqs. (B.23) and (B.24) explain the slow, gradual decrease of cycle heat and work during the period of perfect operation (see Fig. 3.5a). At the same time, Eq. (B.25) explains the reason why the efficiency does not accumulate errors and stays very close to η_O throughout the perfect performance. Indeed, as mentioned above, the isochoric works, $W_{ih}^{(k)}$ and $W_{ic}^{(k)}$, stay of the order of α^2 and the acyclicity, as given by $T_c^{(k)} - T_c^{(k-1)}$, being an effect a *single* WM-bath interaction session has on the bath, remains almost unchanged throughout the perfect cycles and is small compared to the cycle heat. Another important consequence of Eq. (B.25) is that Eq. (3.31) needs to be slightly modified for $k \geq 2$. Indeed, although $\eta_{\alpha=0}^{(k)} - \eta^{(k)} \propto \alpha^2$ still holds, one needs additionally account for

$$\eta_O - \eta_{\alpha=0}^{(k)} = \mathcal{O}(T_c^{(k)} - T_c^{(k-1)}), \tag{B.26}$$

and if, for the first cycle, this term can be eliminated by adjusting the interaction time, it will be nonzero for the subsequent cycles. However, as mentioned above, the correction (B.26) is very small and does not increase as the cycles proceed.

Lastly, let us remark that whenever the machine approaches the Carnot efficiency $\eta_C = 1 - T_c/T_h$—namely, when ω_c/ω_h approaches T_c/T_h from above—the work output of the engine tends to zero [as can be seen from Eq. (B.24)] and, exactly at the point when $\eta_O = \eta_C$, $W^{(k)} < 0$.

B.3 Relative entropy

In order to show the evolution of the states of the baths, we choose the relative entropy as a distance quantifier. Take, for example, the hot bath, denoting its state at the beginning of the i^{th} cycle by $\rho_h^{(i)}$. Then, the quantity of interest is

$$S(\rho_h^{(i)}||\rho_h^{(1)}) = S\left(\rho_h^{(i)}\middle|\middle|\frac{1}{Z_h}e^{-H_h/T_h}\right), \tag{B.27}$$

where H_h is the Hamiltonian of the hot bath, $Z_h = \operatorname{Tr} e^{-H_h/T_h}$, and [NC10]

$$S(\rho||\sigma) = \operatorname{Tr}\left[\rho(\ln\rho - \ln\sigma)\right]. \tag{B.28}$$

The relative entropy has several features desirable for a distance quantifier. In particular, those that are of interest for the case studied are that

$$S(\rho||\sigma) \geq 0, \tag{B.29}$$

i.e., that it is a nonnegative quantity, and that

$$S(\rho||\sigma) = 0 \quad \text{iff} \quad \rho = \sigma. \tag{B.30}$$

Although the relative entropy does not satisfy the triangle inequality and is not symmetric, which means it is not a distance measure in the proper sense, it decreases monotonically under completely positive trace preserving operations [NC10], which makes it a distinguishability measure of choice in many situations.

Whenever the second argument in $S(\bullet||\bullet)$ is a Gibbs state, as, e.g., is the case in Eq. (B.27),

$$S(\rho_h^{(i)}||\rho_h^{(1)}) = T_h^{-1}\left(E_h^{(i)} - E_h^{(1)}\right) - S(\rho_h^{(i)}) + S(\rho_h^{(1)}) \tag{B.31}$$

$$= T_h^{-1}F[\rho_h^{(i)}] - T_h^{-1}F[\rho_h^{(1)}], \tag{B.32}$$

where S is the von Neumann entropy, $E_h^{(i)}$ is the energy of the hot bath at the beginning of the i^{th} cycle, and $F = E - TS$ is the free energy. Eq. (B.32) means that relative entropy measures the "thermodynamic" distance between $\rho_h^{(i)}$ and $\rho_h^{(1)}$, which additionally motivates our choice of relative entropy as a distance quantifier.

The quantities in Eq. (B.31) can be readily calculated directly in the covariance-matrix picture via Eqs. (3.5), (3.15), and (3.16). An example of such a calculation is presented in Fig. 3.5b, where it can be seen that the relative entropy distance of the state of of the bath from the initial state of the bath increases at a steady rate as the "perfect" cycles proceed. This is contrasted by how the efficiency and work per cycle remain almost constant during these cycles.

Appendices of Chapter 4

C.1 Coherent element-wise operations

In this appendix we give a more formal description of the quantum method for approximately computing the polynomial $P^n_{(\odot,\oslash)}(K^0)$ in Section 4.3. The main results needed are well summarized by the following Lemmas 1 and 2, and Theorem 1.

Lemma 1. *Given $\mathcal{O}(t^2/\epsilon)$ copies of N-qubit density matrices, ρ_1 and ρ_2, let $\rho_1 \odot \rho_2$ denote the Hadamard product between ρ_1 and ρ_2. There exists a quantum algorithm to implement the unitary $e^{-\mathrm{i}\rho_1\odot\rho_2 t}$ on an N-qubit input state, σ, for a time t to accuracy $1-\epsilon$ in operator norm.*

Proof. The usual SWAP matrix employed in quantum principal component analysis [RML14] is given by $S = \sum_{j,k} |j\rangle\langle k| \otimes |k\rangle\langle j|$. Here we modify this operator to $S_\odot = \sum_{j,k} |j\rangle\langle k| \otimes |j\rangle\langle k| \otimes |k\rangle\langle j|$. With an arbitrary input state σ, the operation

$$\mathrm{Tr}_{1,2}\left[e^{-\mathrm{i}S_\odot\delta}(\rho_1 \otimes \rho_2 \otimes \sigma)e^{\mathrm{i}S_\odot\delta}\right] \tag{C.1}$$

can be efficiently performed with a small parameter δ. The symbol $\mathrm{Tr}_{1,2}$ represents the trace over the subspaces of ρ_1 and ρ_2. Expanding Eq. (C.1) to first order in δ leads to

$$\begin{aligned}&\mathrm{Tr}_{1,2}\left[e^{-\mathrm{i}S_\odot\delta}(\rho_1 \otimes \rho_2 \otimes \sigma)e^{\mathrm{i}S_\odot\delta}\right]\\ &\quad= \mathbb{1} - \mathrm{i}\,\mathrm{Tr}_{1,2}[S_\odot(\rho_1 \otimes \rho_2 \otimes \sigma)]\delta + \mathrm{i}\,\mathrm{Tr}_{1,2}[(\rho_1 \otimes \rho_2 \otimes \sigma)S_\odot]\delta + \mathcal{O}(\delta^2).\end{aligned} \tag{C.2}$$

Examining the first element linear in the parameter δ reveals

$$\begin{aligned}
&\mathrm{Tr}_{1,2}[S_{\odot}(\rho_1 \otimes \rho_2 \otimes \sigma)] \\
&\quad = \mathrm{Tr}_{1,2}\left[\sum_{j,k} |j\rangle\langle k| \otimes |j\rangle\langle k| \otimes |k\rangle\langle j| \, (\rho_1 \otimes \rho_2 \otimes \sigma)\right] \\
&\quad = \sum_{n,m,j,k} \langle n|j\rangle \langle k|\rho_1|n\rangle \langle m|j\rangle \langle k|\rho_2|m\rangle \, |k\rangle\langle j| \, \sigma \\
&\quad = \sum_{j,k} \langle k|\rho_1|j\rangle \langle k|\rho_2|j\rangle \, |k\rangle\langle j| \, \sigma \\
&\quad = (\rho_1 \odot \rho_2)\, \sigma. \qquad \text{(C.3)}
\end{aligned}$$

In the same manner, we have that

$$\mathrm{Tr}_{1,2}[(\rho_1 \otimes \rho_2 \otimes \sigma) S_{\odot}] = \sigma(\rho_1 \odot \rho_2). \qquad \text{(C.4)}$$

Thus in summary, we have shown that

$$\mathrm{Tr}_{1,2}\left[e^{-\mathrm{i}S_{\odot}\delta}(\rho_1 \otimes \rho_2 \otimes \sigma)e^{\mathrm{i}S_{\odot}\delta}\right] = \sigma - \mathrm{i}[(\rho_1 \odot \rho_2), \sigma]\delta + \mathcal{O}(\delta^2). \qquad \text{(C.5)}$$

The above is equivalent, up to $\mathcal{O}(\delta)$, to applying the unitary $\exp[-\mathrm{i}(\rho_1 \odot \rho_2)\delta]$ to the density matrix σ:

$$\begin{aligned}
e^{-\mathrm{i}(\rho_1 \odot \rho_2)\delta}\sigma e^{\mathrm{i}(\rho_1 \odot \rho_2)\delta} &= [\mathbb{1} - \mathrm{i}(\rho_1 \odot \rho_2)\delta + O(\delta^2)]\, \sigma [\mathbb{1} + \mathrm{i}(\rho_1 \odot \rho_2)\delta + O(\delta^2)] \\
&= \sigma - \mathrm{i}[(\rho_1 \odot \rho_2), \sigma]\delta + \mathcal{O}(\delta^2). \qquad \text{(C.6)}
\end{aligned}$$

The above completes the derivation of Eq. (4.10). Note that if the small time parameter is taken to be $\delta = \epsilon/t$, and the above procedure is implemented $\mathcal{O}(t^2/\epsilon)$ times, the overall effect amounts to implementing the desired operation, $e^{-\mathrm{i}\rho t}\sigma e^{\mathrm{i}\rho t}$ up to an error $\mathcal{O}(\delta^2 t^2/\epsilon) = \mathcal{O}(\epsilon)$, while consuming $\mathcal{O}(t^2/\epsilon)$ copies of ρ_1 and ρ_2. This concludes the proof of Lemma 1. □

Lemma 2. *Given $\mathcal{O}(t^2/\epsilon)$ copies of N-qubit density matrices, ρ_1 and ρ_2, let $\rho_1 \oslash \rho_2$ denote the outer product between the diagonal entries of ρ_1 and ρ_2. There exists a quantum algorithm to implement the unitary $e^{-\mathrm{i}\rho_1 \oslash \rho_2 t}$ on an N-qubit input state, σ, for a time t to accuracy $1-\epsilon$ in operator norm.*

Proof. By changing indices in the modified SWAP operator $S_{\odot}$, one obtains $S_{\oslash} = \sum_{j,k} |j\rangle\langle j| \otimes |k\rangle\langle k| \otimes |k\rangle\langle j|$. Analogously with the proof of Lemma 1, we have that

$$\mathrm{Tr}_{1,2}\left[e^{-\mathrm{i}S_{\oslash}\delta}(\rho_1 \otimes \rho_2 \otimes \sigma)e^{\mathrm{i}S_{\oslash}\delta}\right] = \sigma - \mathrm{i}[(\rho_1 \oslash \rho_2), \sigma]\delta + O(\delta^2). \qquad \text{(C.7)}$$

The above can be compared with

$$e^{-\mathrm{i}(\rho_1 \oslash \rho_2)\delta}\sigma e^{\mathrm{i}(\rho_1 \oslash \rho_2)\delta} = \sigma - \mathrm{i}[(\rho_1 \oslash \rho_2), \sigma]\delta + O(\delta^2). \tag{C.8}$$

The equivalence up to the linear term in δ confirms the validity of Eq. (4.11). Similarly with Lemma 1, with a $\mathcal{O}(t^2/\epsilon)$ repetition consuming $\mathcal{O}(t^2/\epsilon)$ copies of ρ_1 and ρ_2, the desired $e^{-\mathrm{i}\rho t}\sigma e^{\mathrm{i}\rho t}$ can be implemented up to error ϵ. □

Given the density matrix $\rho = K^0$ which encodes the base case covariance matrix, we approximate the nonlinear kernel function at l^{th} layer with the order-n polynomial, $P^n_{(\odot,\oslash)}(\rho) = \sum_r^n c_r \rho^{(\odot,\oslash)r}$. Here the label $(\odot, \oslash)$ indicates that we work in the setting where the types of product operation involved for taking the r^{th} power of ρ are arbitrary combinations of Hadamard and diagonal outer products. Now we are in the position of presenting the main theorem required to implement the kernel function at the l^{th} layer.

Theorem 1. *Given $\mathcal{O}(n^2t^2/\epsilon)$ copies of the N-qubit density matrix ρ, and the order-n polynomial of Hadamard and diagonal outer products*

$$P^n_{(\odot,\oslash)}(\rho) = \sum_r^n c_r \rho^{(\odot,\oslash)r},$$

there exists a quantum algorithm to implement the unitary $e^{-\mathrm{i}P^n_{(\odot,\oslash)}(\rho)t}$ on an N-qubit input state, σ, for a time t to accuracy $1-\epsilon$ in operator norm.

Proof. We first address how to implement the unitary $e^{-\mathrm{i}\rho^{(\odot,\oslash)r}t}$. Intuitively, this can be achieved by constructing a generalized $\tilde{S}$ operator with tensor product components of $|j\rangle\langle j|$, $|j\rangle\langle k|$, $|k\rangle\langle k|$ and $|k\rangle\langle j|$, corresponding to the contributing elements in the matrices in each term. We now give a recursive procedure to determine such $\tilde{S}$.

In the case of $r = 2$, we have already shown in Lemma 1 and Lemma 2 the desired operation can be achieved using $S_\odot$ and $S_\oslash$ corresponding to the $\odot$ and $\oslash$ cases respectively. Thus we can write the base case of the recursive procedure as

$$\tilde{S}^{(r=2)} = \sum_{j,k} T^{(2)}(j,k) \otimes |k\rangle\langle j|,$$

where $T^{(2)}(j,k)$ denotes the possible combinations of tensor products, $|j\rangle\langle k| \otimes |j\rangle\langle k|$ or $|j\rangle\langle j| \otimes |k\rangle\langle k|$. Now, in the case of $r = 3$, the additional factor of ρ will come in two possible cases. If it comes as a $\odot$ product, the updated operator $\tilde{S}^{(r=3)}_\odot$ is simply given by

$$\tilde{S}^{(r=3)}_\odot = \sum_{j,k} T^{(2)}(j,k) \otimes |j\rangle\langle k| \otimes |k\rangle\langle j|.$$

If the additional ρ comes in as a $\oslash$ product, the updated operator $\tilde{S}^{(r=3)}_{\oslash}$ is instead given by

$$\tilde{S}^{(r=3)}_{\oslash} = \sum_{j,k} |j\rangle\langle j| \otimes |j\rangle\langle j| \otimes |k\rangle\langle k| \otimes |k\rangle\langle j| .$$

This can be seen by observing that the contributing elements to a $\oslash$ product are exclusively diagonal, which we use $|j\rangle\langle j|$ to pick up. Any off-diagonal information about the previous element-wise product operations is irrelevant. In general, if we have the r^{th} order $\tilde{S}$ operator given by

$$\tilde{S}^{(r)} = \sum_{j,k} T^{(r)}(j,k) \otimes |k\rangle\langle j| ,$$

the operators $\tilde{S}^{(r+1)}_{\odot}$ and $\tilde{S}^{(r+1)}_{\oslash}$ can be generated as follows:

$$\tilde{S}^{(r+1)}_{\odot} = \sum_{j,k} T^{(r)}(j,k) \otimes |j\rangle\langle k| \otimes |k\rangle\langle j| , \tag{C.9}$$

$$\tilde{S}^{(r+1)}_{\oslash} = \sum_{j,k} (|j\rangle\langle j|)^{\otimes r} \otimes |k\rangle\langle k| \otimes |k\rangle\langle j| . \tag{C.10}$$

Therefore, we have a recursive procedure to construct $\tilde{S}^{(r)}$ up to $r = n$ such that

$$\mathrm{Tr}_{1\ldots r}\left[e^{-\mathrm{i}\tilde{S}^{(r)}\delta}(\rho^{\otimes r} \otimes \sigma)e^{\mathrm{i}\tilde{S}^{(r)}\delta}\right] = e^{-\mathrm{i}\rho^{(\odot,\oslash)r}\delta}\sigma e^{\mathrm{i}\rho^{(\odot,\oslash)r}\delta} + \mathcal{O}(\delta^2), \tag{C.11}$$

for a small evolution δ. Analogously with Lemma 1 and Lemma 2, with a $\mathcal{O}(t^2/\epsilon)$ repetition consuming $\mathcal{O}(rt^2/\epsilon)$ copies of ρ, the desired $e^{-\mathrm{i}\rho^{(\odot,\oslash)r}t}\sigma e^{\mathrm{i}\rho^{(\odot,\oslash)r}t}$ can be implemented up to an error ϵ.

Finally one makes use of the Lie product formula for summing the terms in the polynomial [Suz92, CCD+03, WBHS10]:

$$e^{\mathrm{i}\delta(A+B)+\mathcal{O}(\delta^2/m)} = (e^{\mathrm{i}\delta A/m}e^{\mathrm{i}\delta B/m})^m, \tag{C.12}$$

where A and B are taken to be different terms in $P^n_{(\odot,\oslash)}(\rho) = \sum_r^n c_r\rho^{(\odot,\oslash)r}$, and the factors c_r simply amount to multiplying the $S^{(r)}$ matrices with the respective coefficients. The parameter m can be chosen further suppress the error by repeating the entire procedure. However, for the purpose of implementing $e^{-\mathrm{i}P^n_{(\odot,\oslash)}(\rho)t}\sigma e^{\mathrm{i}P^n_{(\odot,\oslash)}(\rho)t}$ to the desired accuracy $1-\epsilon$, $\mathcal{O}(n^2t^2/\epsilon)$ copies of ρ are required. The quadratic dependency in the order of the polynomial, n^2, stems from implementing the unitary $\exp[-\mathrm{i}\rho^{(\odot,\oslash)r}t]$ up to $r = n$, each consuming $\mathcal{O}(nt^2/\epsilon)$ copies as argued before. □

Appendices of Chapter 5

D.1 Compatibility with classical and mixed scenarios

Both techniques described in Chapter 5, scalar extension and quantum inflation, can be easily adapted for solving problems of causal compatibility with classical causal structures. It is known that any correlation achievable with only classical latent variables can be realized in terms of commuting measurements acting on a quantum state [BCWA17]. Thus, in order to detect correlations incompatible with classical scenarios, one must generalize the commutation relations in Eqs. 5.24a and 5.24b to the constraint that any pair of measurement operators commute. That is,

$$\hat{O}^{\boldsymbol{s}_1|k_1}_{i_1|m_1}\cdot \hat{O}^{\boldsymbol{s}_2|k_2}_{i_2|m_2} = \hat{O}^{\boldsymbol{s}_2|k_2}_{i_2|m_2}\cdot \hat{O}^{\boldsymbol{s}_1|k_1}_{i_1|m_1} \quad \forall\, \boldsymbol{s}_1, \boldsymbol{s}_2, k_1, k_2, m_1, m_2, i_1, i_2. \tag{D.1}$$

The rest of the procedures remain unaltered. In particular, one can define hierarchies of extensions or inflations where one should be able to find a quantum state and (commuting) measurements that reproduce the correlations under scrutiny, and for answering this question an NPO hierarchy in the spirit of that of [NPA07] can be defined.

The associated NPO hierarchies, which now only involve commuting operators, are guaranteed to converge at a finite level. In fact, for hierarchy levels higher than $N \cdot m \cdot (d-1)$—where N is the number of parties, m is the number of settings per party and d is the number of outcomes per measurement—application of the commutation relations allows one to reduce any product of the operators involved into one of shorter length. In the case of quantum inflation, for a fixed inflation level, the problem solved at the highest level of the NPO hierarchy is analogous to the linear program solved in classical inflation [WSF19] at the same inflation level.

The application of quantum inflation to scenarios with classical latent variables is of particular interest, as it conforms a general method to characterize the correlations achievable in any classical causal structure. In contrast with the original classical inflation technique, the classical variant of quantum inflation uses semidefinite programming, and exhibits far more efficient scaling with the inflation hierarchy than the original linear programming approach [NW17].

One must note that this gain in efficiency comes at the expense of introducing further relaxations in the problem. Nevertheless, this classical variant of quantum inflation is capable of recovering a variety of seminal results of classical causal inference, such as the incompatibility of the W and GHZ distributions with classical realizations in the triangle scenario. It also identifies the distribution described by Fritz [Fri12], compatible with a quantum realization in the triangle scenario, as classically infeasible. For all these results, the relaxed SDP formulation is far less memory-demanding than the raw linear programming formulation.

Furthermore, the SDP approach is the only method that can be used when considering inflations or extensions, and thus causal compatibility, in the presence of terminal nodes which can take continuous values, even when the latent variables are classical. Therefore, not only can quantum inflation and scalar extension be leveraged to obtain results for networks with classical sources, but in fact it can be argued that considering their commuting-variables version is the most suitable method to proceed for addressing causal compatibility with classical realizations in large causal structures.

Moreover, considering commuting measurements of a same party also allows one to study mixed scenarios where there are both classical and quantum latent variables. As a general rule, commutation between all measurement operators of a same party should be imposed whenever the associated node is only affected by classical latent variables. It can also be the case, however, that a visible node is influenced by both classical and quantum nodes. An example of this could be the node A in the triangle scenario of Fig. 5.1c, whenever only the variable U_{AB} is classical, while U_{BC} and U_{AC} are quantum. In such a case, imposing the commutation of all measurement operators E_a (or all its inflations, $E_a^{i,j}$) constrains the share received from U_{AC} to be classical as well. One must instead impose commutation of only the operators that act on the Hilbert space associated to the classical system, leaving those operators defined on genuinely quantum subspaces of the global Hilbert space as noncommuting. In the example, this would mean imposing the constraints $[E_a^{i,j}, E_{a'}^{i',j}] = 0$ whenever $i \neq i'$, while in general the operators $E_a^{i,j}$ and $E_{a'}^{i,j'}$ would not commute as they act on the same copy i of the quantum state U_{AC}.

D.2 Nonlinear constraints in compatibility problems

In order to solve causal compatibility problems using quantum inflation, one imposes the moment identification constraints of Eq. (5.26) to associate certain expectation values with the distribution one wishes to test. In this appendix we take this identification a step further, noting that analogue constraints can also be imposed in scalar extension.

Eq. (5.26) only deals with the case in which the initial expectation value can be factorized in expectation values over disjoint copies of the original causal structure. In general one will also find expectation values that factorize in some expectation values over disjoint copies of the original scenario, along with other expectation values that, for some reason, can not be computed from P_{obs}, and therefore would remain as variables. For instance, in the example of inflations of the triangle scenario in Section 5.4, one of these expectation values could be

$$\begin{aligned}
&\left\langle E_1^{\{\rho^1_{\text{AB}},\rho^1_{\text{AC}}\}} E_0^{\{\rho^2_{\text{AB}},\rho^2_{\text{AC}}\}} F_0^{\{\rho^1_{\text{BC}},\rho^1_{\text{AB}}\}} F_1^{\{\rho^3_{\text{BC}},\rho^2_{\text{AB}}\}} G_1^{\{\rho^1_{\text{AC}},\rho^1_{\text{BC}}\}} G_0^{\{\rho^3_{\text{AC}},\rho^3_{\text{BC}}\}} \right\rangle_\rho \\
&= \left\langle E_1^{\{\rho^1_{\text{AB}},\rho^1_{\text{AC}}\}} F_0^{\{\rho^1_{\text{BC}},\rho^1_{\text{AB}}\}} G_1^{\{\rho^1_{\text{AC}},\rho^1_{\text{BC}}\}} \right\rangle_\rho \left\langle E_0^{\{\rho^2_{\text{AB}},\rho^2_{\text{AC}}\}} F_1^{\{\rho^3_{\text{BC}},\rho^2_{\text{AB}}\}} G_0^{\{\rho^3_{\text{AC}},\rho^3_{\text{BC}}\}} \right\rangle_\rho \\
&= P_{\text{obs}}(1,0,1) \left\langle E_0^{\{\rho^2_{\text{AB}},\rho^2_{\text{AC}}\}} F_1^{\{\rho^3_{\text{BC}},\rho^2_{\text{AB}}\}} G_0^{\{\rho^3_{\text{AC}},\rho^3_{\text{BC}}\}} \right\rangle_\rho \\
&\qquad\qquad \text{apply } \rho^1_{\text{AB}} \leftrightarrow \rho^2_{\text{AB}}, \rho^1_{\text{AC}} \leftrightarrow \rho^2_{\text{AC}}, \rho^1_{\text{BC}} \leftrightarrow \rho^3_{\text{BC}} \\
&= P_{\text{obs}}(1,0,1) \left\langle E_0^{\{\rho^1_{\text{AB}},\rho^1_{\text{AC}}\}} F_1^{\{\rho^1_{\text{BC}},\rho^1_{\text{AB}}\}} G_0^{\{\rho^3_{\text{AC}},\rho^1_{\text{BC}}\}} \right\rangle_\rho .
\end{aligned} \tag{D.2}$$

Given that, in compatibility problems, P_{obs} is given, constraints as that in Eq. (D.2) can be understood as linear constraints between two unknown variables—in the particular case of Eq. (D.2), relating the variables representing the original expectation value and that of $E_0^{\{\rho^1_{\text{AB}},\rho^1_{\text{AC}}\}} F_1^{\{\rho^1_{\text{BC}},\rho^1_{\text{AB}}\}} G_0^{\{\rho^3_{\text{AC}},\rho^1_{\text{BC}}\}}$—and therefore can be enforced in the semidefinite programs associated to the corresponding NPO feasibility problem.

This type of constraints allows one to perform tighter approximations to the sets of correlations admitting a given inflation. We explicitly demonstrate this by showing how the values of $v_{\max}$ in Tables 5.2 and 5.3 are modified when considering these new constraints. For both quantum and classical problems—in Tables D.1 and D.2, respectively—the upper bounds on threshold visibilities for the W and GHZ distributions are improved when inserting mixed-factorization constraints, while they seem not to make a difference for the Mermin-GHZ distribution. It is possible that this is a sign that the values found for the Mermin-GHZ distribution are tight, but it does not constitute a proof of it.

Distribution	$v_{\max}$	Inflation order	NPO level
GHZ	0.4141	2	$\mathfrak{L}_2^*$
W	0.7472	2	$\mathfrak{L}_2^*$
Mermin-GHZ	0.7685	2	$\mathcal{S}_2 \cup \mathfrak{L}_1$

Table D.1: Noise resistance of realizations in the quantum triangle scenario, using mixed-factorization constraints. For $v \geq v_{\max}$, the corresponding noisy distribution, in the sense of Eq. (5.18), does not have a realization in the triangle scenario when the latent variables are quantum. The set $\mathfrak{L}_2^*$ represents the subset of $\mathfrak{L}_2$ composed of operator strings of length not larger than 4.

Distribution	$v_{\max}$	Inflation order	NPO level
GHZ	0.4125	3	$\mathfrak{L}_1$
W	0.6939	2	$\mathfrak{L}_2$
Mermin-GHZ	0.5000	2	$\mathcal{S}_2 \cup \mathfrak{L}_1$

Table D.2: Noise resistance of realizations in the classical triangle scenario, using mixed-factorization constraints. For $v \geq v_{\max}$, the corresponding noisy distribution, in the sense of Eq. (5.18), cannot be realized in the triangle scenario when the latent variables are classical.

While it is clear that imposing these constraints is in some cases beneficial, it is not known whether they are required in the asymptotic limit. An intuition motivating an answer in the negative comes from the classical inflation technique [NW17], where convergence of the method to the set of correlations compatible with classical realizations does not need of the analogue of these constraints to be imposed. In that case, the analogue of Eq. (5.26) are sufficient to ensure that the method converges.

Bibliography

[AAB+19] Gadi Aleksandrowicz, Thomas Alexander, Panagiotis Barkoutsos, Luciano Bello, Yael Ben-Haim, David Bucher, ... Christa Zoufal. Qiskit: An open-source framework for quantum computing, (2019).

[Aar15] Scott Aaronson. Read the fine print. *Nat. Phys.* **11** (2015).

[ABG+07] Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani. Device-independent security of quantum cryptography against collective attacks. *Phys. Rev. Lett.* **98**, 230501 (2007).

[ABH+17] John-Mark A. Allen, Jonathan Barrett, Dominic C. Horsman, Ciarán M. Lee, and Robert W. Spekkens. Quantum common causes and quantum causal models. *Phys. Rev. X* **7**, 031021 (2017).

[AHJM11] Armen E. Allahverdyan, Karen V. Hovhannisyan, Dominik Janzing, and Günter Mahler. Thermodynamic limits of dynamic cooling. *Phys. Rev. E* **84**, 041109 (2011).

[AHM10] Armen E. Allahverdyan, Karen V. Hovhannisyan, and Günter Mahler. Optimal refrigerator. *Phys. Rev. E* **81**, 051129 (2010).

[AHMG13] Armen E. Allahverdyan, Karen V. Hovhannisyan, Alexey V. Melkikh, and Sasun G. Gevorkian. Carnot cycle at finite power: Attainability of maximal efficiency. *Phys. Rev. Lett.* **111**, 050601 (2013).

[AI07] Gerardo Adesso and Fabrizio Illuminati. Entanglement in continuous-variable systems: Recent advances and current perspectives. *J. Phys. A: Math. Theor.* **40** (28), 7821 (2007).

[AJM08] Armen E. Allahverdyan, Ramandeep S. Johal, and Günter Mahler. Work extremum principle: Structure and function of quantum heat engines. *Phys. Rev. E* **77**, 041118 (2008).

[AKMM14] Álvaro M. Alhambra, Achim Kempf, and Eduardo Martín-Martínez. Casimir forces on atoms in optical cavities. *Phys. Rev. A* **89**, 033835 (2014).

[Ali79] Robert Alicki. The quantum open system as a model of the heat engine. *J. Phys. A: Math. Gen.* **12** (5), L103–L107 (1979).

[Amb10] Andris Ambainis. Variable time amplitude amplification and a faster quantum algorithm for solving systems of linear equations, *arXiv:1010.4458*.

[AMPS13] Ahmed Almheiri, Donald Marolf, Joseph Polchinski, and James Sully. Black holes: complementarity or firewalls? *JHEP* **2013** (2), 62 (2013).

[ASB16] Martin Arjovsky, Amar Shah, and Yoshua Bengio. Unitary evolution recurrent neural networks. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48*, ICML'16, pages 1120–1128, (2016).

[ASWdF16] Yannis M. Assael, Brendan Shillingford, Shimon Whiteson, and Nando de Freitas. LipNet: End-to-end sentence-level lipreading, *arXiv:1611.01599*.

[Avi04] David Avis. A revised implementation of the reverse search vertex enumeration algorithm. In Gil Kalai and Günter M. Ziegler, editors, *Polytopes — Combinatorics and Computation, DMV Seminar, vol. 29*, pages 177–198. Birkhäuser, Basel, (2004).

[AW08] Janet Anders and Andreas J. Winter. Entanglement and separability of quantum harmonic oscillator systems at finite temperature. *Quantum Inf. Comput.* **8** (3–4), 245–262 (2008).

[BB84] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing*, volume 175, page 8, (1984).

[BC12] Dominic W. Berry and Andrew M. Childs. Black-box Hamiltonian simulation and unitary implementation. *Quantum Inf. Comput.* **12** (1–2), 29–62 (2012).

[BCK15] Dominic W. Berry, Andrew M. Childs, and Robin Kothari. Hamiltonian simulation with nearly optimal dependence on all parameters. In *Proceedings of the 56th Annual Symposium on Foundations of Computer Science*, pages 792–809, (2015).

[BCKW15] Charles Blundell, Julien Cornebise, Koray Kavukcuoglu, and Daan Wierstra. Weight uncertainty in neural networks. In *Proceedings of the 32nd International Conference on International Conference on Machine Learning*, volume 37, pages 1613–1622, (2015).

[BCMdW10] Harry Buhrman, Richard Cleve, Serge Massar, and Ronald de Wolf. Nonlocality and communication complexity. *Rev. Mod. Phys.* **82**, 665–698 (2010).

[BCP+14] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.* **86**, 419–478 (2014).

[BCSSF19] Ileana-Cristina Benea-Chelmus, Francesca Fabiana Settembrini, Giacomo Scalari, and Jérôme Faist. Electric field correlation measurements on the electromagnetic vacuum state. *Nature* **568**, 202–206 (2019).

[BCWA17] Flavio Baccari, Daniel Cavalcanti, Peter Wittek, and Antonio Acín. Efficient device-independent entanglement detection for multipartite systems. *Phys. Rev. X* **7**, 021042 (2017).

[BD84] Nicholas D. Birrell and Paul C. W. Davies. *Quantum Fields in Curved Space*. Cambridge University Press, (1984).

[BDSW96] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A* **54**, 3824–3851 (1996).

[Bell64] John S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics* **1** (3), 195–200 (1964).

[Ber14] Dominic W. Berry. High-order quantum algorithm for solving linear differential equations. *J. Phys. A: Math. Theor.* **47** (10), 105301 (2014).

[BFH16] Eric G. Brown, Nicolai Friis, and Marcus Huber. Passivity and practical work extraction using gaussian operations. *New J. Phys.* **18** (11), 113028 (2016).

[BGMBMM15a] Ana Blasco, Luis J. Garay, Mercedes Martín-Benito, and Eduardo Martín-Martínez. The quantum echo of the early universe. *Can. J. Phys.* **93** (9), 968–970 (2015).

[BGMBMM15b] Ana Blasco, Luis J. Garay, Mercedes Martín-Benito, and Eduardo Martín-Martínez. Violation of the strong Huygens' principle and timelike signals from the early universe. *Phys. Rev. Lett.* **114**, 141103 (2015).

[BGMBMM16] Ana Blasco, Luis J. Garay, Mercedes Martín-Benito, and Eduardo Martín-Martínez. Timelike information broadcasting in cosmology. *Phys. Rev. D* **93**, 024055 (2016).

[BGP10] Cyril Branciard, Nicolas Gisin, and Stefano Pironio. Characterizing the nonlocal correlations created via entanglement swapping. *Phys. Rev. Lett.* **104**, 170401 (2010).

[BHK05] Jonathan Barrett, Lucien Hardy, and Adrian Kent. No Signaling and Quantum Key Distribution. *Phys. Rev. Lett.* **95**, 010503 (2005).

[BJ03] Brian H. Bransden and Charles J. Joachain. *Physics of Atoms and Molecules*. Prentice Hall, (2003).

[BLM$^+$05] Jonathan Barrett, Noah Linden, Serge Massar, Stefano Pironio, Sandu Popescu, and David Roberts. Nonlocal correlations as an information-theoretic resource. *Phys. Rev. A* **71**, 022101 (2005).

[BMG17] John Bradshaw, Alexander G. de G. Matthews, and Zoubin Ghahramani. Adversarial examples, uncertainty, and transfer testing robustness in Gaussian process hybrid deep networks, *arXiv:1707.02476*.

[BMMMM13] Eric G. Brown, Eduardo Martín-Martínez, Nicolas C. Menicucci, and Robert B. Mann. Detectors for probing relativistic quantum physics beyond perturbation theory. *Phys. Rev. D* **87**, 084062 (2013).

[Bor26] Max Born. Zur Quantenmechanik der Stoßvorgänge. *Z. Physik* **37**, 863–867 (1926).

[BP02] Heinz-Peter Breuer and Francesco Petruccione. *The theory of open quantum systems.* Oxford University Press, (2002).

[BPŻ13] Samuel L. Braunstein, Stefano Pirandola, and Karol Życzkowski. Better late than never: information retrieval from black holes. *Phys. Rev. Lett.* **110**, 101301 (2013).

[Bra02] Daniel Braun. Creation of entanglement by interaction with a common heat bath. *Phys. Rev. Lett.* **89**, 277901 (2002).

[Bra05] Daniel Braun. Entanglement from thermal blackbody radiation. *Phys. Rev. A* **72**, 062324 (2005).

[BRGP12] Cyril Branciard, Denis Rosset, Nicolas Gisin, and Stefano Pironio. Bilocal versus nonbilocal correlations in entanglement-swapping experiments. *Phys. Rev. A* **85**, 032119 (2012).

[Bro13] Eric G. Brown. Thermal amplification of field-correlation harvesting. *Phys. Rev. A* **88**, 062336 (2013).

[BRS12] Stephen D. Bartlett, Terry Rudolph, and Robert W. Spekkens. Reconstruction of Gaussian quantum mechanics from Liouville mechanics with an epistemic restriction. *Phys. Rev. A* **86**, 012103 (2012).

[BSS14] Jean-Daniel Bancal, Lana Sheridan, and Valerio Scarani. More randomness from the same data. *New J. Phys.* **16** (3), 033011 (2014).

[BW92] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.* **69**, 2881–2884 (1992).

[BWP+17] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature* **549** (7671), 195–202 (2017).

[CA75] Frank L. Curzon and Boye K. Ahlborn. Efficiency of a Carnot engine at maximum power output. *Am. J. Phys.* **43** (1), 22–24 (1975).

[Cal85] Herbert B. Callen. *Thermodynamics and an Introduction to Thermostatistics*. John Wiley, 2nd edition, (1985).

[Cas48] Hendrik B. G. Casimir. On the attraction between two perfectly conducting plates. *Proc. Akad. Wet. Amsterdam* **51**, 793–795 (1948).

[CASA11] Daniel Cavalcanti, Mafalda L. Almeida, Valerio Scarani, and Antonio Acín. Quantum networks reveal quantum nonlocality. *Nat. Commun.* **2**, 184 (2011).

[CCA17] Rafael Chaves, Daniel Cavalcanti, and Leandro Aolita. Causal hierarchy of multipartite bell nonlocality. *Quantum* **1**, 23 (2017).

[CCA+18] Rafael Chaves, Gonzalo Carvacho, Iris Agresti, Valerio Di Giulio, Leandro Aolita, Sandro Giacomini and Fabio Sciarrino. Quantum violation of an instrumental test. *Nat. Phys.* **14** (3), 291 (2018).

[CCD+03] Andrew M. Childs, Richard Cleve, Enrico Deotto, Edward Farhi, Sam Gutmann, and Daniel A. Spielman. Exponential algorithmic speedup by a quantum walk. In *Proceedings of the 35th Annual ACM Symposium on Theory of computing*, pages 59–68, (2003).

[CDFK12] Yudong Cao, Anmer Daskin, Steven Frankel, and Sabre Kais. Quantum circuit design for solving linear systems of equations. *Mol. Phys.* **110** (15–16), 1675—-1680 (2012).

[CF12] Rafael Chaves and Tobias Fritz. Entropic approach to local realism and noncontextuality. *Phys. Rev. A* **85**, 032113 (2012).

[Cha16] Rafael Chaves. Polynomial Bell inequalities. *Phys. Rev. Lett.* **116**, 010402 (2016).

[Chi10] Andrew M. Childs. On the relationship between continuous- and discrete-time quantum walk. *Commun. Math. Phys.* **294** (2), 581–603 (2010).

[Cir80] Boris S. Cirel'son. Quantum generalizations of Bell's inequality. *Lett. Math. Phys.* **4** (2), 93–100 (1980).

[CKBG15] Rafael Chaves, Richard Kueng, Jonatan Bohr Brask, and David Gross. Unifying framework for relaxations of the causal assumptions in Bell's theorem. *Phys. Rev. Lett.* **114**, 140403 (2015).

[CKS17] Andrew M. Childs, Robin Kothari, and Rolando D. Somma. Quantum algorithm for systems of linear equations with exponentially improved dependence on precision. *SIAM J. Comput.* **46** (6), 1920–1950 (2017).

[CKW00] Valerie Coffman, Joydip Kundu, and William K. Wootters. Distributed entanglement. *Phys. Rev. A* **61**, 052306 (2000).

[CLM+14] Rafael Chaves, Lukas Luft, Thiago O. Maciel, David Gross, Dominik Janzing, and Bernhard Schölkopf. Inferring latent structures via information inequalities. In *Proceedings of the 30th Conference on Uncertainty in Artificial Intelligence*, pages 112–121, (2014).

[CMG15] Rafael Chaves, Christian Majenz, and David Gross. Information–theoretic implications of quantum causal structures. *Nat. Commun.* **6**, 5766 (2015).

[Col75] George E. Collins. Quantifier elimination for the elementary theory of real closed fields by cylindrical algebraic decomposition. *Lect. Notes Comput. Sci.* **33**, 134–183 (1975).

[Col07] Roger Colbeck. Ph.D. Thesis. *University of Cambridge* (2007).

[CPP+13] Yudong Cao, Anargyros Papageorgiou, Iasonas Petras, Joseph Traub, and Sabre Kais. Quantum algorithm and circuit design solving the Poisson equation. *New J. Phys.* **15** (1), 013021 (2013).

[CRGB18] Esteban Castro-Ruiz, Flaminia Giacomini, and Časlav Brukner. Dynamics of quantum causal structures. *Phys. Rev. X* **8**, 011047 (2018).

[CS09] Youngmin Cho and Lawrence K. Saul. Kernel methods for deep learning. In *Proceedings of the Advances in Neural Information Processing Systems conference*, pages 342–350, (2009).

[CS16] Fabio Costa and Sally Shrapnel. Quantum causal modelling. *New J. Phys.* **18** (6), 063032 (2016).

[CTDRG87] Claude Cohen-Tannoudji, Jacques Dupont-Roc, and Gilbert Grynberg. *Photons and Atoms: Introduction to Quantum Electrodynamics*. Wiley-VCH, (1987).

[CV95] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning* **20** (3), 273–297 (1995).

[Deu91] David Deutsch. Quantum mechanics near closed timelike lines. *Phys. Rev. D* **44**, 3197–3217 (1991).

[DFS16] Amit Daniely, Roy Frostig, and Yoram Singer. Toward deeper understanding of neural networks: The power of initialization and a dual view on expressivity. In *Proceedings of the Advances in Neural Information Processing Systems 29th conference*, pages 2253–2261, (2016).

[DHI79] Bryce S. DeWitt, Stephen W. Hawking, and Werner Israel. *General Relativity: An Einstein Centenary Survey*. Cambridge University Press, (1979).

[Eke91] Artur K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.* **67**, 661–663 (1991).

[EOG15] Massimiliano Esposito, Maicol A. Ochoa, and Michael Galperin. Quantum thermodynamics: A nonequilibrium Green's function approach. *Phys. Rev. Lett.* **114**, 080602 (2015).

[Eva12] Robin J. Evans. Graphical methods for inequality constraints in marginalized DAGs. In *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing*. IEEE, (2012).

[Eva18] Robin J. Evans. Margins of discrete Bayesian networks. *Ann. Stat.* **46** (6A), 2623–2656 (2018).

[FFH14] Michael Frey, Ken Funo, and Masahiro Hotta. Strong local passivity in finite quantum systems. *Phys. Rev. E* **90**, 012127 (2014).

[FGG14] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A Quantum Approximate Optimization Algorithm, *arXiv:1411.4028*.

[FGN06] Reinhard Furrer, Marc G. Genton, and Douglas Nychka. Covariance tapering for interpolation of large spatial datasets. *J. Comput. Graph. Stat.* **15** (3), 502–523 (2006).

[FN18] Edward Farhi and Hartmut Neven. Classification with quantum neural networks on near term processors, *arXiv:1802.06002*.

[FP14] Nahuel Freitas and Juan Pablo Paz. Analytic solution for heat flow through a general harmonic network. *Phys. Rev. E* **90**, 042128 (2014).

[Fri12] Tobias Fritz. Beyond Bell's theorem: correlation scenarios. *New J. Phys.* **14** (10), 103001 (2012).

[FSA+13] Tobias Fritz, Ana Belén Sainz, Remigiusz Augusiak, Jonatan Bohr Brask, Rafael Chaves, Antony Leverrier and Antonio Acín. Local orthogonality as a multipartite principle for quantum correlations. *Nat. Commun.* **4**, 2263 (2013).

[GBB11] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. Deep sparse rectifier neural networks. In *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics*, volume 15, pages 315–323, (2011).

[GC01] Daniel Gottesman and Isaac Chuang. Quantum digital signatures, *arXiv:quant-ph/0105032*.

[GG16] Yarin Gal and Zoubin Ghahramani. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48, pages 1050–1059, (2016).

[Gha10] Sevag Gharibian. Strong NP-hardness of the quantum separability problem. *Quantum Inf. Comput.* **10** (3–4), 343–360 (2010).

[Gis91] Nicolas Gisin. Bell's inequality holds for all non-product states. *Phys. Lett. A* **154** (5-6), 201–202 (1991).

[Gis17] Nicolas Gisin. The Elegant Joint Quantum Measurement and some conjectures about n-locality in the triangle and other configurations, *arXiv:1708.05556*.

[GK92] Eitan Geva and Ronnie Kosloff. A quantum-mechanical heat engine operating in finite time. A model consisting of spin-1/2 systems as the working fluid. *J. Chem. Phys.* **96** (4), 3054–3067 (1992).

[GKAG15] David Gelbwaser-Klimovsky and Alán Aspuru-Guzik. Strongly coupled quantum heat machines. *J. Phys. Chem. Lett.* **6** (17), 3477–3482 (2015).

[GKW+18] Koon Tong Goh, Jędrzej Kaniewski, Elie Wolfe, Tamás Vértesi, Xingyao Wu, Yu Cai, ... Valerio Scarani. Geometry of the set of quantum correlations. *Phys. Rev. A* **97**, 022104 (2018).

[GM98] Dan Geiger and Christopher Meek. Graphical models and exponential families. In *Proceedings of the 14th conference on Uncertainty in Artificial Intelligence*, pages 156–165, (1998).

[GM99] Dan Geiger and Christopher Meek. Quantifier elimination for statistical problems. In *Proceedings of the 15th conference on Uncertainty in Artificial Intelligence*, pages 226–235, (1999).

[Göd31] Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. *Monatsh. Math* **38** (1), 173–198 (1931).

[GP92] Alberto Galindo and Pedro Pascual. *Quantum Mechanics I.* Springer-Verlag, (1992).

[GPC+16] Varun Gulshan, Lily Peng, Marc Coram, Martin C. Stumpe, Derek Wu, Arunachalam Narayanaswamy, ... Dale R. Webster. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **316** (22), 2402–2410 (2016).

[GPSB17] Kathrin Grosse, David Pfaff, Michael Thomas Smith, and Michael Backes. How wrong am I? - studying adversarial examples and their impact on uncertainty in Gaussian process machine learning models, *arXiv:1711.06598*.

[GRE14] Rodrigo Gallego, Arnau Riera, and Jens Eisert. Thermal machines beyond the weak coupling regime. *New J. Phys.* **16** (12), 125009 (2014).

[GVW+15] Marissa Giustina, Marijn A. M. Versteegh, Sören Wengerowsky, Johannes Handsteiner, Armin Hochrainer, Kevin Phelan, ... Anton Zeilinger. Significant-loophole-free test of Bell's theorem with entangled photons. *Phys. Rev. Lett.* **115**, 250401 (2015).

[H⊗396] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. *Phys. Lett. A* **223** (1–2), 1–8 (1996).

[H⊗409] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.* **81**, 865–942 (2009).

[Haw75] Stephen W. Hawking. Particle creation by black holes. *Commun. Math. Phys.* **43** (3), 199–220 (1975).

[Haw76] Stephen W. Hawking. Breakdown of predictability in gravitational collapse. *Phys. Rev. D* **14**, 2460–2473 (1976).

[HBD+15] Bas Hensen, Hannes Bernien, Anaïs E. Dréau, Andreas Reiserer, Norbert Kalb, Machiel S. Blok, ... Ronald Hanson. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526**, 682–686 (2015).

[Hei27] Werner Heisenberg. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Z. Physik* **43**, 172–198 (1927).

[HHL09] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.* **103**, 150502 (2009).

[Hin07] Geoffrey E. Hinton. Learning multiple layers of representation. *Trends Cogn. Sci.* **11**, 428–434 (2007).

[HLP14] Joe Henson, Raymond Lal, and Matthew F. Pusey. Theory-independent limits on correlations from generalized Bayesian networks. *New J. Phys.* **16** (11), 113043 (2014).

[Hot08] Masahiro Hotta. Quantum measurement information as a key to energy extraction from local vacuums. *Phys. Rev. D* **78**, 045006 (2008).

[Hot09] Masahiro Hotta. Quantum Energy Teleportation in spin chain systems. *J. Phys. Soc. Jpn.* **78** (3), 034001 (2009).

[HPS16] Stephen W. Hawking, Malcolm J. Perry, and Andrew Strominger. Soft hair on black holes. *Phys. Rev. Lett.* **116**, 231301 (2016).

[HR17] Stephanie Hyland and Gunnar Rätsch. Learning unitary operators with help from u(n). In *Proceedings of the AAAI Conference on Artificial Intelligence*, (2017).

[HW01] Alexander S. Holevo and Reinhard F. Werner. Evaluating capacities of bosonic Gaussian channels. *Phys. Rev. A* **63**, 032312 (2001).

[IMB+15] Francesco Intravaia, Vanik E. Mkrtchian, Stefan Y. Buhmann, Stefan Scheel, Diego A. R. Dalvit, and Carsten Henkel. Friction forces on atoms after acceleration. *J. Phys. Condens. Matter* **27** (21), 214020 (2015).

[IO14] Yuki Izumida and Koji Okuda. Work output and efficiency at maximum power of linear irreversible heat engines operating with a finite-sized heat source. *Phys. Rev. Lett.* **112**, 180603 (2014).

[Jac02] John D. Jackson. From Lorenz to Coulomb and other explicit gauge transformations. *Am. J. Phys.* **70** (9), 917 (2002).

[JAL14] Benito A. Juárez-Aubry and Jorma Louko. Onset and decay of the 1 + 1 Hawking–Unruh effect: what the derivative-coupling detector saw. *Class. Quantum Gravity* **31** (24), 245007 (2014).

[JCD+19] Max Jaderberg, Wojciech M. Czarnecki, Iain Dunning, Luke Marris, Guy Lever, Antonio Garcia Castañeda, ... Thore Graepel. Human-level performance in 3D multiplayer games with population-based reinforcement learning. *Science* **364** (6443), 859–865 (2019).

[JMMK15] Robert H. Jonsson, Eduardo Martín-Martínez, and Achim Kempf. Information transmission without energy exchange. *Phys. Rev. Lett.* **114**, 110505 (2015).

[Kit95] Alexei Y. Kitaev. Quantum measurements and the Abelian stabilizer problem, *arXiv:quant-ph/9511026*.

[KLL+17] Shelby Kimmel, Cedric Yen-Yu Lin, Guang Hao Low, Maris Ozols, and Theodore J. Yoder. Hamiltonian simulation with optimal sample complexity. *npj Quantum Inf.* **3** (1), 13 (2017).

[KR17] Ronnie Kosloff and Yair Rezek. The quantum harmonic Otto cycle. *Entropy* **19** (4), 136 (2017).

[Kra10] Alexander E. Kramida. A critical compilation of experimental data on spectral lines and energy levels of hydrogen, deuterium, and tritium. *At. Data. Nucl. Data Tables* **96** (6), 586–644 (2010).

[KVD+18] Amir Khoshaman, Walter Vinci, Brandon Denis, Evgeny Andriyash, and Mohammad H. Amin. Quantum Variational Autoencoder, *arXiv:1802.05779*.

[KvPÅ+17] Aditya Kela, Kai von Prillwitz, Johan Åberg, Rafael Chaves, and David Gross. Semidefinite tests for latent causal structures, *arXiv:1701.00652*.

[LBN+18] Jaehoon Lee, Yasaman Bahri, Roman Novak, Samuel S. Schoenholz, Jeffrey Pennington, and Jascha Sohl-Dickstein. Deep neural networks as Gaussian processes. In *Proceedings of the International Conference on Learning Representations*, (2018).

[LG14] François Le Gall. Powers of tensors and fast matrix multiplication. In *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, pages 296–303, (2014).

[LL80] Lev D. Landau and Evgeny M. Lifshitz. *Statistical Physics, Part I.* Pergamon, New York, (1980).

[Llo96] Seth Lloyd. Universal quantum simulators. *Science* **273** (5278), 1073–1078 (1996).

[LR47] Willis E. Lamb and Robert C. Retherford. Fine structure of the hydrogen atom by a microwave method. *Phys. Rev.* **72**, 241–243 (1947).

[LRB+15] Yeong-Cherng Liang, Denis Rosset, Jean-Daniel Bancal, Gilles Pütz, Tomer Jack Barnea, and Nicolas Gisin. Family of Bell-like inequalities as device-independent witnesses for entanglement depth. *Phys. Rev. Lett.* **114**, 190401 (2015).

[LRW+19] Ding Liu, Shi-Ju Ran, Peter Wittek, Cheng Peng, Raul Blázquez García, Gang Su and Maciej Lewenstein. Machine learning by unitary tensor network of hierarchical tree structure. *New J. Phys.* **21** (7), 073059 (2019).

[LS09] Juan León and Carlos Sabín. Generation of atom-atom correlations inside and outside the mutual light cone. *Phys. Rev. A* **79**, 012304 (2009).

[LS17] Ciarán M. Lee and Robert W. Spekkens. Causal inference via algebraic geometry: feasibility tests for functional causal structures with two binary observed variables. *J. Causal Inference* **5** (2) (2017).

[LSS87] Willis E. Lamb, Rainer R. Schlicher, and Marlan O. Scully. Matter-field interaction in atomic physics and quantum optics. *Phys. Rev. A* **36**, 2763–2772 (1987).

[MBL+13] Tobias Moroder, Jean-Daniel Bancal, Yeong-Cherng Liang, Martin Hofmann, and Otfried Gühne. Device-independent entanglement quantification and related applications. *Phys. Rev. Lett.* **111**, 030501 (2013).

[MCI+10] J. Gonzalo Muga, Xi Chen, Sara Ibáñez, Ion Lizuain, and Andreas Ruschhaupt. Transitionless quantum drivings for the harmonic oscillator. *J. Phys. B: At. Mol. Opt. Phys.* **43** (8), 085509 (2010).

[Mer90] N. David Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.* **65**, 1838–1840 (1990).

[MHR+18] Alexander G. de G. Matthews, Jiri Hron, Mark Rowland, Richard E. Turner, and Zoubin Ghahramani. Gaussian process behaviour in wide deep neural networks. In *Proceedings of the International Conference on Learning Representations*, (2018).

[MM15] Eduardo Martín-Martínez. Causality issues of particle detector models in QFT and quantum optics. *Phys. Rev. D* **92**, 104019 (2015).

[MMBDK13] Eduardo Martín-Martínez, Eric G. Brown, William Donnelly, and Achim Kempf. Sustainable entanglement production from a quantum field. *Phys. Rev. A* **88**, 052310 (2013).

[MML14] Eduardo Martín-Martínez and Jorma Louko. Particle detectors and the zero mode of a quantum field. *Phys. Rev. D* **90**, 024015 (2014).

[MMM12] Eduardo Martín-Martínez and Nicolas C. Menicucci. Cosmological quantum entanglement. *Class. Quantum Gravity* **29** (22), 224003 (2012).

[MMMdR13] Eduardo Martín-Martínez, Miguel Montero, and Marco del Rey. Wavepacket detection with the Unruh-DeWitt model. *Phys. Rev. D* **87**, 064038 (2013).

[MMS16] Eduardo Martín-Martínez and Barry C. Sanders. Precise space–time positioning for entanglement harvesting. *New J. Phys.* **18** (4), 043031 (2016).

[MMST16] Eduardo Martín-Martínez, Alexander R. H. Smith, and Daniel R. Terno. Spacetime structure and vacuum entanglement. *Phys. Rev. D* **93**, 044001 (2016).

[MOS] MOSEK ApS. The MOSEK optimization toolbox for Python manual. Version 8.1, (2017).

[MP87] Michael A. Morrison and Gregory A. Parker. A guide to rotations in quantum mechanics. *Austral. J. Phys.* **40** (4), 465 (1987).

[MP03] Serge Massar and Stefano Pironio. Violation of local realism versus detection efficiency. *Phys. Rev. A* **68**, 062109 (2003).

[MP13] Esteban A. Martínez and Juan Pablo Paz. Dynamics and thermodynamics of linear quantum open systems. *Phys. Rev. Lett.* **110**, 130406 (2013).

[MRBAG16] Jarrod R. McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.* **18** (2), 023023 (2016).

[MV96] Valter Moretti and Luciano Vanzo. Thermal Wightman functions and renormalized stress tensors in the Rindler wedge. *Phys. Lett. B* **375** (1), 54–59 (1996).

[MY04] Dominic Mayers and Andrew Yao. Self testing quantum apparatus. *Quantum Inf. Comput.* **4** (4), 273–286 (2004).

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information.* Cambridge University Press, (2010).

[Nea94] Radford M. Neal. Priors for infinite networks. Technical Report crg-tr-94-1, University of Toronto, (1994).

[NGHA15] Miguel Navascués, Yelena Guryanova, Matty J. Hoban, and Antonio Acín. Almost quantum correlations. *Nat. Commun.* **6**, 6288 (2015).

[NPA07] Miguel Navascués, Stefano Pironio, and Antonio Acín. Bounding the set of quantum correlations. *Phys. Rev. Lett.* **98** (1), 010401 (2007).

[NPA08] Miguel Navascués, Stefano Pironio, and Antonio Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New J. Phys.* **10** (7), 073013 (2008).

[NRSS09] Bruno Nachtergaele, Hillel Raz, Benjamin Schlein, and Robert Sims. Lieb-Robinson bounds for harmonic and anharmonic lattice systems. *Commun. Math. Phys.* **286** (3), 1073–1098 (2009).

[NS10] Bruno Nachtergaele and Robert Sims. Lieb-Robinson bounds in quantum many-body physics. *Contemp. Math.* **529**, 141 (2010).

[NW09] Miguel Navascués and Harald Wunderlich. A glance beyond the quantum model. *Proc. R. Soc. A* **466** (2009).

[NW17] Miguel Navascués and Elie Wolfe. The inflation technique solves completely the classical inference problem, *arXiv:1707.06476.*

[OOL+] Frank. W. J. Olver, Adri B. Olde Daalhuis, Daniel W. Lozier, Barry I. Schneider, Ronalf F. Boisvert, Charles W. Clark, ... Bonita V. Saunders. NIST digital library of mathematical functions. http://dlmf.nist.gov/.

[OR11] S. Jay Olson and Timothy C. Ralph. Entanglement between the future and the past in the quantum vacuum. *Phys. Rev. Lett.* **106**, 110404 (2011).

[OR12] S. Jay Olson and Timothy C. Ralph. Extraction of timelike entanglement from the quantum vacuum. *Phys. Rev. A* **85**, 012306 (2012).

[PAA+14] Francesco Plastina, Antonio Alecce, Tony J. G. Apollaro, Giovanni Falcone, Gianluca Francica, Fernando Galve, ... Roberta Zambrini. Irreversible work and inner friction in quantum thermodynamic processes. *Phys. Rev. Lett.* **113**, 260601 (2014).

[PAAM17] Valerio Pruneri, Carlos Abellán, Waldimar Amaya, and Morgan W. Mitchell. Process for quantum random number generation in a multimode laser cavity. Patent no. WO2017071901A1 (2017).

[PAM+10] Stefano Pironio, Antonio Acín, Serge Massar, Antoine Boyer de la Giroday, Dimitri N. Matsukevich, Peter Maunz, ... Christopher Monroe. Random numbers certified by Bell's theorem. *Nature* **464**, 1021 (2010).

[PB15] Jacques Pienaar and Časlav Brukner. A graph-separation theorem for quantum causal models. *New J. Phys.* **17** (7), 073020 (2015).

[PBS11] Stefano Pironio, Jean-Daniel Bancal, and Valerio Scarani. Extremal correlations of the tripartite no-signaling polytope. *J. Phys. A: Math. Theor.* **44** (6), 065303 (2011).

[Pea09] Judea Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, (2009).

[Per96] Asher Peres. Separability criterion for density matrices. *Phys. Rev. Lett.* **77**, 1413–1415 (1996).

[Pie17] Jacques Pienaar. Which causal structures might support a quantum–classical gap? *New J. Phys.* **19** (4), 043021 (2017).

[PK19] Alejandro Pozas-Kerstjens. Computational appendix of *Bounding the sets of classical and quantum correlations in networks*. Zenodo 2646262, (2019).

[PKBH18] Alejandro Pozas-Kerstjens, Eric G. Brown, and Karen V. Hovhannisyan. A quantum Otto engine with finite heat baths: energy, correlations, and degradation. *New J. Phys.* **20** (4), 043034 (2018).

[PKHB17] Alejandro Pozas-Kerstjens, Karen V. Hovhannisyan, and Eric G. Brown. Computational appendix of *A quantum Otto engine with finite heat baths: energy, correlations, and degradation*. Zenodo 847182, (2017).

[PKLMM17] Alejandro Pozas-Kerstjens, Jorma Louko, and Eduardo Martín-Martínez. Degenerate detectors are unable to harvest spacelike entanglement. *Phys. Rev. D* **95**, 105009 (2017).

[PKMM15] Alejandro Pozas-Kerstjens and Eduardo Martín-Martínez. Harvesting correlations from the quantum vacuum. *Phys. Rev. D* **92**, 064042 (2015).

[PKMM16] Alejandro Pozas-Kerstjens and Eduardo Martín-Martínez. Entanglement harvesting from the electromagnetic vacuum with hydrogenlike atoms. *Phys. Rev. D* **94**, 064074 (2016).

[PKRR+19] Alejandro Pozas-Kerstjens, Rafael Rabelo, Łukasz Rudnicki, Rafael Chaves, Daniel Cavalcanti, Miguel Navascués and Antonio Acín. Bounding the sets of classical and quantum correlations in networks, *arXiv:1904.08943*. Accepted in Physical Review Letters.

[PKW18] Alejandro Pozas-Kerstjens and Peter Wittek. Code for *Bayesian Deep Learning on a Quantum Computer*. GitLab repository: bayesian-dl-quantum, (2018).

[PLBH+17] Martí Perarnau-Llobet, Elisa Bäumer, Karen V. Hovhannisyan, Marcus Huber, and Antonio Acín. No-go theorem for the characterization of work fluctuations in coherent quantum systems. *Phys. Rev. Lett.* **118**, 070601 (2017).

[PLWR+18] Martí Perarnau-Llobet, Henrik Wilming, Arnau Riera, Rodrigo Gallego, and Jens Eisert. Strong coupling corrections in quantum thermodynamics. *Phys. Rev. Lett.* **120**, 120602 (2018).

[PMA+15] Lorenzo M. Procopio, Amir Moqanaki, Mateus Araújo, Fabio Costa, Irati Alonso Calafell, Emma G. Dowd, ... Philip Walther. Experimental superposition of orders of quantum gates. *Nat. Commun.* **6**, 7913 (2015).

[PMS+14] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, ... Jeremy L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).

[PNA10] Stefano Pironio, Miguel Navascués, and Antonio Acín. Convergent relaxations of polynomial optimization problems with non-commuting variables. *SIAM J. Optim.* **20**, 2157–2180 (2010).

[PR94] Sandu Popescu and Daniel Rohrlich. Quantum nonlocality as an axiom. *Found. Phys.* **24**, 379 (1994).

[Pre18] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).

[PSG16] Jukka P. Pekola, Samu Suomela, and Yuri M. Galperin. Finite-size bath in qubit thermodynamics. *J. Low Temp. Phys.* **184** (5-6), 1015–1029 (2016).

[PV91] Judea Pearl and Thomas Verma. A theory of inferred causation. In *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 441–452, (1991).

[PV09] Károly F. Pál and Tamás Vértesi. Quantum bounds on Bell inequalities. *Phys. Rev. A* **79**, 022120 (2009).

[QLSN07] Hai-Tao Quan, Yu-Xi Liu, Chang-Pu Sun, and Franco Nori. Quantum thermodynamic cycles and quantum heat engines. *Phys. Rev. E* **76**, 031105 (2007).

[RAM18] Jonathan G. Richens, Álvaro M. Alhambra, and Lluis Masanes. Finite-bath corrections to the second law of thermodynamics. *Phys. Rev. E* **97**, 062132 (2018).

[RAV+15] Katja Ried, Megan Agnew, Lydia Vermeyden, Dominik Janzing, Robert W. Spekkens, and Kevin J. Resch. A quantum advantage for inferring causal structure. *Nat. Phys.* **11** (5), 414 (2015).

[RDT+16] Johannes Roßnagel, Samuel T. Dawkins, Karl N. Tolazzi, Obinna Abah, Eric Lutz, Ferdinand Schmidt-Kaler and Kilian Singer. A single-atom heat engine. *Science* **352** (6283), 325–329 (2016).

[Rei29] Henri B. Reitlinger. *Sur l'utilisation de la chaleur dans les machines á feu.* Vaillant-Carmanne, Liége, Belgium, (1929).

[Rez00] Benni Reznik. Distillation of vacuum entanglement to EPR pairs, *arXiv:quant-ph/0008006.*

[Rez03] Benni Reznik. Entanglement from the Vacuum. *Found. Phys.* **33** (1), 167–176 (2003).

[RGW18] Denis Rosset, Nicolas Gisin, and Elie Wolfe. Universal bound on the cardinality of local hidden variables in networks. *Quantum Inf. Comput.* **18** (11–12), 910–926 (2018).

[Rib15] Grégoire Ribordy. Apparatus and method for quantum key distribution with enhanced security and reduced trust requirements. Patent no. EP3185463 (2015).

[RK06] Yair Rezek and Ronnie Kosloff. Irreversible performance of a quantum harmonic heat engine. *New J. Phys.* **8** (5), 83 (2006).

[RML14] Patrick Rebentrost, Masoud Mohseni, and Seth Lloyd. Quantum support vector machine for big data classification. *Phys. Rev. Lett.* **113**, 130503 (2014).

[Ros18] Denis Rosset. SymDPoly: symmetry-adapted moment relaxations for noncommutative polynomial optimization, *arXiv:1808.09598.*

[RR13] Thomas S. Richardson and James M. Robins. *Single World Intervention Graphs (SWIGs) : A Unification of the Counterfactual and Graphical Approaches to Causality.* Now Publishers Inc, (2013). Working Paper #128, Center for Statistics and the Social Sciences, University of Washington.

[RRS05] Benni Reznik, Alex Retzker, and Jonathan Silman. Violating Bell's inequalities in vacuum. *Phys. Rev. A* **71**, 042104 (2005).

[RSML18] Patrick Rebentrost, Adrian Steffens, Iman Marvian, and Seth Lloyd. Quantum singular-value decomposition of nonsparse low-rank matrices. *Phys. Rev. A* **97** (1), 012327 (2018).

[RSW+19] Patrick Rebentrost, Maria Schuld, Leonard Wossnig, Francesco Petruccione, and Seth Lloyd. Quantum gradient descent and Newton's method for constrained polynomial optimization. *New J. Phys.* **21** (7), 073023 (2019).

[RW06] Carl E. Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning.* MIT Press, (2006).

[RW14] David Reeb and Michael M. Wolf. An improved Landauer principle with finite-size corrections. *New J. Phys.* **16** (10), 103011 (2014).

[SA15] Bastian Steudel and Nihat Ay. Information-theoretic inference of common ancestors. *Entropy* **17** (4), 2304–2327 (2015).

[SAT+17] Alexia Salavrakos, Remigiusz Augusiak, Jordi Tura, Peter Wittek, Antonio Acín, and Stefano Pironio. Bell inequalities tailored to maximally entangled states. *Phys. Rev. Lett.* **119**, 040402 (2017).

[ŠB19] Ivan Šupić and Joseph Bowles. Self-testing of quantum systems: a review, *arXiv:1904.10042.*

[Sch86] Bonny L. Schumaker. Quantum mechanical pure states with Gaussian wave functions. *Phys. Repts.* **135**, 317 (1986).

[Sch04] Sebastian Schlicht. Considerations on the Unruh effect: causality and regularization. *Class. Quantum Gravity* **21** (19), 4647–4660 (2004).

[Scu98] Horia Scutaru. Fidelity for displaced squeezed thermal states and the oscillator semigroup. *J. Phys. A: Math. Gen.* **31** (15), 3659 (1998).

[SCZ16] Robert S. Smith, Michael J. Curtis, and William J. Zeng. A practical quantum instruction set architecture, *arXiv:1608.03355*.

[SGAN18] Ana Belén Sainz, Yelena Guryanova, Antonio Acín, and Miguel Navascués. Almost-quantum correlations violate the no-restriction hypothesis. *Phys. Rev. Lett.* **120**, 200402 (2018).

[Sho94] Peter W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, (1994).

[SIFW10] Mohan Sarovar, Akihito Ishizaki, Graham R. Fleming, and K. Birgitta Whaley. Quantum entanglement in photosynthetic light-harvesting complexes. *Nat. Phys.* **6**, 462–467 (2010).

[SK18] Maria Schuld and Nathan Killoran. Quantum machine learning in feature Hilbert spaces. *Phys. Rev. Lett.* **122**, 040504 (2018).

[Slo19] William Slofstra. The set of quantum correlations is not closed. *Forum of Mathematics, Pi* **7**, e1 (2019).

[SM09] Greg Ver Steeg and Nicolas C. Menicucci. Entangling power of an expanding universe. *Phys. Rev. D* **79**, 044027 (2009).

[SMM15] Grant Salton, Robert B. Mann, and Nicolas C. Menicucci. Acceleration-assisted entanglement harvesting and rangefinding. *New J. Phys.* **17** (3), 035001 (2015).

[SMM18] Petar Simidzija and Eduardo Martín-Martínez. Harvesting correlations from thermal and squeezed coherent states. *Phys. Rev. D* **98**, 085007 (2018).

[SMSC+15] Lynden K. Shalm, Evan Meyer-Scott, Bradley G. Christensen, Peter Bierhorst, Michael A. Wayne, Martin J. Stevens, ... Sae Woo Nam. Strong loophole-free test of local realism. *Phys. Rev. Lett.* **115**, 250402 (2015).

[SO08] S. Taylor Smith and Roberto Onofrio. Thermalization in open classical systems with finite heat baths. *Eur. Phys. J. B* **61**, 271 (2008).

[SPdRMM12] Carlos Sabín, Borja Peropadre, Marco del Rey, and Eduardo Martín-Martínez. Extracting past-future vacuum correlations using circuit QED. *Phys. Rev. Lett.* **109**, 033602 (2012).

[SSD59] Henry E. D. Scovil and Erich O. Schulz-DuBois. Three-level masers as heat engines. *Phys. Rev. Lett.* **2**, 262–263 (1959).

[SSP14] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. The quest for a quantum neural network. *Quantum Inf. Process.* **13** (11), 2567–2586 (2014).

[SSS+17] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, ... Demis Hassabis. Mastering the game of Go without human knowledge. *Nature* **550**, 354–359 (2017).

[SST16] Naoto Shiraishi, Keiji Saito, and Hal Tasaki. Universal trade-off relation between power and efficiency for heat engines. *Phys. Rev. Lett.* **117**, 190601 (2016).

[STH00] Ken Sekimoto, Fumiko Takagi, and Tsuyoshi Hondou. Carnot's cycle for small systems: Irreversibility and cost of operations. *Phys. Rev. E* **62**, 7759–7768 (2000).

[Sto18] E. Miles Stoudenmire. Learning relevant features of data with multi-scale tensor networks. *Quantum Sci. Technol.* **3** (3), 034003 (2018).

[Stu99] Jos F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optim. Methods Softw.* **11–12**, 625–653 (1999). Version 1.05 available from `https://github.com/sqlp/sedumi`.

[Suz92] Masuo Suzuki. General theory of higher-order decomposition of exponential operators and symplectic integrators. *Phys. Lett. A* **165** (5-6), 387–395 (1992).

[SV18] Christopher J. Shallue and Andrew Vanderburg. Identifying exoplanets with deep learning: A five-planet resonant chain around Kepler-80 and an eighth planet around Kepler-90. *Astron. J.* **155** (2), 94 (2018).

[SVA+07] Guillem Sagué, Eugen Vetsch, Wolfgang Alt, Dieter Meschede, and Arno Rauschenbeutel. Cold-atom physics using ultrathin optical fibers: Light-induced dipole forces and surface interactions. *Phys. Rev. Lett.* **99**, 163602 (2007).

[Sve87] George Svetlichny. Distinguishing three-body from two-body nonseparability by a Bell-type inequality. *Phys. Rev. D* **35**, 3066–3069 (1987).

[Sve10] Orazio Svelto. *Principles of Lasers*. Springer Science & Business Media, (2010).

[SW85] Stephen J. Summers and Reinhard F. Werner. The vacuum violates Bell's inequalities. *Phys. Lett. A* **110** (5), 257–259 (1985).

[SW87] Stephen J. Summers and Reinhard F. Werner. Bell's inequalities and quantum field theory. I. General setting. *J. Math. Phys.* **28** (10), 2440 (1987).

[SZ97] Marlan O. Scully and M. Suhail Zubairy. *Quantum Optics*. Cambridge University Press, (1997).

[Tak86] Shin Takagi. Vacuum noise and stress induced by uniform acceleration: Hawking-Unruh effect in Rindler manifold of arbitrary dimension. *Prog. Theor. Phys. Suppl.* **88**, 1–142 (1986).

[TBMM16] Le Phuc Thinh, Jean-Daniel Bancal, and Eduardo Martín-Martínez. Certified randomness from a two-level system in a relativistic quantum field. *Phys. Rev. A* **94**, 022321 (2016).

[TGR19] Erik Torrontegui and Juan José García-Ripoll. Unitary quantum perceptron as efficient universal approximator. *EPL* **125** (3), 30004 (2019).

[TH17] Hiroyasu Tajima and Masahito Hayashi. Finite-size effect on optimal efficiency of heat engines. *Phys. Rev. E* **96**, 012128 (2017).

[Tie08] Tijmen Tieleman. Training Restricted Boltzmann Machines using approximations to the likelihood gradient. In *Proceedings of the 25th International Conference on Machine learning*, pages 1064–1071, (2008).

[TKU19] Zhoudunming Tu, Dmitri Kharzeev, and Thomas Ullrich. The EPR paradox and quantum entanglement at sub-nucleonic scales, *arXiv:1904.11974*.

[TL19] Mingxing Tan and Quoc V. Le. EfficientNet: Rethinking model scaling for convolutional neural networks, *arXiv:1905.11946*.

[TMGB19] Francesco Tacchino, Chiara Macchiavello, Dario Gerace, and Daniele Bajoni. An artificial neuron implemented on an actual quantum processor. *npj Quantum Inf.* **5** (1), 26 (2019).

[TTT99] Kim-Chuan Toh, Michael J. Todd, and Reha H. Tutuncu. SDPT3—a MATLAB software package for semidefinite programming. *Optim. Methods Softw.* **11**, 545–581 (1999). Version 4 available from `http://www.math.nus.edu.sg/~mattohkc/sdpt3.html`.

[Val91] Antony Valentini. Non-local correlations in quantum electrodynamics. *Phys. Lett. A* **153** (6-7), 321–325 (1991).

[VBB17] Guillaume Verdon, Michael Broughton, and Jacob Biamonte. A quantum algorithm to train neural networks using low-depth circuits, *arXiv:1712.05304*.

[vdODZ+16] Aäron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonya, Oriol Vinyals, Alex Graves, ... Koray Kavukcuoglu. WaveNet: A generative model for raw audio, *arXiv:1609.03499*.

[vHBP+18] Thomas van Himbeeck, Jonatan Bohr Brask, Stefano Pironio, Ravishankar Ramanathan, Ana Belén Sainz, and Elie Wolfe. Quantum violations in the Instrumental scenario and their relations to the Bell scenario, *arXiv:1804.04119*.

[VPB18] Guillaume Verdon, Jason Pye, and Michael Broughton. A universal training algorithm for quantum deep learning, *arXiv:1806.09729*.

[VRMCH97] Santiago Velasco, José M. M. Roco, Alejandro Medina, and Antonio Calvo Hernández. New performance bounds for a finite-time Carnot refrigerator. *Phys. Rev. Lett.* **78**, 3241–3244 (1997).

[VW02] G. Vidal and Reinhard F. Werner. Computable measure of entanglement. *Phys. Rev. A* **65**, 032314 (2002).

[WBHS10] Nathan Wiebe, Dominic Berry, Peter Høyer, and Barry C. Sanders. Higher order decompositions of ordered operator exponentials. *J. Phys. A: Math. Theor.* **43** (6), 065203 (2010).

[WBL12] Nathan Wiebe, Daniel Braun, and Seth Lloyd. Quantum algorithm for data fitting. *Phys. Rev. Lett.* **109**, 050505 (2012).

[Weh06] Stephanie Wehner. Tsirelson bounds for generalized Clauser-Horne-Shimony-Holt inequalities. *Phys. Rev. A* **73**, 022110 (2006).

[Wei99] Ulrich Weiss. *Quantum dissipative systems.* World Scientific, 2nd edition, (1999).

[Wit15] Peter Wittek. Algorithm 950: Ncpol2sdpa—sparse semidefinite programming relaxations for polynomial optimization problems of noncommuting variables. *ACM Trans. Math. Softw.* **41**, 21 (2015).

[WLYZ18] Yuanhao Wang, Ying Li, Zhang-Qi Yin, and Bei Zeng. 16-qubit IBM universal quantum computer can be fully entangled. *npj Quantum Inf.* **4** (1), 46 (2018).

[WNW19] Mischa P. Woods, Nelly Huei Ying Ng, and Stephanie Wehner. The maximum efficiency of nano heat engines depends on more than temperature. *Quantum* **3**, 177 (2019).

[Woo98] William K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.* **80**, 2245–2248 (1998).

[WPKG+19] Elie Wolfe, Alejandro Pozas-Kerstjens, Matan Grinberg, Denis Rosset, Antonio Acín, and Miguel Navascués. Quantum inflation: a general approach to quantum causal compatibility. *To appear* (2019).

[WS15] Christopher J. Wood and Robert W. Spekkens. The lesson of causal discovery algorithms for quantum correlations: causal explanations of Bell-inequality violations require fine-tuning. *New J. Phys.* **17** (3), 033002 (2015).

[WSF19] Elie Wolfe, Robert W. Spekkens, and Tobias Fritz. The inflation technique for causal inference with latent variables. *J. Causal Inference* **7** (2) (2019).

[WT11] Peter Wittek and Chew Lim Tan. Compactly supported basis functions as support vector kernels for classification. *Trans. Pattern Anal. Mach. Intell.* **33** (10), 2039–2050 (2011).

[WW01] Reinhard F. Werner and Michael M. Wolf. All-multipartite Bell-correlation inequalities for two dichotomic observables per site. *Phys. Rev. A* **64**, 032112 (2001).

[WY12] Elie Wolfe and Susanne F. Yelin. Quantum bounds for inequalities involving marginal expectation values. *Phys. Rev. A* **86**, 012123 (2012).

[WZP18] Leonard Wossnig, Zhikuan Zhao, and Anupam Prakash. Quantum linear system algorithm for dense matrices. *Phys. Rev. Lett.* **120**, 050502 (2018).

[XAI] Ethics guidelines for trustworthy AI. `https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai`. Accessed: 2019-06-05.

[YY15] Nengkun Yu and Mingsheng Ying. Optimal simulation of Deutsch gates and the Fredkin gate. *Phys. Rev. A* **91**, 032302 (2015).

[ŻB02] Marek Żukowski and Časlav Brukner. Bell's Theorem for general N-qubit states. *Phys. Rev. Lett.* **88**, 210401 (2002).

[ZFF19] Zhikuan Zhao, Jack K. Fitzsimons, and Joseph F. Fitzsimons. Quantum-assisted Gaussian process regression. *Phys. Rev. A* **99**, 052331 (2019).

[ZFO+19] Zhikuan Zhao, Jack K. Fitzsimons, Michael A. Osborne, Stephen J. Roberts, and Joseph F. Fitzsimons. Quantum algorithms for training Gaussian processes. *Phys. Rev. A* **100**, 012304 (2019).

[ZPKRW19] Zhikuan Zhao, Alejandro Pozas-Kerstjens, Patrick Rebentrost, and Peter Wittek. Bayesian deep learning on a quantum computer. *Quantum Mach. Intell.* **1** (1), 41–51 (2019).