PhD Thesis

Optimization and Geometry for Quantum Information tasks

Author:
Paolo Abiuso

Supervisor:
Prof. Dr. Antonio Acín

November 4, 2022

Abstract

In this thesis, we study the optimization of operational tasks that involve the manipulation of quantum resources. In most cases, such optimizations are aided by understanding the geometric properties of the physical objects involved. We split our results in a first part concerning Thermodynamics, and a second part concerning Information Theory.

In the context of Thermodynamics, we first study the optimization of thermal machines. That is, we look for those periodic control protocols, performed on a quantum working fluid, that maximize figures of merit based on power and efficiency. By making small assumptions on the dynamical regimes (of low-dissipation/slow-driving, or fast-driving), we are able to construct and characterize optimal protocols that are valid for large classes of quantum (and classical) thermal machines.
Secondly, we study how to design quantum thermal probes that optimize the precision in temperature estimation when put in contact with a thermal bath. The resulting optimal configurations are simple and physically feasible, and show an Heisenberg-like scaling of the optimal sensitivity.

In the context of Information Theory, initially we study how to characterize memory effects (information backflows) in the dynamics of open quantum systems, how to detect them and operationally exploit them.
Furthermore, in the subfield of Nonlocality, we study relaxations and generalisation of the canonical Bell scenario, which allow us to bring the realization of nonlocal experiments closer to simple, table-top quantum optics. In particular, by considering nonlocality in quantum networks, we are able to design an experiment which only involves simple passive optics and single-photon entangled states, in which it is possible to certify nonlocality without measurement inputs. Likewise, a different relaxation consists in allowing trusted quantum inputs in a Bell experiment. This permits certifying nonlocality of any entangled state, without trusting the measurement device. We study this measurement-device-independent framework to design simple protocols of entanglement detection for continuous-variable states.

The results of the thesis are relevant both from the theoretical point of view and for the efficient realisation of the operational tasks analysed.

Resumen

En esta tesis, estudiamos la optimización de tareas operacionales que involucran manipulación de recursos cuánticos. En la mayoría de los casos, dichas optimizaciones pueden simplificarse al tener en cuenta las propriedades geométricas de los objetos físicos involucrados. Dividimos nuestros resultados en una primera parte que concierne a la Termodinámica, y una segunda parte que concierne a la Teoría de la Información.

En el contexto de la Termodinámica, estudiamos en primer lugar la optimización the máquinas térmicas. Es decir, buscamos protocolos de control perdiódico, efectuados con fluidos de trabajo cuánticos, que maximizen cantidades de interés basadas en la potencia y la eficiencia. Realizando pequeñas suposiciones relativas al régimen dinámico, por ejemplo de manipulación lenta o rápida (slow-driving, fast-driving), logramos construir y caracterizar protocolos óptimos válidos para una amplia gama de máquinas térmicas cuánticas (y clásicas).
En segundo lugar, estudiamos como diseñar sondas térmicas que optimizen la precisión en la estimación de la temperatura, cuando son puestas en contacto con un baño térmico. Las configuraciones óptimas resultantes son sencillas y realizables físicamente, y muestran un crecimiento escalado de la sensibilidad óptima de tipo Heisenberg.

En el contexto de Teoría de la Información, inicialmente estudiamos cómo caracterizar los efectos de memoria (reflujos de información) en la dinámica de sistemas cuánticos abiertos, cómo detectarlos y explotarlos operacionalmente. Además, en el campo de la no-localidad, estudiamos relajamientos y generalizaciones del escenario canónico de Bell, que nos permiten acercarnos a la realización de experimentos de no-localidad con esquemas sencillos de óptica cuántica. En particular, considerando escenarios de no-localidad en redes cuánticas, logramos diseñar un experimento que requiere solamente elementos de óptica pasiva y estados entrelazados de fotones individuales, en el que es posible certificar no-localidad sin utilizar inputs aleatorios en las medidas. Asimismo, una generalización diferente consiste en permitir inputs cuánticos de confianza en un experimento de Bell. Esto permite certificar la no-localidad de todos los estados entrelazados, sin necesidad de confiar en el

aparato de medida. Estudiamos este marco independiente del aparato de medida (measurement-device-independent) para diseñar protocolos sencillos de detección de entrelazamiento cuántico, empleando estados de variable continua.

Los resultados de la tesis son relevantes tanto desde el punto de vista teórico, así como para la realización eficiente de las tareas operacionales analizadas.

List of publications

Peer-reviewed publications forming part of this thesis:

[1] P. Abiuso and M. Perarnau-Llobet, *Optimal cycles for low-dissipation heat engines*, Phys. Rev. Lett. **124**, 110606 (2020).

[2] P. Abiuso, H. J. D. Miller, M. Perarnau-Llobet and M. Scandi, *Geometric optimization of quantum thermodynamic processes*, Entropy **22**(10) (2020).

[3] P. Abiuso, S. Bäuml, D. Cavalcanti and A. Acín, *Measurement-device-independent entanglement detection for continuous-variable systems*, Phys. Rev. Lett. **126**, 190502 (2021).

[4] V. Cavina, P. A. Erdman, P. Abiuso, L. Tolomeo and V. Giovannetti, *Maximum-power heat engines and refrigerators in the fast-driving regime*, Phys. Rev. A **104**, 032226 (2021).

[5] P. Abiuso, T. Kriváchy, E.-C. Boghiu, M.-O. Renou, A. Pozas-Kerstjens and A. Acín, *Single-photon nonlocality in quantum networks*, Phys. Rev. Research **4**, L012041 (2022).

[6] P. Terrén Alonso, P. Abiuso, M. Perarnau-Llobet and L. Arrachea, *Geometric optimization of nonequilibrium adiabatic thermal machines and implementation in a qubit system*, PRX Quantum **3**, 010326 (2022).

Pre-prints forming part of this thesis:

[7] P. Abiuso, M. Scandi, D. De Santis and J. Surace, *Characterizing (non-)Markovianity through Fisher Information*, arXiv preprint 2204.04072, (2022).

[8] P. Abiuso, P.A. Erdman, M. Ronen, G. Haack, F. Noé and M. Perarnau-Llobet, *Discovery of Optimal Thermometers with Spin Networks aided by Machine-Learning*, arXiv preprint 2211.01934, (2022).

Peer-reviewed publications and pre-prints relevant to this thesis, but not forming part of it:

[9] P. Abiuso and V. Giovannetti, *Non-markov enhancement of maximum power for quantum thermal machines*, Phys. Rev. A **99**, 052106 (2019).

[10] P. Abiuso, V. Holubec, J. Anders, Z. Ye, F. Cerisola and M. Perarnau-Llobet, *Thermodynamics and optimal protocols of multidimensional quadratic brownian systems*, Journal of Physics Communications **6**(6), 063001 (2022).

[11] Z. Ye, F. Cerisola, P. Abiuso, J. Anders, M. Perarnau-Llobet and V. Holubec, *Optimal finite-time heat engines under constrained control*, arXiv preprint 2202.12953, (2022).

[12] P.E. Erdman, A. Rolandi, P. Abiuso, M. Perarnau-Llobet and F. Noé, *Pareto-optimal cycles for power, efficiency and fluctuations of quantum heat engines using reinforcement learning*, arXiv preprint 2207.13104, (2022).

This page is intentionally left blank.

Contents

Abstract **iii**

Resumen **v**

List of publications **vii**

INTRO **5**

Motivation **5**

Overview of thesis results **9**

Notation and conventions **19**

I Tasks in Thermodynamics: geometry and optimization **21**

1 Optimising slowly-driven engines through geometry **27**
1.1 Background: driven thermal machines 27
1.2 Background: Low-dissipation regime and the thermodynamic length 34
1.3 General traits and principles of low-dissipation machines optimization 36
1.4 The case of low-dissipation including heat leaks 46
1.5 Discussion 51

2 Optimising fastly-driven engines **53**
2.1 Background: power maximization and fast machines 53
2.2 Framework and fast-driving regime 54
2.3 Fast-driving optimization 58
2.4 Applications 65
2.5 Discussion 68

3 Engineering good thermometers **71**
3.1 Background: equilibrium thermometry 71
3.2 Spectrum-based thermometry optimization 73
3.3 Optimal spin thermal probes 75
3.4 Spin probes comparison 80
3.5 Discussion 83

II Tasks in Information Theory: geometry and optical proposals **85**

4 The geometry of (non-)Markovianity and information flow **91**
4.1 Background: Markovianity 91
4.2 Background: Fisher metric and Fisher Information 95
4.3 Rate matrix decomposition of Markovian evolutions 97
4.4 Contractivity of the Fisher metric and the detection of non-Markovianity 98
4.5 Bayesian retrodiction and the meaning of information backflow 103
4.6 Discussion 106

5 Network nonlocality with passive optics and single-photons **109**
5.1 Background: Bell nonlocality 109
5.2 Background: Single-photon nonlocality 111
5.3 Network nonlocality and the Triangle Network 112
5.4 Witnessing Single-Photon nonlocality in the Triangle network 115
5.5 Optical realisation and noise analysis 118
5.6 Discussion 119

6 MDI certification of quantum properties **121**
6.1 Background: Entanglement detection, Nonlocality and Device-Independent certifications 121
6.2 Measurement-device-independent entanglement witnessing . 122
6.3 MDI entanglement witnessing for continuous variables states 125
6.4 Other MDI certification tasks: quantum memory verification . 133
6.5 Discussion 137

OUTRO **141**

Conclusions **141**

Acknowledgements **145**

A Slow-driving derivation of the thermodynamic length **147**
A.1 Thermodynamic metric for single or multiple time scales . . . 150

B Proof of Lemma 1 (Chap. 3). **155**

C Fisher distance and Fisher metric **159**
C.1 Classical Fisher metric and Fisher information 159
C.2 Quantum Fisher metric 163

Bibliography **169**

Lo que dejo por escrito
no está tallado en granito,
yo apenas suelto en el viento
presentimientos.
Pido lo que necesito:
tinta y tiempo,
tinta y tiempo.

Jorge Drexler

INTRO

Motivation

Physics can be seen as the human endeavour of understanding how nature works, the analysis and comprehension of natural phenomena that are not yet explained. At the same time, physics is also the *search* of novel, unexplored circumstances and behaviours of physical systems that might, among other things, be exploited to perform useful tasks. This tension between the explanatory and exploratory/utilitarian view of physics has often led to progress in both.
On one side, it is clear that theoretical breakthroughs in physics eventually lead to enormous technological advances: from geometrical optics, which was born in ancient Greece and brought us tools to explore nature at different scales, to nuclear physics, thanks to which it was possible to develop most of the medical appliances one can find in hospitals nowadays, it is evident how innovation (and consequently society as a whole) benefits from the understanding of physics. So much so that when Faraday in 1831 was questioned, by the British chancellor of the state treasury, about the usefulness of the newly discovered "electricity", he famously replied "I know not, but I wager that one day your government will tax it".
At the same time, the history of science abounds of instances in which the need of exploitation of a resource, or simply some pragmatic operational question, drove to big developments in our knowledge reality. Most famously, Thermodynamics was born at the beginning of the XIX century out of engineers trying to boost the performance of the early steam engines. This led to active research that decades later revolutionized the statistical description of microscopic physics. Another example is the modern field of Informatics, which flourished after the work of Turing, as he laid the groundwork of the field by answering the practical question of what it means, for generic functions, to be computable.

In this balance of *foundations vs applications*, the broadly-defined field of *Quantum Information Sciences* can be seen as paradigmatic: Quantum Mechanics was first born and fully developed in the first half of the XX century. This theory revolutionized our understanding of nature at the microscopic scale and led to enormous technological progress that we still see today in e.g. material

science and nanotechnology. Throughout time, in the second half of the XX century scientists started recognising the potential of applying the mathematical structure of Quantum Mechanics as a new tool in Information Theory, and explore its consequences: first, in the 60s a big push to the field was given by the study of pragmatic tasks in communication theory with quantum states, until in the 80-90s researchers studying cryptography and generalizing Shannon information theory to the field of quantum mechanics gave rise to the modern field of Quantum Information Theory, which has now been thriving for more than 30 years.
To put it shortly, Quantum Information was born out of the theoretical study of communication tasks, but gave important contributions to our understanding of the foundations of quantum mechanics, thermodynamics, and led to progress in the field of metrology, computer science and many others.

Figure 1: Artistic representation of the general motif of this thesis. A *user* that is able to manipulate a *quantum resource*, might use it for different operational tasks. According to each task, we study its optimization, in terms of efficiency, practicality, or other desiderata (see Table 1). Image generated using the DALL·E artificial intelligence system, by OpenAI.

The above discussion introduces the *leitmotif* of this thesis. We study different *operational tasks* a user might perform exploiting physical (quantum) resources (Fig. 1), in the area of Quantum Information Sciences. For each task, different *desiderata* might be listed. Accordingly, we study schemes to achieve or optimise such objectives. It is clear that each desideratum strongly depends, in general, on the task at hand (a train should be *fast*, a telescope should be

precise). In Table 1 we provide a summary-list of tasks and optimization objectives for the works of this thesis, as well as the dedicated Chapters.
Along this journey we study geometrical features of the physical objects and operational protocols involved. Therefore the title, *"Optimization and Geometry for Quantum Information tasks"*. In many cases, geometry helps in simplifying optimization that would otherwise require complex numerical routines and/or control theory (see Chapters 1,2). In other cases, simple geometrical patterns arise from unconstrained optimizations (Chap.s 1,2,3), and demonstrate being efficient in practical implementations of tasks (Chap.s 3,5). In general, geometry underlies some of the interesting properties of the set of physical states and physical dynamics (Chap. 4), and helps quantifying operational bounds for practical tasks, such as in metrology (Chap.s 3,4,6).
We divide the results of this thesis in two broad fields of (quantum) physics, namely, Thermodynamics and Information Theory, to which are dedicated Part I and Part II respectively. Table 1 represents a schematic list of the tasks we treated and the considered desiderata. In the following Chapter we provide a rundown of the results included in this thesis and their context.

TASKS	DESIDERATA	CHAP.
Driving a thermal machine	Maximising Power–Efficieny trade-offs	1-2
Measuring temperature	Engineering the most precise thermometer	3
Characterizing Markovian and non-Markovian dynamics	Identify the geometry of information backflows and their operational meaning	4
Experiments of nonlocality in networks	Using only passive optics and single-photon states	5
Semi-Device-Independent certifications of entanglement and other quantum properties	Finding experimentally-friendly proposals for continuous-variable optical systems	6

Table 1: The different operational tasks and optimizations that we studied in this thesis. Chapters 1, 2 and 3 form Part I, dedicated to tasks in Termodynamics. Chapters 4, 5 and 6 form Part II, dedicated to tasks in Information Theory. For an overview of our results go to next page.

Overview of thesis results

Here we briefly summarize the context and main results forming part of this thesis. We follow the same order of the main Chapters in Parts I and II (see also Table 1). Each Chapter represents a different line of research, a separate *task* to optimize or analyse geometrically. The results within each Chapter are based on single or multiple papers that we authored.

Part I – Tasks in Thermodynamics: geometry and optimization

Chapters 1-2: Geometric optimization of driven engines.

Context. Quantum Thermodynamics [13, 14] was born and rapidly grew in the last decades. Fuelled by high experimental control of quantum systems and engineering at microscopic scales – it is now possible to fabricate devices which behave as qubits and other few-level quantum systems, and couple them to thermal baths [15–19] – one of the central goals of physicists is to push the limits of conventional thermodynamics, and the extension of standard models and cycles to include quantum effects and small ensemble sizes. Beyond the drive to clarify fundamental physical issues, these models may also turn out to be relevant from a more practical point of view, for example in nanotechnology and biology. Therefore the development and implementation of thermodynamic processes in few-level quantum systems is currently a very active area of research [20]. As a straightforward consequence, large theoretical efforts are devoted to the characterization and optimization at the microscopic scale machines, which is typically translated to maximising the efficiency or the power of the machine, or any trade-off between them. Such problem is however in general hard, as any objective function becomes, for thermal cycles, a functional of the chosen protocol $\vec{\lambda}(t)$ that is performed on the driving parameters λ_i of the machine. This means that finding optimal solutions, even for simple thermal machines, in general requires variational calculus and optimal control techniques [21, 22] (recently, reinforcement learning techniques have been also used to tackle this problem [12, 23]), and (partial)

solutions are known only in specific simple systems [24], or for simple dynamics, such as overdamped [25, 26] and underdamped [27, 28] brownian motion in elementary potentials.

Our contributions (Chap. 1). There is, however, a vast class of thermal machines on which it is possible to obtain pseudo-universal results for what concerns power/efficiency tradeoffs, that is the regime of *low-dissipation* thermal machines. These are machines that work close to equilibrium, typically because the driving of the control parameters is slow when compared to the thermalization scale (*slow-driving* regime). As a consequence, their thermodynamics can be expanded around equilibrium in a perturbation theory [29], where the first order corrections correspond to energetic losses (or irreversible entropy production) that are inversely proportional to the duration of the protocol, i.e. to $\frac{1}{\tau}$. Such losses can in turn be characterized via a *metric* on the thermodynamic manifold, and its induced distance, called the *thermodynamic length* [30–35]. Differential geometry therefore offers a powerful framework for optimising and characterising finite-time thermodynamic processes, both classical and quantum.
In Ref. [2] we reviewed and connected different frameworks where the thermodynamic length emerges in the quantum regime: closed quantum systems driven adiabatically, time-dependent Lindblad master equations, and discrete processes. Following this, we established general principles and bounds for the optimization of thermodynamic processes in the slow-driving regime. These include constant speed of control variation according to the thermodynamic metric, absence of quantum coherence generation during thermodynamic protocols, and optimality of small cycles around the point of maximal ratio between heat capacity and relaxation time for Carnot engines. In particular we derived, in Ref. [1] a simple bound that identifies the maximum power at any given efficiency for Carnot-like engines. Such bound can be efficiently computed via the maximization of a scalar function for any working fluid, and is governed by the behaviour of the heat capacity and the thermalization timescales of the system. Moreover, it is possible to identify the optimal cycles that saturate the bound, consisting in small cycles around the point in which the above-mentioned scalar is maximal. These are general results which can be employed for a variety of settings and applications, such as benchmarking of thermal machines, the comparison of interacting vs. non-interacting working fluids, or the study of non-Markovian effects on the engine performance. Moreover, we found that most of our results can be applied to classical macroscopic thermal machines, as they do not depend on the specific quantum thermodynamics setting.

Finally, in Ref. [6], we studied a more general regime of low-dissipation, that is the case in which the thermal machine is slightly out of equilibrium not only due to the finite-time driving, but also because it is in contact with different thermal baths at different temperatures simultaneously. For the dissipation to be small in such case, besides the slow-driving, the thermal bias ΔT between the different reservoirs needs to be small. A linear-response expansion can then be performed in ΔT as well as $\frac{1}{\tau}$ (cf. Sec. 1.4). Again, a geometric framework arises in this context, and the behaviour of the thermal machine can be characterized in terms of a thermal geometric tensor [36], which includes the thermodynamic length metric within. Applying such geometric framework to the maximisation of power and efficiency in a quantum thermal machine, in Ref. [6] we found that such problem, both for the heat engine and refrigerator operational modes, is reduced to an *isoperimetric problem* with nontrivial underlying metrics and curvature. This corresponds to the maximization of the ratio between the area enclosed by a closed curve and its corresponding length. Our framework connects thermodynamic optimization to a fundamental problem of geometry, and allows for practical numerical optimization of the driving cycles. We illustrated this procedure in a qubit coupled to two reservoirs, operating both as a thermal engines and as a refrigerator.

Our contributions (Chap. 2). Additionally, in Ref. [4], we studied a dynamical regime which is opposite to slow-driving. In fact, recent evidence [24, 37, 38] shows that, when focusing only on the power output of thermal machines, cycles in which the modulation of the driving parameters is performed infinitely fast, are optimal. This regime is dubbed *fast-driving* regime. Contrary to the slow-driving regime, in fast-driving the working fluid is constantly far from equilibrium. At the same time, the dynamics can nevertheless be simplified, as the state of the working fluid tends to a "steady state" that does not have time to relax due to the short timescale of the driving.
For these reasons, in [4] we studied maximization of the "generalized power" of arbitrary periodically driven thermal machines. Within the fast-driving assumption, we derived the optimal class of cycles that universally maximize the extracted power of heat engines, the cooling power of refrigerators, and in general any linear combination of the heat currents. We denoted these optimal solutions as “generalized Otto cycles” since they share the basic structure with the standard Otto cycle, while being characterized, in general, by a greater number of (fast) strokes. We bounded this number in terms of the dimension of the Hilbert space of the system used as working fluid. As in the case of slow-driving (Chap. 1), the generality of these results allows for a widespread range of applications, such as reducing the computational complexity for numerical approaches, or obtaining the explicit form of the optimal protocols when the

system-baths interactions are characterized by a single thermalization scale. Additionally, in [4] we illustrated our general results studying paradigmatic models such as a qutrit-based heat engines, as well as comparing simple models of interacting and non-interacting working fluids used as engine and refrigerators, which show nontrivial diversities in behaviour.

Chapter 3: Thermal probes based on spin networks.

Context. Given a system at thermal equilibrium, one can define its *thermal sensitivity* as the variation in the mean energy of such a system per temperature change unit. Such a figure of merit is also termed *heat capacity* $\mathcal{C}$ of the system. The heat capacity is a fundamental thermodynamic property of physical systems: it has been proven proportional to the maximum power a system can output in Carnot-like engines [1,39], and its critical behaviour is studied in phase transitions. Moreover, $\mathcal{C}$ quantifies the maximum precision one can get in estimating the temperature T via measuring the thermalized probe. That is, given an estimator $\hat{T}$ of the temperature, the well known Cramér-Rao bound, specified to the case of temperature estimation [40, 41], bounds the precision of any $\hat{T}$ as $\frac{\langle(\hat{T}-T)^2\rangle}{T^2} \geq (\nu\mathcal{C})^{-1}$, where ν is the number of i.i.d. repetitions of the experiment and $\mathcal{C}$ the heat capacity of the thermal probe. Moreover the bound can be saturated by measuring the energy of the probe. The best thermometer, according to the above scenario, is therefore a probe with maximum heat capacity. Assuming only the (finite) Hilbert space dimension of the thermal probe, the spectrum of the theoretical system achieving maximum $\mathcal{C}$ has already been worked out in [42], albeit assuming a fully controllable Hamiltonian spectrum, without physical constraints. The resulting optimal spectrum consist in an effective two-level system, with a single ground state and an exponential degeneracy of the excited level.

Our contributions. Though the energy spectrum of an optimal temperature probe has been derived theoretically in [42], it is unclear if any physical probe can exhibit such thermal sensitivity. In particular the maximum heat capacity $\mathcal{C}$ found in [42] shows a "Heisenberg-like" scaling, proportional to N^2 when the probe is assumed to be formed by N constituents. With this preamble, in Ref. [8] we aimed at answering whether such optimal probes can be built on physical grounds, specifically with interacting constituents having only *(i)* 2-body interactions, as well as *(ii)* short-ranged. Surprisingly, we answered both questions affirmatively. To tackle the problem, we chose as a platform a generic system of spins, such as those currently programmable in quantum

annealers [1], whose thermodynamics is starting to be studied recently [43–45], motivated by the flexible design and high degree of control such devices. We maximized the heat capacity of thermal probes based on standard spin Hamiltonians containing local magnetic field terms and 2-spins interactions. The models emerging from such optimisation show a simple architecture of the interactions, which can be embedded in currently available annealers. Their resulting maximal value of $\mathcal{C}$ is mathematically smaller than the theoretical bound, which can nevertheless be approximated asymptotically well, achieving the desired Heisenberg-like scaling. Our work paves the way to the physical realisation of such ultra-sensitive thermometers in spin systems and other physical platforms. Moreover, from the point of view of spectral engineering, the resulting optimal probes show an exponential degeneracy of the first excited level. This kind of spectrum can have application in other fields, such as protein folding modelling [46, 47], adiabatic Grover's search [48, 49], and energy based boolean computation [50], and thermal machines optimization (cf. Chapters 1 and 2).

Part II – Tasks in Information Theory: geometry and optical proposals

Chapter 4: Markovian evolutions and Fisher Information.

Context. The study of open systems dynamics is of paramount importance in physics, as the postulate of unitary evolution is lost whenever a description cut is imposed between the system of interest and the environment it interacts with. A non-isolated physical system typically loses information to its environment [51], and when such loss is irreversible the evolution is said to be Markovian, or memoryless. Non-Markovian (or memory) effects are of great relevance in information theory and in the study of open systems dynamics in general. Markovianity is studied by monitoring how information quantifiers evolve in time, that is, observing if such quantifiers monotonously decrease or if temporary information *backflows* happen. One of the canonical quantifiers studied in the community is the distance between states, particularly the trace distance, due to the fact that it quantifies the 1-shot distinguishability between states [52], and has the desired monotone property under Markovian evolutions. That is, in Markovian evolutions, the trace distance between any

[1] See for example the D-Wave annealers.

two states decreases monotonically, and a local increase is seen as a backflow of distinguishability, and therefore, of information. Most of the studies in the community therefore focus on such framework when studying (quantum) non-Markovianity [53, 54]. At the same time, the use of trace distance has its own mathematical and operational issues. For example, such distance is not a smooth function of its arguments (it induces a Banach space but not an Hilbert space), and in general needs the use of ancillas to witness all possible non-Markovian dynamics [55].

Our contributions. For these reasons, in Ref. [7] we looked for a notion of distance that is suited to characterize completely the (non-)Markovianity of a physical evolution while having well-behaved mathematical properties and task-operational meaning. We showed that the Fisher information distance [56] emerges as the natural candidate in this context. Such distance is well-known to have numerous mathematical properties and operational meanings in metrology, statistical analysis and communication theory. In our work [7] we fully analysed the relation between its contractivity properties and Markovianity. We proved, for classical dynamics, that Markovianity, identified as stochastic-divisibility of the evolution maps, is equivalent to the monotonous contraction of the Fisher metric at all points of the set of states, by extending the thesis of a continuous-time version of the famous Chentsov's Theorem [57]. At the same time, we proved that operational witnesses of non-Markovianity based on the dilation of the Fisher distance cannot, in general, detect all non-Markovian evolutions, although this can be fixed when specific physical postprocessing is applied to the dynamics. Finally, we showed for the first time that non-Markovian dilations of Fisher distance between states at any time correspond to backflow of information about the initial state of the dynamics at time $t = 0$, via Bayesian retrodiction. All the mentioned results can be lifted to the case of quantum dynamics by considering the standard CP-divisibility framework. Our results corroborate the idea that the Fisher Information metric is the natural object whose contractivity properties characterize memory effects in open system dynamics, both from the mathematical and operational point of view.

Chapter 5: Network nonlocality with passive optics.

Context. The notion of nonlocality was formalized by Bell [58] in the 60s for a bipartite scenario involving a central source shared by two parties. His

work showed that a local hidden variable theory cannot describe reality according to quantum physics. Nonlocality later became a subfield of quantum information theory on its own [59], both for foundational reasons and applications in quantum technologies. Besides straightforward generalizations of the standard Bell scenario to cases considering multiple parties, nonlocality was recently generalized to more complex *network scenarios*, involving multiple sources distributed according to a given network topology. Such scenarios show nontrivial novel theoretical properties and are subject of intense studies in recent years [60]. We use this framework to study the nonlocality of a single-photon maximally entangled state, which is obtained when a photon impinges on a balanced beamsplitter, resulting in a superposition of the photon being sent in one direction or an other, $|\psi\rangle = \frac{|10\rangle+|01\rangle}{\sqrt{2}}$. This is arguably one of the simplest entangled states that can be created in the laboratory, and its nonlocal properties have been intensively debated in the quantum optics and foundations communities, as $|\psi\rangle$ is formally entangled, but its Bell nonlocality cannot be shown without recurring to the transfer of such state to different degrees of freedom, or additional photons in the measurement apparatus. That is, a standard bipartite Bell test made only of passive optical elements cannot reveal the nonlocality of such state.

Our contributions. In Ref. [5], we show that the nonlocality of single-photon entangled states can nevertheless be revealed in a quantum network, providing at the same time the first proposal of network nonlocality without measurement inputs [61, 62]. The considered setup only requires passive optical elements, namely beamsplitters, phase shifters and photodetectors, and involves a single measurement per observer. In our protocol, three single-photon entangled states are distributed in a triangular network (i.e. three parties $\{A, B, C\}$ share three copies of $|\psi\rangle$, one per couple), introducing indeterminacy in the photons' paths and creating nonlocal correlations without the need for measurements choices. We discuss a concrete experimental realisation and provide numerical evidence that the nonlocality of such proposal has (small) noise-tolerance to natural noises that can arise in its implementation, through a machine learning analysis. Our results show that single-photon entanglement may constitute a promising solution to generate genuine network-nonlocal correlations useful for Bell-based quantum information protocols, as the optical architecture involved is minimal and no measurement choice is needed.

Chapter 6: Measurement-device-independent certification of quantum properties.

Context. Devising practical methods to detect entanglement and other non-classical properties in quantum systems is one of the cornerstones of quantum information science. To strengthen the reliability of the detection, it is essential (e.g. in adversarial scenarios), to design methods requiring the minimum level of characterization of the devices used in the implementation. The nonlocal properties of quantum mechanics offer the strongest possible certifications from this point of view (that is, with minimal assumptions). The violation of Bell inequalities for example, can be seen as a certification of entanglement that only uses the statistics provided by the experiment, without making any assumptions on the real implementation. This is the so-called device-independent (DI) scenario [59]. Violating Bell inequalities is however experimentally challenging: for example, widely used entangled Gaussian states of light cannot violate Bell inequalities with standard phase-space measurements, as they have a local hidden variables model represented by their (positive) Wigner distribution [63]. However, the situation changes if one allows a minimal level of trust on the preparation devices that are used. In particular, there exist methods that do not require any characterization of the measurement implemented for entanglement detection, known as measurement device independent (MDI) [64,65]. These results have been anyway obtained only for finite-dimensional systems and cannot be used when considering infinitely dimensional quantum states, e.g. of light.

Our contributions. In Ref. [3] we studied the MDI detection of continuous-variable (CV) entanglement. We first generalised, to the continuous variable regime, the seminal results by Ref.s [64,65], showing that all entangled states can be detected in this scenario. Most importantly, we then found a practical protocol that allows for the measurement-device-independent certification of entanglement of all two-mode entangled Gaussian states. This protocol is feasible with current technology as it makes only use of standard optical setups such as coherent states and homodyne measurements.
Our works opens the path to the use of CV quantum systems for MDI tasks of entanglement detection and beyond, such as channel certification (in the same Chapter 6 we discuss a proposal for CV quantum memory verification [66]), randomness generation, secure communication, and certification of other quantum properties. It remains an open problem that of characterising the largest set of properties that can be certified by MDI protocols. Moreover the MDI scenario can be seen as a simple quantum user - quantum provider scenario in

which the (honest) user is hardware-limited compared to the remote (and untrusted) provider of quantum services. This scenario is therefore interesting for near-term quantum technologies.

Notation and conventions

Here we list recurrent notation choices for this manuscript, as well as conventional terms and acronyms widely used in the literature.

Units [2]

The Boltzmann constant, $k_B = 1.380649 \cdot 10^{-23}$ J K^{-1}, is set to 1.
The Planck constant, $\hbar = 1.054571817... \cdot 10^{-34}$ J s, is set to 1.

Acronyms

CPTP	Completely Positive Trace-Preserving (linear map)
CV	Continuous-variables
DI	Device-Independent
d.o.f.	degree of freedom
FTT	Finite Time Thermodynamics
GAP	Generalized Average Power
i.i.d.	Independent and Identically Distributed (random variables)
PTP	Positive Trace-Preserving (linear map)
MDI	Measurement-Device-Independent
ML	Machine Learning
n-M	non-Markovian
POVM	Positive-Operator Valued Measure
w.l.o.g.	without loss of generality

[2] Taken from The NIST reference on Constants, Units, and Uncertainty based on CODATA 2018

Thermodynamics

ω_β or $\omega_\beta(H)$	The canonical thermal state given by $\frac{e^{-\beta H}}{\mathrm{Tr}[e^{-\beta H}]}$.
U	The mean internal energy of a system $\mathrm{Tr}[\rho H]$.
$\mathrm{d}U = \mathrm{d}Q + \mathrm{d}W$	The 1st law of Thermodynamics (work is performed *on* the system).
S	The Von Neumann entropy $-\mathrm{Tr}[\rho \ln \rho]$.
F	The free energy $F = U - TS$.
$\eta_{\mathrm{Carnot}} = 1 - \frac{T_\mathrm{c}}{T_\mathrm{h}}$	The efficiency of an ideal Carnot cycle.
$\vec{\lambda}_t$, $\dot{\vec{\lambda}}_t := \frac{\mathrm{d}}{\mathrm{d}t}\vec{\lambda}_t$	Time-parametrization of a set of controls λ^i.
$\vec{\lambda}(s) := \vec{\lambda}_{s\tau}$, $0 \leq s \leq 1$	Decoupling the shape $\vec{\lambda}(s)$ of the trajectory,
$\vec{\lambda}' := \frac{\partial}{\partial s}\vec{\lambda} = \tau\dot{\vec{\lambda}}$	from its duration τ.

Information Theory

A, B, C, E	Alice, Bob, Charlie, Eve.
Correlation	The (conditional) probability distribution among correlated parties $p(a_1, \ldots, a_n \vert x_1, \ldots, x_n)$.
$\mathcal{L}$	The local set of correlations.
$\mathcal{Q}$	The quantum set of correlations.

Part I

Tasks in Thermodynamics: geometry and optimization

This first part of the thesis is dedicated to the mathematical analysis and optimization of operational tasks in Thermodynamics. More precisely, we consider two of the fundamental tasks related to the interaction of externally controlled systems with thermal environments, that are *i)* finding the best cycles for thermal machines (achieving e.g. high power, high efficiency), and *ii)* designing the most precise thermometers for temperature estimation.
The framework we will move in is that of open quantum systems interacting with large thermal environments at equilibrium [51]. In this context, the field of Quantum Thermodynamics [13, 14] has developed at a high rate in the last three decades, pushed both by the technological goal of manipulating the energy flow at the nanoscale [67], as well as understanding the very foundations and emergence of thermodynamic behaviour, from the underlying microscopic quantum dynamics [68]. Enormous progress has been made on the experimental side, where the thermodynamics of small quantum systems has now been tested on several platforms [15–19, 69–71]. As a consequence, in recent years Quantum Thermodynamics has flourished thanks to such technological possibilities, as well as the push of new theoretical tools, and different frameworks that linked thermodynamics to (quantum) Information Theory [72, 73], resource theories [74], and others.

In Chapters 1 and 2 we study the problem of thermal machines optimization. We consider quantum *working fluids* that can be externally controlled and put in contact with different thermal sources. Depending on the specific control, which is typically expressed in cycles, this constitute a thermal machine which can be used in different *operating modes*. For example, a thermal engine extracts work from the interaction with thermal reservoirs at different temperatures, while a refrigerator consumes external work to extract heat from a given thermal source. Optimizing the performance of these machines consists in identifying the cycle that maximizes one or more desiderata. In particular, the most common choices are efficiency and power, whose exact definition can change based on the specific thermal machine under consideration. The problem of maximising the efficiency alone is easy to solve, and typically corresponds to perform only transformations at equilibrium, which are reversible and saturate the second law of thermodynamics; however such goal can only be achieved for infinitely slow transformations, of little usefulness from a practical point of view. Maximising the power (or any trade-off between power and efficiency) of a given thermal machine is instead much more complex, and no universal solutions are known. However, we study regimes in which general results can be obtained for large classes of machines.
In Chap. 1 we consider the regime in which the losses are small, (*low-dissipation* regime [75]). For machines that operate by coupling sequentially to different

thermal baths, this means that the transformations are performed in a long time τ (although not infinite), and the dissipation is proportional to $\frac{1}{\tau}$, while the working fluid is always close to equilibrium. In such regime it is possible to define, on the underlying manifold of thermal equilibrium states, a metric that characterizes the losses via its integrated length, called *thermodynamic length* (see [35] and references therein). Using these geometric tools in Ref. [1] we were able to find the full solution to the problem of optimal control in Carnot-like engines, while in [2] we reviewed general principles of optimization for generic low-dissipation thermal machines. Finally, in Ref. [6] we studied the case which includes an additional mechanism of dissipation, given by heat leaks between thermal reservoirs. To keep losses small in this case, not only τ needs to be large, but also the thermal bias ΔT between the heat sources needs to be small. It has been found [36] that in said regime a thermodynamic tensor (which includes and generalizes the thermodynamic length metric), characterizes all energy fluxes and dissipation. By using such geometric framework, in [6] we optimized the power and efficiency of a simple thermal machine, and found that the problem of performance optimization gets mapped to an *isoperimetric problem*, i.e. the problem of maximising the ratio between area and length, on a given non-trivial geometric manifold.
In Chap. 2 we study the opposite approximation, that of *fast-driving*, which is useful to identify maximum-power protocols of certain classes of thermal machines. Even though in such regime the working fluid is constantly far from equilibrium, a different perturbative expansion can be applied and used to find general solutions for maximum power protocols, as we did in Ref. [4]. As it turns out, the general characterization of such max-power protocols can be constrained based on the geometry of the dynamical space of the working fluid, and its Hilbert space dimension. The resulting optimal protocols can be described in terms of a finite (bounded) number of scalar variables.

In Chapter 3 we move to the characteristic problem of thermometry [40, 41]. The goal of any thermometer is clear: estimating the temperature of a given thermal sample with the best possible precision. The most common choice to frame this is to minimize the mean square error between the estimated temperature and the "true" temperature of the sample. Theoretical tools of metrology allow to bound mathematically such maximal precision via the famous Cramér-Rao inequalities [76]. In particular, in the context of equilibrium thermometry, minimizing the estimate's error corresponds to maximising the heat capacity of the thermometer that is taken to equilibrate with the thermal sample, and later measured.
For this reason, in Ref. [8] we studied how to design a thermal probe that

has maximum heat capacity. In fact, albeit the energy spectrum of an optimal probe had already been derived theoretically by Correa et al. [42], no specific schematics for physical systems to exhibit such thermal sensitivity was provided. We therefore approached the same problem with the additional constrain of considering a realistic physical thermometer with a high degree of control. We chose in particular a probe made of interacting spins, such as those programmable in available quantum annealers, and considered the problem of maximising the heat capacity based on the available local magnetic fields and 2-body interactions. We therefore found the best configurations of interactions for such probes, showing that they can have a thermometric performance arbitrarily close to the mathematical maximal bound. Moreover, such optimal configurations show a simple geometrical structure.

Chapter 1

Optimising slowly-driven engines through geometry

This chapter is based on Ref.s [1, 2, 6]:
" P. Abiuso and M. Perarnau-Llobet, *Optimal cycles for low-dissipation heat engines*, Phys. Rev. Lett. **124**, 110606 (2020)"
" P. Abiuso, H. J. D. Miller, M. Perarnau-Llobet and M. Scandi, *Geometric optimisation of quantum thermodynamic processes*, Entropy **22**(10) (2020) "
" P. Terrén Alonso, P. Abiuso, M. Perarnau-Llobet and L. Arrachea, *Geometric optimization of nonequilibrium adiabatic thermal machines and implementation in a qubit system*, PRX Quantum **3**, 010326 (2022)"

1.1 Background: driven thermal machines

In this section we review some basic facts about quantum *driven* thermal machines, setting the notation and the tools that we shall later employ to analyse optimal control protocols. Notice that a different class of thermal machines is that of steady-state engines [77] (or, more in general, autonomous engines [78, 79]), which do not require an external control in order to function; their optimisation can therefore be framed as a function of the internal parameters and couplings of the thermal machine itself. Quite differently, the behaviour of driven thermal machine depends on a set of controls $\vec{\lambda}_t$ that are externally modulated in time, which makes the problem more challenging, as we shall see below.
The general framework we consider is that depicted in Fig. 1.1. A quantum *working fluid* $\mathcal{S}$ is described by its state ρ and its Hamiltonian H, which can be controlled externally via some parameters $\vec{\lambda}$. Moreover, a set of large thermal reservoirs can be coupled to $\mathcal{S}$, sequentially (or more than one at a time) in any order, while $\vec{\lambda}$ is varied. The contact of each of these reservoirs with $\mathcal{S}$ will be described in the standard weak-coupling/Markov approximations [51], i.e. by means of first order master equations. That is, to each bath at inverse

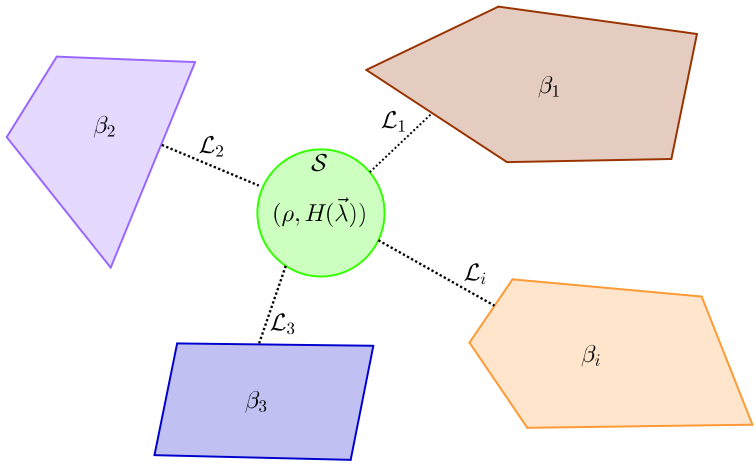

Figure 1.1: A generic quantum thermal machine is described by a system with density matrix ρ, an Hamiltonian $H(\vec{\lambda})$ that is controlled in time via external parameters $\vec{\lambda}_t$, and the possibility of being couple to different thermal baths ad different temperatures $T_i = 1/\beta_i$.

.

temperature $\beta_i = 1/T_i$, is associated a Lindbladian $\mathcal{L}_i$ having, as a unique null eigenstate, the Gibbs state,

$$\mathcal{L}_i[\omega_{\beta_i}] = 0 \ , \quad \omega_{\beta_i} := \frac{e^{-\beta_i H}}{\mathrm{Tr}[e^{-\beta_i H}]} \ . \tag{1.1}$$

This ensures thermodynamic consistency, i.e. that when $\mathcal{S}$ is attached to only one of the thermal baths, its state ρ will relax to the corresponding ω_{β_i}.

1.1.1 The first and second law of thermodynamics in the weak coupling regime

The concept of internal energy U of a system has its natural extension to the quantum regime, that is, for a given state ρ with Hamiltonian H, we can identify U with its mean value,

$$U = \mathrm{Tr}[\rho H] \ . \tag{1.2}$$

The microscopic definition of work and heat is more troubled, and still debated today [73]. We will however use the most common definition that is canonically used in the context of open quantum systems dynamics [13,80,81],

which identifies work as the energy variation due to the external change of the Hamiltonian, while heat as the energy variation that is associated to the change of ρ,

$$\mathrm{d}W := \mathrm{Tr}[\rho\,\mathrm{d}H]\ , \quad \mathrm{d}Q := \mathrm{Tr}[\mathrm{d}\rho\,H]\ , \tag{1.3}$$

such that

$$\mathrm{d}U = \mathrm{d}W + \mathrm{d}Q\ . \tag{1.4}$$

In order to illustrate the general formalism, we consider here a system whose Hamiltonian H_t can be externally driven and which is weakly coupled to a thermal bath. Without loss of generality, we will decompose the system Hamiltonian as $H_t = \sum_i \lambda_t^i X_i$, where $\left\{\lambda_t^i\right\}$ is a family of time dependent external parameters, and $\{X_i\}$ are the corresponding observables. This means that the average work performed on the system is given by:

$$W = \int_\gamma \mathrm{d}W = \int_\gamma \mathrm{d}t\ \mathrm{Tr}[\dot{H}_t\rho_t] = \int_\gamma \mathrm{d}t\ \sum_i \dot{\lambda}_t^i\ \mathrm{Tr}[X_i\rho_t], \tag{1.5}$$

where γ is the path in the parameters space, and ρ_t is the evolved system density matrix at time $t \in (0,\tau)$. We know from equilibrium thermodynamics that if the process is infinitely slow the system is always at equilibrium, that is $\rho_t \simeq \omega_t := \omega_{\beta_t}(H_t)$. Consequently, the work is then given by the difference of free energy at the endpoints of the transformation. Indeed, it is easy to see that in this formalism one regains the usual quasistatic result:

$$W_{\mathrm{eq}} = \int_\gamma \mathrm{d}t\ \mathrm{Tr}[\dot{H}_t\omega_t] = \int_\gamma \mathrm{d}t\ \dot{F} = \Delta F, \tag{1.6}$$

where we used the definition of the free energy

$$F = U - TS\ , \tag{1.7}$$

S being the Von Neumann entropy, $S(\rho) \equiv -\mathrm{Tr}[\rho\ln\rho]$. Similarly, for quasistatic equilibrium transformations, one also has

$$Q_{\mathrm{eq}} = \Delta U - W_{\mathrm{eq}} = \Delta U - \Delta F = T\Delta S\ . \tag{1.8}$$

The second law of thermodynamics, which asserts the irreversibility of general (finite-time) processes can be expressed, equivalently, in several forms. In particular it follows from Clausius theorem (at constant temperature) that for

general processes one has

$$T\Delta S \geq Q \ , \quad \text{or equivalently} \quad W \geq \Delta F \ . \tag{1.9}$$

Given this basic introduction, it is then natural to define the dissipated work as

$$W_{\rm diss} := (W - W_{\rm eq}) = (W - \Delta F) \geq 0 \ , \tag{1.10}$$

in order to isolate the role of the dissipation arising from finite-time effects.

1.1.2 Quantum thermodynamic cycles

Similarly to standard classical thermodynamics, it is possible to describe cycles in which a quantum working fluid in contact with thermal reservoirs is used, e.g. as an engine, that is to extract energy from the thermal bias. We describe the most common (and relevant to us) engine cycles in the following.

Carnot Cycle A Carnot Cycle is identified with a 4-steps process, that is two isothermal strokes alternated with two isoentropic (adiabatic) strokes (cf. Fig. 1.2). Consider a system with a controlled Hamiltonian $H(\vec{\lambda})_t$ which can be coupled independently to two reservoirs with temperature $T_{\rm h} > T_{\rm c}$. In the ideal quasistatic limit, the operations are performed slowly enough to allow the system to be in thermal equilibrium $\rho_t = \omega_{\beta_t}(H(\vec{\lambda}_t))$ at every instant. The 4 steps are:

1. while being coupled to the cold reservoir, the Hamiltonian is modified continuously from $H^{(X)}$ to $H^{(Y)}$, such that $\mathrm{Tr}[\dot{\omega}_{\beta_{\rm c}} H]$ is negative, in order for heat to be released to the cold source.

2. with the system isolated from the reservoirs, a quench is performed taking $H^{(Y)} \to H^{(Y)} \frac{T_{\rm h}}{T_{\rm c}}$.

3. while being coupled to the hot reservoir, the Hamiltonian is modified continuously from $H^{(Y)} \frac{T_{\rm h}}{T_{\rm c}}$ to $H^{(X)} \frac{T_{\rm h}}{T_{\rm c}}$.

4. again isolating the system a quench is performed to restore $H^{(X)} \frac{T_{\rm h}}{T_{\rm c}} \to H^{(X)}$.

Note the factors $\frac{T_{\rm h}}{T_{\rm c}}$ are chosen in order for the state to be continuous during the quenches. In fact the thermal state uniquely depends on βH (cf. (1.1)).

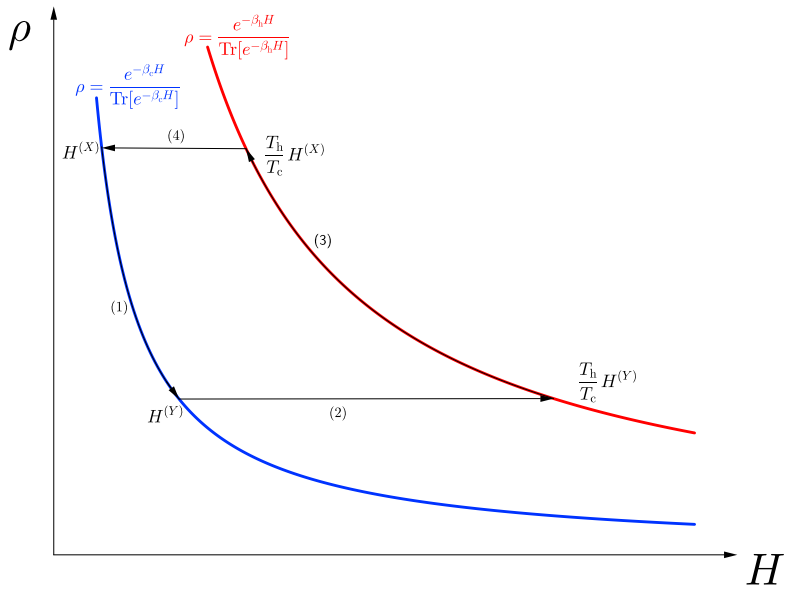

Figure 1.2: Schematics of a Quantum Carnot cycle. Two isothermal strokes in which the state is at equilibrium $\rho = \omega_{\beta_i}$ (1.1) are alternated with two quenches of the Hamiltonian, in which the state ρ is does not change.

.

Otto cycle Similarly, an Otto cycle is given by two adiabatic strokes alternating with two thermalizing strokes (classically isochores) in which the control Hamiltonian is fixed. See Fig 1.3. Therefore the steps of the cycle are:

1. Starting from an initial state ρ_{H} keeping fixed the Hamiltonian H_{C} the system is allowed to thermalize in contact with the cold reservoir, reaching the state $\rho_{\mathrm{c}} \equiv \omega_{\beta_{\mathrm{c}}}(H_{\mathrm{c}})$.

2. After isolating the system from the bath, a quench is performed taking $H_{\mathrm{c}} \to H_{\mathrm{h}}$.

3. While the gap is fixed, the system is allowed to thermalize in contact with the hot reservoir, reaching $\rho_{\mathrm{h}} \equiv \omega_{\beta_{\mathrm{h}}}(H_{\mathrm{h}})$.

4. A final quench restores $H_{\mathrm{h}} \to H_{\mathrm{c}}$.

Notice that, unless considering infinitesimal transformations where $\beta_{\mathrm{h}} H_{\mathrm{h}} \sim \beta_{\mathrm{c}} H_{\mathrm{c}}$, it is clear that at variance with the Carnot cycle, in the Otto cycle the working medium $\mathcal{S}$ is always in an out-of-equilibrium state.

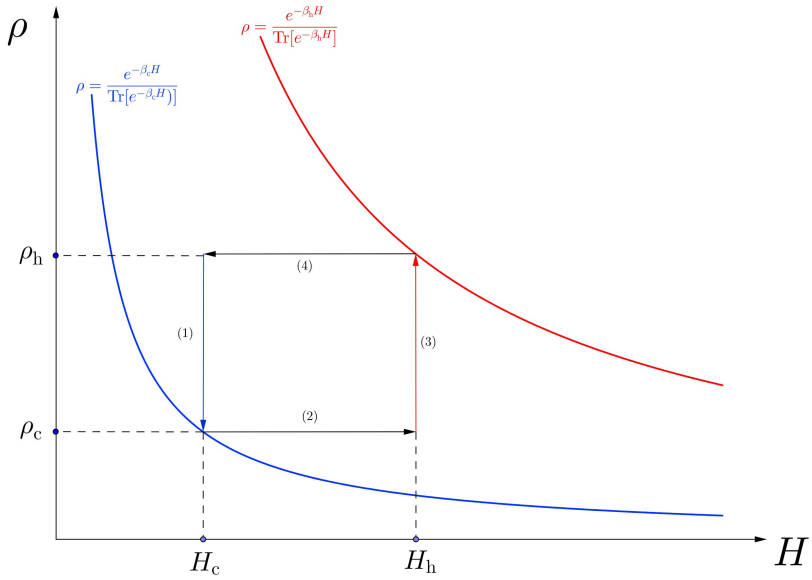

Figure 1.3: Schematics of a Quantum Otto cycle. In this case, two quenches of the Hamiltonian are alternated with thermalizing strokes in which the Hamiltonian is kept constant.

.

1.1.3 Optimisation of a thermal machine: what does that mean?

Before passing to the main treatment and results of this Chapter, it is useful to remark what is intended as *optimisation* of a thermodynamic cycle. Different thermal machines can be used in different *operating modes*. The most typical is the engine mode, in which the controls are cycled in order to extract work from thermal sources at different temperatures [1]. However other modes are possible such as refrigerators, accelerators, etc. [2, 4]. We will refer to the engine mode for clarity of exposition. Optimisation of control in engines typically consists in maximising a given objective function that results from the cycling operation of the system. The most typical objectives are *power* and *efficiency*. These are defined, for an engine, as

$$P = \frac{-W}{\tau} \ , \quad \eta = \frac{-W}{Q_h} \ . \tag{1.11}$$

[1] Both the Carnot cycle and Otto cycle are engines, when operated in the order of steps described in the text in Sec. 1.1.2. Similarly, by inverting the cycles, they can be used as refrigerators extracting heat from the reservoir at smaller temperature T_c.

that is, power is the amount of output work per cycle, divided by the duration of the cycle itself, while the efficiency is the ratio between output work and heat absorbed from the hot source. When considering the efficiency alone, it follows from standard thermodynamics that the maximum efficiency is that obtained from an equilibrium Carnot cycle 1.1.2

$$\eta_{\max} = \eta_{\text{Carnot}} := 1 - \frac{T_{\text{h}}}{T_{\text{c}}} \, . \tag{1.12}$$

This is a universal result that follows from the second law of thermodynamics. The problem of power maximization is instead more complicated and model-dependent. However, typically increasing the power decreases the efficiency, while e.g. the maximum efficiency (1.12) is obtained in the infinite-time limit, which implies the power is null. One could therefore desire to search for hybrid figures of merit, in a trade-off aimed at not sacrificing completely the power for the sake of efficiency, nor vice versa. In the context of multi-objective optimization, the full characterization of such optimal solution is given by the identification of the so called *Pareto front* of the machine. This is defined by taking into account all possible values of the objectives (in our case $\{\eta, P\}$) that the engine can achieve by different operating controls. The Pareto front is then given by those values of $\{\eta, P\}$ such that no other cycle, on the same machine, can achieve both higher efficiency and higher power (cf. Fig. 1.4). The cycles on such boundary are called *Pareto-optimal* cycles. For standard machines, reconstructing the Pareto front is equivalent to finding the curve $\eta_{\max}(P)$ of maximal efficiency at fixed power P (or viceversa).

Is that all about optimization? In all that follows, we will focus on P and η and the associated Pareto front, as the typical desiderata from a thermal machines are its output rate and its efficiency. At the same time, an additional important feature of a thermal machine is its *reliability*. In fact, both power and efficiency are defined on the basis of average quantities. In principle it is desirable that the machine also has small fluctuations in its output. The problem of tradeoff between power, efficiency, and fluctuations is more complex. In the context of steady-state machines, the thermodynamic-uncertainty-relations (TURs [82–86]) partially answer this issue with pseudo-universal bounds, which have been extended in specific cases to driven thermal machines [87–90]. More in general a Pareto-front for the 3 quantities $\{\eta, P, \Delta P\}$ would be required. For this, we refer the reader to our work [12] which is not included in this thesis.

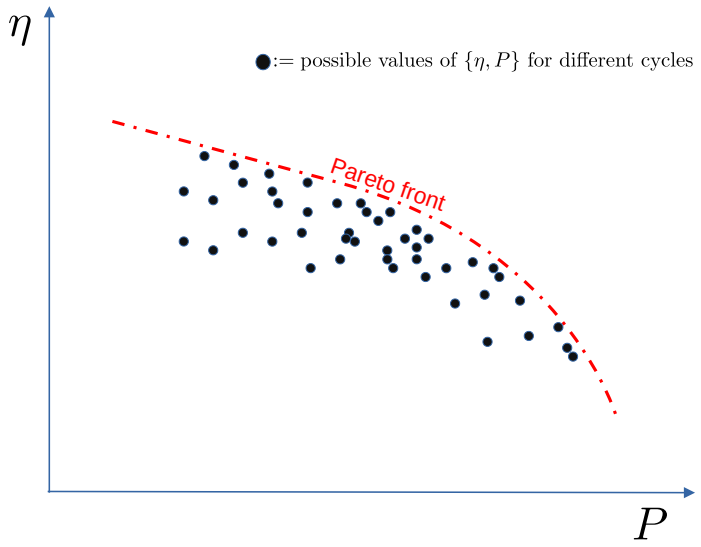

Figure 1.4: Schematic representation of a Pareto front for a thermal engine. A cycle that achieves values of $\{\eta, P\}$ on the Pareto front is considered to be optimal, as it is impossible to obtain at the same time higher power and higher efficiency. The full optimization of a thermal machine consists in finding the Pareto-optimal cycles.

1.2 Background: Low-dissipation regime and the thermodynamic length

One of the reason of the success of equilibrium thermodynamics is its universality: the work performed on a system only depends on the values of the controllable parameters at the endpoints of the transformation via the free energy difference (1.6), the maximum efficiency of an engine is indifferent to the details of the baths (1.12), depending only on their temperature, and in general all the quantities of interest are given by functions of state [91]. These results are extremely strong, but their usefulness is hindered by the necessity of performing all the protocols in infinite time, in order to ensure the equilibrium of the system at all times. On the other hand, if one wants to study arbitrary finite-time, fully out-of-equilibrium processes, the situation considerably complicates and it is hard to believe there can be any universal results. Noticeably, it was found a middle ground between the two situations above, given by the case in which the protocol is performed in long but finite time, which can be characterised by few geometrical quantities. The main ideas

were introduced for classical systems in a series of seminal papers in the 80s by Weinhold and Andresen, Berry and Salamon, among others [30–33]. More recently, the field saw a revival following a series of papers initiated by Crooks [29, 34, 35, 92], leading to several applications.
The first main ingredient of such *low-dissipation* regime is that of assuming the dissipation produced in a finite-time transformation being proportional to the inverse time duration. This can be expressed both in terms of irreversible entropy production

$$\Delta S_{\rm irr} = \frac{\sigma}{\tau} \geq 0 \,, \tag{1.13}$$

or of work dissipation, which perturbatively makes the second law (1.9) explicit, i.e.

$$W - \Delta F = \frac{T\sigma}{\tau} \,, \quad \text{or equivalently} \quad T\Delta S - Q = \frac{T\sigma}{\tau} \,. \tag{1.14}$$

This low-dissipation assumption can be taken as empiric [75, 93] (as the first correction term of a power-law expansion in terms of the speed of the protocol) if no information on the system–bath interaction is given, or it can be justified and derived dynamically using perturbation theory. In Ref. [2] we reviewed some of the approaches that can be used to describe the dynamics of generic quantum working fluids and how to perform the perturbative expansion around equilibrium. In particular, in Appendix A we provide a derivation which is based on the standard Markovian framework and Lindbladian dynamics, together with the assumption of *slow driving* of the system under control [29].
Independently on the dynamical scheme, this means that it is possible to expand $\dot{W}_{\rm diss}$ in terms of $\{\dot{\lambda}^i_t\}$ around the quasistatic limit ($\dot{\lambda}^i_t \equiv 0$), and obtain:

$$\dot{W}_{\rm diss} = \sum_{ij} \dot{\lambda}^i_t \left(\partial_i \partial_j \, \dot{W}_{\rm diss}\big|_{\dot{\lambda}_t \equiv 0} \right) \dot{\lambda}^j_t + \mathcal{O}\left(||\dot{\lambda}||^3 \right) , \tag{1.15}$$

where the first derivative cancels since we are expanding around a minimum. For the same reason, we know that the Hessian $g_{ij} = \beta \partial_i \partial_j \, \dot{W}_{\rm diss}\big|_{\dot{\lambda}_t \equiv 0}$ is positive definite. From these considerations we see that the dissipated work can be written as:

$$W_{\rm diss} = \frac{1}{\beta} \int_\gamma {\rm d}t \, \sum_{ij} \dot{\lambda}^i_t \, (g_{ij})_t \dot{\lambda}^j_t , \tag{1.16}$$

up to higher order corrections. Linear response theory tells us that the matrix

g depends smoothly on the thermal state ω_t. Therefore, from now on, we will use the abbreviation

$$g \text{ or } g_t \text{ or } g_\lambda \text{ to indicate } g(\omega_\beta(H(\vec{\lambda}_t)) \ . \tag{1.17}$$

Moreover, we can deduce that it is positive definite and symmetric, being the Hessian of a function around its minimum. These are the defining properties of a metric. This interpretation is particularly useful thanks to the following fact. If one defines the length of γ as:

$$l_\gamma = \int_\gamma \mathrm{d}t \ \sqrt{\sum_{ij} \dot{\lambda}^i_t \ (g_{ij})_t \dot{\lambda}^j_t}, \tag{1.18}$$

we have the Cauchy–Schwarz like expression

$$\beta W_{\text{diss}} \geq l^2_\gamma / \tau, \tag{1.19}$$

which takes the name of "thermodynamic length inequality" [31]. Among the curves connecting two endpoints, $\vec{\lambda}_0$ and $\vec{\lambda}_\tau$, we call γ geodesic if it minimises the distance between the two points as measured by Equation (1.18). A geodesic is also characterised by the property that it keeps the product $\sum_{ij} \dot{\lambda}^i_t \ (g_{ij})_t \dot{\lambda}^j_t$ constant along its path, implying that the Cauchy–Schwarz inequality in Equation (1.19) is saturated if γ is a geodesic. This means that in order to design minimal dissipating protocols in the slow driving regime for fixed endpoints – a problem that in general would require an ad-hoc minimisation procedure – it is sufficient to solve a system of differential equations, i.e., the geodesic equations.

1.3 General traits and principles of low-dissipation machines optimization

In Ref. [2], we derived and reviewed generic principles of optimisation for finite-time thermal machines in the low-dissipation regime [31, 75]. We consider a thermal machine made up of a (quantum) working substance and several thermal baths at different temperatures. The level of control consists of n experimental parameters λ^i ($i = 1, \ldots, n$) of the machine that can be driven (typically Hamiltonian parameters, as $H = \sum_i \lambda^i_t X_i$), together with the possibility to put the machine in contact with one of the thermal baths. The n control parameters for a vector which is parametrised in time as $\vec{\lambda}(s) \equiv \vec{\lambda}_{s\tau}$ with $s \in (0, 1)$—note that this notation decouples the duration τ of each process from its shape $\vec{\lambda}(s)$. We assume in very general terms that the low-dissipation

condition holds and it is described by an underlying thermodynamic metric, as presented above in (1.16). That is, for an isothermal transformation at temperature $T = \beta^{-1}$, we rewrite Equation (1.16) as

$$\Delta Q = T\left(\Delta S - \frac{\sigma}{\tau}\right) , \tag{1.20}$$

$$\sigma = \int_0^1 \mathrm{d}s\ \vec{\lambda}'^T(s) g_{\vec{\lambda}} \vec{\lambda}'(s) , \tag{1.21}$$

which follows from identifying $W_{\mathrm{diss}} = W - \Delta F = T\Delta S - \Delta Q = T\sigma/\tau$ and by recalling $\vec{\lambda}(s) \equiv \vec{\lambda}_{s\tau}$, which has derivative $\vec{\lambda}' \equiv \frac{\partial}{\partial s}\vec{\lambda} = \tau\dot{\vec{\lambda}}$. Notice that in most of what follows, the exact form of $g_{\vec{\lambda}}$ does not significantly change the results. In this sense, most of the derivations are common to any system that has first-order losses described by some quadratic form, as in linear response theory.

We consider a machine performing M transformations close to equilibrium (in general with different baths), each described by some heat exchange and some dissipation in the low-dissipation regime, with an output

$$\Delta W_{\mathrm{out}} = \sum_{\alpha}^{M} \Delta Q_\alpha = \sum_{\alpha=1}^{M} T_\alpha \Delta S_\alpha - \frac{T_\alpha \sigma_\alpha}{\tau_\alpha} . \tag{1.22}$$

The output being a sum of heat exchanges is guaranteed when considering cycling machines, or when the output of interest is the heat extraction from a subset of the sources (cf. Chap. 2, [4]). This framework thus includes a variety of tasks: cooling, work extraction, Landauer erasure, Carnot cycles, and generalised Carnot engines with multiple baths or finite-size baths. In any such a process, three main features define the control over the machine, and can therefore be optimized:

1. **The speed of the trajectory**: that is, the duration τ, which characterises the average speed of the process, plus any rescaling of the instantaneous velocity along the trajectory. This can be formalised as a change of coordinates $\vec{\lambda}(s) \to \vec{\lambda}(\mathfrak{s}(s))$ with $\mathfrak{s}$ smooth monotonous and $\mathfrak{s}(0) = 0,\ \mathfrak{s}(1) = 1$.

2. **The path of the trajectory**: i.e., the (ordered) set of points swept by $\vec{\lambda}$, for fixed $\vec{\lambda}(0)$ and $\vec{\lambda}(1)$. This identifies a curve γ in $\mathbb{R}^n$.

3. **The extremal points** of γ, or the "location" of the process in the control space.

In the following, we elaborate on the above features and show how to optimise them, in their application to thermodynamic processes. In particular, following the above order, in Section 1.3.1 we optimize the time duration of each transformation τ_i and show a principle of constant dissipation rate optimality; in Section 1.3.2 we discuss consequences of the considerations presented in Section 1.2 when the experimental control is such to allow variations of the curve γ defined by $\vec{\lambda}(s)$; and in Section 1.3.3 we discuss the cases in which a full optimisation can be carried out, so that all the degrees of freedom listed above can be optimised.

1.3.1 Tuning the Speed: Optimality of Constant Dissipation Rate

Here, we suppose initially that the only control available on the machine (1.22) is the time tuning of each step τ_α. We wish to maximise the power output $P = \Delta W_{\text{out}} / \sum_\alpha \tau_\alpha$ for a given loss, or equivalently we fix the (maximum) amount of dissipated work,

$$\sum_\alpha \frac{T_\alpha \sigma_\alpha}{\tau_\alpha} \equiv W_{\text{diss}} \tag{1.23}$$

and maximize P. The power can be written as

$$P = \frac{(\sum_\alpha T_\alpha \Delta S_\alpha) - W_{\text{diss}}}{\sum_{\alpha'} \tau_{\alpha'}} \,, \tag{1.24}$$

hence, maximising it is equivalent to minimising $\sum_\alpha \tau_\alpha$ with the constraint (1.23). This can be stated as

Principle 0. Maximising the power at fixed dissipation is equivalent to minimising the dissipation at given duration.

This remark is important as the main result of this subsection (the optimality of constant thermodynamic speed, or dissipation rate) will thus be valid for all machines performing tasks that are limited by the above trade-off. Examples are: maximising the power, minimising the dissipation (or entropy production) with fixed total time, or hybrid figures of merit combinations, such as maximising the power with a fixed amount of total loss. For a discussion of what machines maximise their outputs when the irreversible entropy production is minimised see [94].

The maximisation of (1.24) can be done differentiating w.r.t τ_α and using Lagrange multipliers, or directly with a Cauchy–Schwarz inequality

$$W_{\rm diss}\sum_\alpha \tau_\alpha = \left(\sum_\alpha \frac{T_\alpha\sigma_\alpha}{\tau_\alpha}\right)\left(\sum_\alpha \tau_\alpha\right) \geq \left(\sum_\alpha \sqrt{T_\alpha\sigma_\alpha}\right)^2 \tag{1.25}$$

which is saturated when all $T_\alpha\sigma_\alpha/\tau_\alpha^2$ are equal, that is

$$\tau_\alpha = \frac{\sqrt{T_\alpha\sigma_\alpha}(\sum_{\alpha'}\sqrt{T_{\alpha'}\sigma_{\alpha'}})}{W_{\rm diss}} \tag{1.26}$$

$$P_{W_{\rm diss}} = \frac{W_{\rm diss}(\sum_\alpha T_\alpha\Delta S_\alpha) - W_{\rm diss}^2}{(\sum_{\alpha'}\sqrt{T_{\alpha'}\sigma_{\alpha'}})^2}\,. \tag{1.27}$$

Notice that the fact that $T_\alpha\sigma_\alpha/\tau_\alpha^2$ is the same $\forall\alpha$ means that the rate of dissipation is constant for each of the M steps of the protocol. In particular, when the dissipation is described by an underlying thermodynamic metric (1.21), this implies the optimality of constant thermodynamic velocity $\vec{\lambda}'^T g_{\vec{\lambda}}\vec{\lambda}' = \text{const.}$, which can be seen by dividing each transformation into infinitesimal steps, i.e., expressing

$$T_\alpha\Delta S_\alpha - \frac{T_\alpha\sigma_\alpha}{\tau_\alpha} = \int_{\gamma^{(\alpha)}} T\mathrm{d}S - \frac{T\mathrm{d}\vec{\lambda}^T g_{\vec{\lambda}}\mathrm{d}\vec{\lambda}}{\mathrm{d}\tau} \tag{1.28}$$

and applying the above reasoning, which concludes that each of the infinitesimal $\frac{T\mathrm{d}\vec{\lambda}^T g_{\vec{\lambda}}\mathrm{d}\vec{\lambda}}{\mathrm{d}\tau^2}$ must be equal. The "thermodynamic length inequality" (1.19) ([31, 33, 95]) is indeed saturated when its integrand is constant, and coincides with the continuous version of (1.25). These considerations can be summed up saying that for the class of machines considered here

Principle 1. In optimal protocols, the speed of the control variation is constant (as measured from the underlying thermodynamic metric), leading to a constant entropy production rate.

The optimality of constant entropy production rate was already noted in the first seminal papers [96] in the context of endoreversible engines, and appeared in many works thereafter (for an historical perspective, see also [97, 98]). The above formulation manifests the universality of this principle whenever a trade-off between output rate and losses is present in the regime where losses are linear in the average speed of the process.

The power (1.26) can be further maximised choosing $W_{\rm diss} = \frac{1}{2}\sum_\alpha T_\alpha \Delta S_\alpha$ to obtain the durations leading to the maximum power, in this case

$$P_{\max} = \frac{(\sum_\alpha T_\alpha \Delta S_\alpha)^2}{4(\sum_{\alpha'} \sqrt{T_{\alpha'}\sigma_{\alpha'}})^2} \,. \tag{1.29}$$

At maximum power the losses thus correspond to half of the quasistatic output: this corresponds to the "7th principle of control thermodynamics" pointed out by Salamon et al. in [97], whose general validity was unknown: we can state it holds (at least) for all machines described by (1.22).
We give here an example of application of the time tuning optimisation just described.

Example: Multi-Bath Carnot Engine. A generalised Carnot engine consists of a sequence of isotherms in contact with different thermal baths, alternated with adiabatic strokes as in the standard Carnot cycle. The total work output can be expressed as the sum of the heat exchanges due to cycling conditions, as in Equation (1.22), with $\sum_\alpha \Delta S_\alpha = 0$. All the results described above apply and the maximum power obtainable by tuning the time durations of the isotherms is thus as in Equation (1.29). Moreover, in [2] we further analyze this result assuming that all the baths have the same spectral density $\propto \omega^\theta$, described by the ohmicity parameter θ (see also App. A). Under this hypothesis and the assumption that all the isotherms are small enough (cf. [2]), this can be translated in the maximum power being expressed by

$$P_{\max}^{\rm multi-Carnot} = \frac{(\sum_\alpha T_\alpha {\rm d}S_\alpha)^2}{4\kappa_0 T_0 \left(\sum_\alpha (\frac{T_\alpha}{T_0})^{\frac{1-\theta}{2}} |{\rm d}S_\alpha|\right)^2} \tag{1.30}$$

where κ_0 represents the local ratio between σ_0 and $(\Delta S_0)^2$ at some reference temperature T_0, and satisfies $\kappa_\alpha/\kappa_{\alpha'} = (T_\alpha/T_{\alpha'})^{-\theta}$ (cf. App. A). Moreover, it is possible to show how in this case, the power is upper bounded by the same power when it is obtained by the use of the highest and lowest temperature only, which leads to the maximum power of a standard Carnot Engine (cf. Section 1.3.3 or [1])

$$P_{\max}^{\rm multi-Carnot} \leq P_{\max}^{\rm Carnot} = \frac{(\Delta S)^2}{\sigma_{\rm h}} \frac{(T_{\rm h}-T_{\rm c})^2}{4T_{\rm h}\left(1+(\frac{T_{\rm c}}{T_{\rm h}})^{\frac{1-\theta}{2}}\right)^2} \,. \tag{1.31}$$

1.3.2 Path Optimisation: Geodesics and Coherences

When the control over the working fluid allows not only to vary the speed of the transformation, but includes possible modifications of the path γ of the trajectory $\vec{\lambda}(s)$, the machine can be substantially improved. The optimisation over γ is independent from the time tuning considered in the previous section. It consists of finding the shortest path $\sigma = \int_\gamma \vec{\lambda}'^T g_{\vec{\lambda}} \vec{\lambda}'$ between two fixed points for each isotherm (1.21) considered in the cycle. Indeed, when the extremal points of a trajectory are fixed, the quasistatic output is fixed and minimizing σ always improves both power and the efficiency (cf. Section above 1.3.1).
More precisely, with the tools described in Section 1.2, each of the σ_α in Equation (1.22) will be described as in (1.18) by some metric $g^{(\alpha)}$ and some trajectory $\vec{\lambda}_{(\alpha)}$, in the form $\sigma_\alpha = \int_{\gamma^{(\alpha)}} \vec{\lambda}'^T_{(\alpha)} g^{(\alpha)}_{\vec{\lambda}} \vec{\lambda}'_{(\alpha)}$. As mentioned earlier (see Section 1.2 or Section 1.3.1), by choosing the speed to be constant the above expression can be minimised to the thermodynamic length of the path $\gamma^{(\alpha)}$

$$\sigma_\alpha = \left(\int_{\gamma^{(\alpha)}} \mathrm{d}s \, \sqrt{\vec{\lambda}'^T_{(\alpha)} g^{(\alpha)}_{\vec{\lambda}} \vec{\lambda}'_{(\alpha)}} \right)^2 \equiv l^2_{\gamma^{(\alpha)}} \, . \tag{1.32}$$

This quantity depends only on the path $\gamma^{(\alpha)}$ of the trajectory and not on its parametrisation $\vec{\lambda}(s)$, but it can be further minimised by considering its minimum among all the possible paths linking the extremal points, which then defines the geodesics distance between the extremal points

$$d_{\vec{\lambda}(0),\vec{\lambda}(1)} = \min_{\substack{\gamma \text{ with extremals} \\ \{\vec{\lambda}(0),\vec{\lambda}(1)\}}} l_\gamma \, . \tag{1.33}$$

These considerations can be stated as follows:

Principle 2. In optimal protocols, the driving minimises the entropy production, i.e., it follows a geodesic on the thermodynamic manifold.

Notice that, once the metric $g_{\vec{\lambda}}$ is known, the geodesics between any two points $\{\vec{\lambda}(0), \vec{\lambda}(1)\}$ can be found by solving accordingly the geodesics equation [2].
Notice that all considerations so far apply both to classical and quantum driven thermal machines. However, in the quantum case, it is possible to additionally constrain optimal protocols based on coherence creation. That is, assuming full control on the Hamiltonian H_t, it is always possible to split its variation

in a coherence part and in a diagonal part

$$\dot{H}_t = \dot{H}_t^{(d)} + \dot{H}_t^{(c)} \ , \tag{1.34}$$

which can be expressed in the diagonal basis $H_t = \sum_i \epsilon_i(t) \, |i(t)\rangle \, \langle i(t)|$ as

$$\dot{H}_t^{(d)} := \sum_i \dot{\epsilon}_i(t) \, |i(t)\rangle \, \langle i(t)| \ , \tag{1.35}$$

$$\dot{H}_t^{(c)} := \sum_i \epsilon_i(t) \left(|\dot{i(t)}\rangle \, \langle i(t)| + |i(t)\rangle \, \langle \dot{i(t)}| \right) \ . \tag{1.36}$$

As it turns out, thanks to the geometric properties of the thermodynamic length (see App. A), for standard detailed-balanced Lindbladians, we showed in [2] that the irreversible entropy production can be split in two independent parts, one that is solely due to $\dot{H}_t^{(d)}$ and one due to the rotation of the Hamiltonian $\dot{H}_t^{(c)}$, that is

$$W_{\rm diss} = W_{\rm diss}^{(d)} + W_{\rm diss}^{(c)} \ , \tag{1.37}$$

whose detailed definition we provide in App. A as well as Ref. [2]. Most importantly, both components are positive, $W_{\rm diss}^{(x)} \geq 0 \ x = d, c$, but the rotation of the eigenvectors of the Hamiltonian (1.36) has no other effect except that of generating the dissipation $W_{\rm diss}^{(c)}$. In [2] we prove that, following the above discussion, it can be stated, for standard Markovian dynamics,

Principle 3. Quantum coherences are not created in optimal protocols, i.e., non-commutativity $[H_t, H_{t'}] \neq 0$ is avoided.

The effect of coherences inducing losses in the power was noted already in [99] in the context of linear response theory of slowly driven engines with slowly driven temperature.
We show in the following an application of the minimization of the irreversibility parameter σ for a cooling process.

Example: Cooling/Work Extraction Suppose we are interested only in a subset of the heat currents that are part protocol, meaning that relevant output is the heat extracted from one (or multiple) thermal sources, as in a generalised refrigerator model. To fix the ideas for a single bath to be cooled the cooling rate is

$$P^{\rm cooling} = \frac{T_{\rm c}\Delta S_{\rm c} - \frac{T_{\rm c}\sigma_{\rm c}}{\tau_{\rm c}}}{\tau_{\rm ex} + \tau_{\rm c}} \equiv \frac{T_{\rm c}\Delta S_{\rm c} - W_{\rm diss}}{\tau_{\rm ex} + \tau_{\rm c}} \tag{1.38}$$

where now $\tau_{\rm ex}$ is additional time spent on parts of the cycle that do not contribute to the cooling output. The optimisation for fixed loss $W_{\rm diss}$ applies as from (1.26) leading to $\tau_{\rm c} = T_{\rm c}\sigma_{\rm c}/W_{\rm diss}$, and a power

$$P^{\rm cooling}_{W_{\rm diss}} = \frac{T_{\rm c}\Delta S_{\rm c} - W_{\rm diss}}{\tau_{\rm ex} + T_{\rm c}\sigma_{\rm c}W^{-1}_{\rm diss}} \,, \tag{1.39}$$

which clearly increases as $\sigma_{\rm c}$ is minimised. The overall maximum of the cooling rate becomes for a suitable choice of $W_{\rm diss}$

$$P^{\rm cooling}_{\rm max} = T_{\rm c}\sigma_{\rm c}\frac{\left(\sqrt{\Delta S_{\rm c}\tau_{\rm ex}/\sigma_{\rm c}+1}-1\right)^2}{\tau^2_{\rm ex}} = T_{\rm c}\frac{\Delta S^2_{\rm c}}{4\sigma_{\rm c}} + \mathcal{O}(\tau_{\rm ex}) \,. \tag{1.40}$$

The above expressions are all decreasing in the value of $\sigma_{\rm c} \geq 0$, which is minimal when obtained on the geodesics of the transformation, as from Equations (1.32) and (1.33).

1.3.3 Choosing the location: total optimisation of Carnot-like machines

After optimizing the time duration and trajectory of the transformations, the resulting optimal output rates only depend on the end points of the transformations. The final maximisation of such expressions is in general non-trivial. However, we note how the maximum power obtained in (1.31) is proportional to $(\Delta S)^2/\sigma$, which is maximal when σ takes the geodesics value described above in (1.33). Thus, this last quantity

$$\frac{(\Delta S)^2}{\sigma} = \frac{\left(S_{\vec{\lambda}(0)} - S_{\vec{\lambda}(1)}\right)^2}{d^2_{\vec{\lambda}(0),\vec{\lambda}(1)}} \tag{1.41}$$

can be maximised by changing the extremal points of the transformation. The same quantity appears as the leading term for the cooling rate in (1.40). We find this to be a strikingly general feature of all thermal machines whose dynamical information ultimately consists of just one simple isothermal transformation close to equilibrium. This is clearly the case for a single heat extraction from a bath as in (1.40), but it happens also, e.g., for Carnot engines, where all relevant quantities which can be expressed solely in terms of the

two isotherms. For example, power and efficiency of a Carnot engine read:

$$P^{\text{Carnot}} = \frac{\Delta S(T_\text{h} - T_\text{c}) - \left(\frac{T_\text{c}\sigma_\text{c}}{\tau_\text{c}} + \frac{T_\text{h}\sigma_\text{h}}{\tau_\text{h}}\right)}{\tau_\text{c} + \tau_\text{h}}, \tag{1.42}$$

$$\eta = \frac{Q_\text{h} + Q_\text{c}}{Q_\text{h}} = 1 - \frac{T_\text{c}(\Delta S + \frac{\sigma_\text{c}}{\tau_\text{c}})}{T_\text{h}(\Delta S - \frac{\sigma_\text{h}}{\tau_\text{h}})}, \tag{1.43}$$

where ΔS is the variation of entropy during the hot isotherm, and the irreversible entropy productions are proportional to each other on optimal protocols $\sigma_\text{h}/\sigma_\text{c} = (T_\text{c}/T_\text{h})^{-\theta}$, according to the spectral density of the baths $\propto \omega^\theta$ [1, 29]. The two isotherms are thus *symmetric*, in the sense that by construction they have an opposite entropy variation $\Delta S_\text{h} = -\Delta S_\text{c}$, and the trajectories follow the same geodesics to link the endpoints [1, 29]. After time optimisation on $\tau_\text{c}, \tau_\text{h}$ in such a case it is clear from dimensional analysis that the resulting power can only be proportional to $(\Delta S)^2/\sigma_\text{h}$ (or equivalently $(\Delta S)^2/\sigma_\text{c}$ due to proportionality) multiplied by a function with the dimension of temperature.
More explicitly, we showed in [1] that is possible to express the maximum power at any given efficiency

$$\eta = (1-\delta)\eta_\text{Carnot} = (1-\delta)(1 - T_\text{c}/T_\text{h}) \tag{1.44}$$

for a Carnot engine (see also [93, 100]). We report here for simplicity only on the case where $\theta = 0$, thus $\sigma_\text{c} = \sigma_\text{h} = \sigma$, as

$$P_\delta^{\text{Carnot}} = \frac{(\Delta S)^2}{4\sigma}\frac{(T_\text{h} - T_\text{c})^2\delta(1-\delta)}{(1-\delta)T_\text{c} + \delta T_\text{h}} \tag{1.45}$$

The importance of the term $(\Delta S)^2/\sigma$ was noted already in [101] as a natural unit of entropy over time, defining the performance of thermal machines in the low-dissipation regime for any trade-off between power and efficiency. The equivalent optimisation for a refrigerator has been conducted in [102], where one has a cooling power and COP coefficient (this time ΔS is defined to be positive on the cold isotherm)

$$P^{\text{Refrigerator}} = \frac{T_\text{c}\Delta S - \frac{T_\text{c}\sigma_\text{c}}{\tau_\text{c}}}{\tau_\text{c} + \tau_\text{h}}, \tag{1.46}$$

$$\varepsilon = \frac{Q_\text{c}}{|Q_\text{h}| - Q_\text{c}} = \frac{T_\text{c}\left(\Delta S - \frac{\sigma_\text{c}}{\tau_\text{c}}\right)}{T_\text{h}\left(\Delta S + \frac{\sigma_\text{h}}{\tau_\text{h}}\right) - T_\text{c}\left(\Delta S - \frac{\sigma_\text{c}}{\tau_\text{c}}\right)}, \tag{1.47}$$

which leads to a maximum cooling power at given COP (again we report it for flat spectral density $\sigma_{\rm c} = \sigma_{\rm h}$, see [102] for generalisations) $\varepsilon = (1-\delta)\varepsilon_{\rm Carnot} = (1-\delta)T_{\rm c}/(T_{\rm h}-T_{\rm c})$

$$P_{\delta}^{\rm Refrigerator} = \frac{(\Delta S)^2}{4\sigma}\frac{T_{\rm c}(T_{\rm h}-T_{\rm c})\delta}{T_{\rm h}-\delta T_{\rm c}} \,. \tag{1.48}$$

Crucially, both (1.45) and (1.48) represent the full Pareto front – in the low-dissipation approximation – for a Carnot-engine and a Carnot-refrigerator. As discussed in Sec. 1.1.3, the full optimisation of such machines is therefore represented by these expression. In both cases, the magnitude of the power is proportional to $(\Delta S)^2/\sigma$, whose maximisation thus corresponds to the total control optimisation.
In Ref. [1] we showed how such term can always be upper bounded via the use of a Cauchy-Schwarz inequality, that is noticing that

$$\begin{aligned}\frac{(\int {\rm d}S)^2}{\int {\rm d}s \vec{\lambda}'^T g_{\vec{\lambda}}\vec{\lambda}'} &= \frac{\left(\int {\rm d}s\ \vec{\partial} S_{\vec{\lambda}}\cdot\vec{\lambda}'\right)^2}{\int {\rm d}s \vec{\lambda}'^T g_{\vec{\lambda}}\vec{\lambda}'} \\ &\leq \int {\rm d}s\ \vec{\partial} S_{\vec{\lambda}}^T g_{\vec{\lambda}}^{-1}\vec{\partial} S_{\vec{\lambda}} \leq \max_{\vec{\lambda}}\ \vec{\partial} S_{\vec{\lambda}}^T g_{\vec{\lambda}}^{-1}\vec{\partial} S_{\vec{\lambda}} \equiv \max_{\vec{\lambda}}\ C(\vec{\lambda})\ . \end{aligned} \tag{1.49}$$

Moreover the upper bound in (1.49) can be saturated by performing infinitesimal cycles around the point where $C(\vec{\lambda})$ is maximised [1].
Thanks to this bound and the expressions (1.45)-(1.48) the full optimisation of these thermal machines is reduced to a simple scalar maximisation which can be conducted numerically (see examples in our works [1, 2]). Finally, in the meaningful case in which all observables X_i decay with a well defined timescale $\tau_{\rm eq}$ to their thermal value, in [1] we proved that $C(\vec{\lambda})$ is exactly the heat capacity of the system divided by the equilibration time, leading to [1]:

$$\frac{(\Delta S)^2}{\sigma} \leq \max_{G}\frac{\mathcal{C}(G)}{\tau_{\rm eq}}. \tag{1.50}$$

Here, $G = \beta H$ is the adimensional Hamiltonian, so that the thermal state can be expressed as $\omega = e^{-G}/{\rm Tr}e^{-G}$ and the heat capacity corresponds to the adimensional energy variance $\mathcal{C}(G) = {\rm Tr}G^2\pi - {\rm Tr}G\pi^2$. In other words, it is possible to state a final optimisation principle [2] as

Principle 4. Power-efficiency trade-offs are optimised by performing finite-time Carnot cycles around the point where the ratio between heat capacity and relaxation time of the working fluid is maximised.

In Refs [1] and [2] these results are provided together with examples of application to thermal machines based on low-dimensional quantum systems. Moreover, the above analysis shows the importance of identifying systems with large heat capacity in the context of thermal engines. This task is intimately related to equilibrium thermometry, where the heat capacity directly quantifies the precision of temperature estimation of a thermal probe. Its maximisation can be approached by considering system that are close to a phase transition [1, 39] or via interaction engineering, see Chapter 3 [8].

1.4 The case of low-dissipation including heat leaks

Following the results of the above sections (Refs. [1, 2]), in Ref. [6] we considered a low-dissipation regime in which there is an additional mechanism of dissipation: *heat leaks*. The crucial difference with respect to our previous works [1, 2] is the possibility of allowing the working fluid to be in contact with more than one thermal reservoir at a time. It is in fact typically hard to fully isolate a quantum working fluid from the environment, which is required to emulate ideal classical cycles 1.1.2. This motivates the study of non-equilibrium systems, where the driven working fluid is permanently in contact with two or more reservoirs. Unlike standard thermodynamic cycles, these microscopic machines operate away from equilibrium during all the cycle. Thermoelectric devices [77] as well as autonomous machines [78, 79, 103, 104] are seminal examples of this type of operation.
When a driven thermal machine is connected at the same time to two or more thermal reservoirs, it is permanently thread by a heat flux. Hence, the very operation as a machine relies on the mechanism of heat–work conversion in order to overcome this effect as well as the dissipation generated by the driving sources. To consider a low-dissipation regime in this case, besides a slow driving of the control parameters, the thermal bias will need to be small, as in the first approximation the heat leaks are proportional to ΔT. Under these assumptions, the thermodynamics of a quantum system was derived in [36]. In particular it was showed that a tensor R, which locally depends on the equilibrium manifold only ($R \equiv R(T, \vec{\lambda})$), can be identified characterizing all the energy fluxes. That is, given n control parameters λ_i, the tensor R is $(n+1) \times (n+1)$-dimensional and symmetric, and can be decomposed as

$$R_{i,j} = g_{ij} \quad i, j = 1, \dots, n \tag{1.51}$$

$$R_{i,n+1} = \vec{R} \quad i = 1, \dots, n \tag{1.52}$$

$$R_{n+1,n+1} = \kappa \tag{1.53}$$

such that the output work $W_{\rm out}$ of a cyclic driven machine, and the heat transfer $Q_{\rm tran}$ from the hot reservoir (at temperature $T+\Delta T$) to the cold reservoir (at temperature T) can be expressed as

$$W_{\rm out} = \frac{\Delta T}{T}\int_0^\tau \mathrm{d}t\ \vec{R}\cdot\dot{\vec{\lambda}} - \int_0^\tau \mathrm{d}t\ \dot{\vec{\lambda}}\cdot g\cdot\dot{\vec{\lambda}}\,, \tag{1.54}$$

$$Q_{\rm tran} = \int_0^\tau \mathrm{d}t\ \vec{R}\cdot\dot{\vec{\lambda}} + \frac{\Delta T}{T}\int_0^\tau \mathrm{d}t\ \kappa\,. \tag{1.55}$$

These expressions can be derived in the adiabatic linear-response regime as from Ref. [36] and we defer the reader to that paper for further details. For our purposes, it is important to stress that the whole tensor R and its components are all local functions of $\vec{\lambda}_t$, while they also depend on the coupling parameters, the density of states of the thermal baths and T (which however do not vary).
In Eq. (1.54), the first term represents the mechanism of *heat–work* conversion and the second one corresponds to finite-time dissipation developed by the time-dependent controls, i.e. the same dissipation mechanism we considered in Sections 1.2 and 1.3. For this reason we identify $R_{ij}\equiv g_{ij}$, in (1.51), as they coincide mathematically (see [6]). Finally, the terms (in W and Q) proportional to $\int_0^\tau \mathrm{d}t\ \vec{R}\cdot\dot{\vec{\lambda}}$ correspond to the equilibrium quasistatic energy fluxes, and $R_{n+1,n+1}\equiv\kappa$ is the resulting heat conductance.
It is important to notice that the second terms of Eqs. (1.54) and (1.55) have a defined sign, since they are directly related to the entropy production rate [36] (i.e. the second law of thermodynamics) which means they are detrimental. Instead, the line integral $\int_0^\tau \mathrm{d}t\ \vec{R}\cdot\dot{\vec{\lambda}}$ may have any sign, depending on the driving protocol and it is enough to time-reverse the control $\vec{\lambda}_t$ to flip the sign, and change the operating mode of the machine.

1.4.1 A different time scaling, and an old geometric problem

Given the expressions of the work output and heat transfer in Eqs. (1.54) and (1.55), it is possible to interpret their terms geometrically, as explained in the following. This allows for the optimization of the thermodynamic protocols in terms of clear geometrical quantities.
First, we factorize the total duration τ in the expressions Eqs. (1.54) and (1.55), such to decouple the time-rescaling from the geometrical contribution to the different quantities, similarly to the approach taken in Sec. 1.3. Indeed by considering an adimensional time unit s such that

$$\vec{\lambda}(s) = \vec{\lambda}(s\tau)\,, \quad s\in(0,1)\,, \tag{1.56}$$

we can define, identifying from now on the adimensional time derivative $\vec{\lambda}' \equiv \partial\vec{\lambda}/\partial s = \tau \frac{d\vec{\lambda}}{dt}$,

$$A = \int_0^1 \mathrm{d}s \; \vec{R} \cdot \vec{\lambda}' \; , \tag{1.57}$$

$$\ell^2 = \int_0^1 \mathrm{d}s \; \vec{\lambda}' \cdot g \cdot \vec{\lambda}' \; , \tag{1.58}$$

$$\langle \kappa \rangle = \int_0^1 \mathrm{d}s \; \kappa \; . \tag{1.59}$$

Accordingly, Eqs. (1.54) and (1.55) can be expressed as follows,

$$W_{\mathrm{out}} = \frac{\Delta T}{T} A - \frac{\ell^2}{\tau} \tag{1.60}$$

$$Q_{\mathrm{tran}} = A + \frac{\Delta T}{T} \tau \langle \kappa \rangle . \tag{1.61}$$

The names A and ℓ^2 are related the geometrical meaning of the quantities above, as we discuss. Given that $\vec{\lambda}(s)$ represents a closed trajectory in space, we can use Stokes' theorem – in a three-dimensional space or its corresponding generalization in higher dimensions – to re-express the line-integral defining A

$$A = \int_{\partial\Sigma} \vec{R} \cdot \mathrm{d}\vec{\lambda} = \int_{\Sigma} (\vec{\nabla}_\lambda \wedge \vec{R}) \cdot \mathrm{d}\vec{\Sigma} \; , \tag{1.62}$$

where Σ is a surface in the $\vec{\lambda}$ space, having boundary $\partial\Sigma$ which coincides with the control trajectory. We can therefore *think of A as the area of the surface defined by the control trajectory* (with local weight depending on the $\vec{\nabla}_\lambda \wedge \vec{R}$ curvature). Note that this geometrical translation clarifies as well that A depends *only* on the geometry of the trajectory $\vec{\lambda}(s)$: that is, not only A is independent of τ, but it is also invariant under any reparametrization $s'(s)$ which might change the local speed and time spent on different points of the trajectory. Concerning ℓ^2, *it can be interpreted as a square length of the control trajectory* $\vec{\lambda}(s)$. At the same time, given the presence of two time derivatives, l^2 can depend in general on reparametrizations $s'(s)$. However, ℓ^2 represents losses due to dissipation in the driving – see Eq. (1.54) – and we are therefore interested in its minimum value, which can be obtained through the usual Cauchy-Schwarz inequality, or thermodynamic length inequality, used

already in (1.18)-(1.19)

$$\ell^2 \geq \left(\int_0^1 \mathrm{d}s\, \sqrt{\vec{\lambda}' \cdot g \cdot \vec{\lambda}'}\right)^2 = \left(\int_{\partial\Sigma} \sqrt{d\vec{\lambda} \cdot g \cdot d\vec{\lambda}}\right)^2 \equiv l^2 \,. \tag{1.63}$$

The lower bound l is fully geometric (it depends solely on $\partial\Sigma$) and it is always achievable by choosing the time-parametrization s' such that $\vec{\lambda}' \cdot g \cdot \vec{\lambda}'$ is constant. $\mathfrak{L}$ is a natural extension of the standard thermodynamic length [30, 31, 33, 35, 92] (which we used in the sections above and [1, 2]) to non-equilibrium setups where the working fluid is simultaneously interacting with several baths. Finally, it is apparent that $\langle\kappa\rangle$ Eq.(1.59) represents the simple average of a scalar number (the heat conductance) along the trajectory. In general it clearly also depends on reparametrizations of the adimensional time $s'(s)$, as the average can be arbitrarily close to the maximum (minimum) value $\kappa_{\max}$ ($\kappa_{\min}$) of the trajectory, in case s' is such to spend almost all the time close to $\kappa_{\max}$ ($\kappa_{\max}$).

Time-scaling Without diving into the details of the results of our work [6], we notice here two main differences with the standard low-dissipation regime discussed in 1.3. The first crucial difference is that the new loss mechanism, which is due to heat leaks, scales proportionally to the cycle duration $\propto \tau$, instead of $\frac{1}{\tau}$. This has the direct consequence that the ideal Carnot efficiency cannot be achieved by slowing down the operational times of control (contrary to standard Carnot-like machines), as the heat leak losses would grow indefinitely. The direct consequence is that the maximum efficiency of such machines is strictly smaller than Carnot's efficiency and is achieved at finite-times (cf. Fig. 1.5).
This can be seen, for example, in the *engine* operating mode of the system. The power of the heat engine and its efficiency can be written as

$$P = \frac{W_{\text{out}}}{\tau} = \frac{\Delta T}{T}\frac{A(1-\frac{\tau_D}{\tau})}{\tau} \,, \tag{1.64}$$

$$\eta = \frac{W_{\text{out}}}{Q_{\text{tran}}} = \eta_{\text{Carnot}}\frac{1-\frac{\tau_D}{\tau}}{1+\frac{\tau}{\tau_k}} \,, \tag{1.65}$$

where we substituted Eqs. (1.54)-(1.55) and we defined the dissipation and heat leak timescales

$$\tau_D = \frac{T}{\Delta T}\frac{\ell^2}{A} \,, \quad \tau_\kappa = \frac{T}{\Delta T}\frac{A}{\langle\kappa\rangle} \,. \tag{1.66}$$

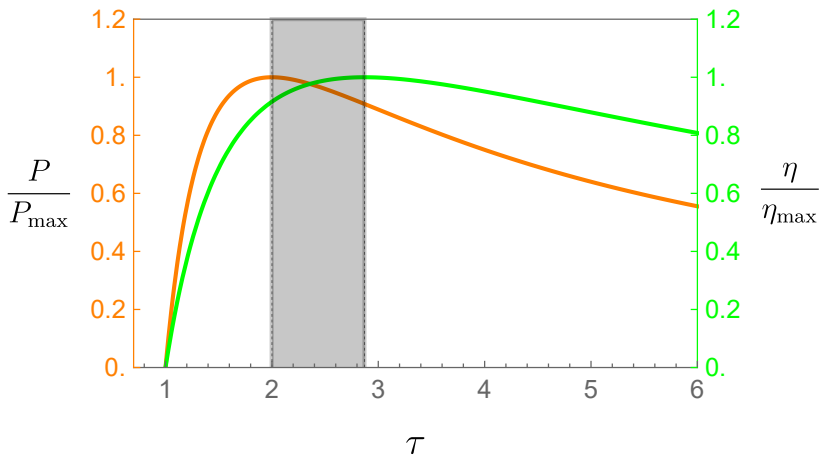

Figure 1.5: Engine mode for a low-dissipatoin engine with heat leaks. Power and efficiency are plotted as a function of the cycle duration, according to (1.64) and (1.65). The optimal operating region is the gray interval between the two dashed lines: indeed for any point outside the region, there is a point inside with both larger efficiency and larger power. In the plot $\tau_D = 1$ and $\tau_\kappa = 2.5$.

In the previous expressions $\eta_{\text{Carnot}} = \Delta T/T$ is the Carnot efficiency. Given the expressions above, we can optimize the duration of the cycles in order to maximize the power or the efficiency, obtaining correspondingly

$$\tau_P = 2\tau_D, \quad \tau_\eta = \tau_D + \sqrt{\tau_D(\tau_D + \tau_\kappa)}\,. \tag{1.67}$$

We see that the duration for maximum efficiency is always larger than the duration for maximum power. The corresponding maximum power and efficiency at maximum power are

$$P_{\max} = \frac{1}{4}\frac{(\Delta T)^2}{T^2}\frac{A^2}{\ell^2}\,, \quad \eta_{P_{\max}} = \frac{\eta_C}{2}\frac{x-1}{x+1} \tag{1.68}$$

while the maximum efficiency and power at maximum efficiency

$$\eta_{\max} = \eta_C\left(1 - \frac{2}{\sqrt{x}+1}\right)\,, \quad P_{\eta_{\max}} = \frac{(\Delta T)^2}{T^2}\langle\kappa\rangle\frac{(\sqrt{x}-1)^2}{\sqrt{x}}\,, \tag{1.69}$$

with

$$x = 1 + \frac{A^2}{\ell^2\langle\kappa\rangle}. \tag{1.70}$$

See Fig. 1.5 for a visual representation.

The isoperimetric problem. Finally we notice that the problem of finding such optimal protocols reduces to an *isoperimetric problem* [105] (also studied as *Cheeger Problem* [106, 107]), that is the task of finding the shape which maximizes the ratio between area and length. This is one of the oldest geometric problems in history, and was solved already by the ancient Greeks in the standard 2-dimensional Euclidean plane [108]. Nevertheless, when the underlying area density or length metrics are nontrivial, no general solution is known. The observation is straightforward, after the manipulation and time-optimization shown in the above section, all the power-efficiency related quantities in (1.68) and (1.69) can be seen as increasing functions of an Area/Perimeter ratio, A/ℓ. The problem of maximising such quantities then translates to maximising such ratio, in which Area and length are defined based on nontrivial metrics, as from (1.62) and (1.63).
In [6] we illustrated these ideas in a prominent quantum system playing the role of the working fluid: a qubit driven by two parameters slowly changing in time and asymmetrically coupled to two thermal reservoirs at different temperature. Such system can be used both in the engine mode and the refrigerator mode, which we optimized by numerically solving their relative isoperimetric problem on different classes of shapes.

1.5 Discussion

While originally developed for macroscopic systems, the geometric approach to finite-time thermodynamics is now finding renewed applications within the emerging fields of stochastic and quantum thermodynamics. In a series of papers [1, 2, 6], we have highlighted its utility for minimising dissipation in small scale systems operating close to equilibrium, characterizing the optimal solutions and showing how geometrical tools can greatly simplify the analysis and search of optimal driving protocols for such machines. Among our contributions, in [2] we summarised a set of key principles needed to optimise finite-time quantum low-dissipation engines in terms of efficiency and power, based on the computation of the thermodynamic metric tensor and length. Taken together, these principles provide a straightforward method for determining optimal thermodynamic processes. Indeed, it is also possible to fully optimize Carnot-like engines and refrigerators, finding the complete Pareto front, that is finding the maximum power at any given efficiency. A simple bound can be obtained characterizing such optimal performance [1] and can be saturated by performing small cycles close to the point where the ratio between heat capacity and thermalization timescale is the highest. Finally, in the case of dissipation arising not only from finite-time driving, but also heat leaks, we showed how exploiting a generalized thermodynamic tensor [36] it

is possible to solve the problem of performance optimization via a mapping to an isoperimetric problem.
Interesting future directions for thermodynamic geometry in the quantum regime include the extension beyond the slow driving regime (see also Chapter 2), the inclusion of fluctuations as an objective [90, 109, 110], which means solving a 3-dimensional Pareto front (see Sec. 1.1.3 and our work [23], which is not included in this thesis), connections with strong coupling and speed-ups to isothermality [111], relations with the third law of thermodynamics, many-body systems, criticality [39] and interaction engineering. Finally, experimental application of our techniques is in principle possible on different platforms, such as state of the art superconducting electronics [17, 37], NMR [69], single-electron quantum dots [112], but most of our optimization schemes can also be applied to larger systems, such as classical gases [113].

Chapter 2

Optimising fastly-driven engines

This chapter is based on Ref. [4]:
" V. Cavina, P. A. Erdman, P. Abiuso, L. Tolomeo and V. Giovannetti, *Maximum-power heat engines and refrigerators in the fast-driving regime*, Phys. Rev. A **104**, 032226 (2021)"

2.1 Background: power maximization and fast machines

In this Chapter we tackle once again the problem of thermal machines optimization. We consider generic driven thermal machines, and for a generic introduction to the framework we refer the reader to the background Section 1.1 of the previous Chapter.
The most important thermal machines that can be constructed utilizing two or more thermal baths are the heat engine and the refrigerator. These machines are mainly characterized by two figures of merit: the efficiency (or coefficient of performance for the refrigerator) and the extracted power (or cooling power). The optimal strategy to maximize the efficiency was identified already in the 19^{th} century, and is closely related to the second law of thermodynamics. As such it is characterized by a universal strategy: *infinitely slow* transformations, known as reversible transformations, must be performed (cf. Chap. 1). On the other hand, the maximization of the extracted power or cooling power requires finite-time thermodynamics, which relies on a microscopic model to describe the evolution of the system. Therefore, the maximization of the power is usually regarded as a model-specific task, thus lacking a universal characterization [25, 80, 114–116] In this framework (cf. Chap. 1), the fundamental question, which has not been tackled in general, is how to optimally drive the control parameters as to maximize the power of periodically

driven classical or quantum thermal machines. This is what we analyse in this Chapter.
In general, this is a formidable task, as it requires to solve the time-dependent dynamics of an open quantum system, coupled to thermal baths, and to perform a functional optimization over all available control parameters. In Chapter 1 we showed that in the low-dissipation, slow-driving regime, a universal strategy to maximize the power at any efficiency can be derived [1, 2]. However, notice that the dynamical approximations involved in the low-dissipation are justified – unsurprisingly – in the regime of high efficiency. That is, considering the Pareto front 1.1.3, in Chap. 1 we showed how to reconstruct it, under a general approximation that however can fail when focusing on the points of maximum power. Beyond this regime, common strategies to improve the power extracted from a quantum engine rely on performing fast and effectively adiabatic quantum operations through the Shortcut to Adiabaticity technique [117–119] or using Floquet engineering [120, 121]. The variety of frameworks employed span from the optimization of finite time Carnot cycles to Otto cycles [122, 123], endoreversible cycles, etc..
However within this *mare magnum* of frameworks and methods, in the context of systems described by Markovian dynamics, recent evidence suggests that the optimal strategy to extract maximum power may consist of varying the control parameters *infinitely fast* [12, 21, 24, 26, 124, 125]. This observation would imply a profound "duality" between efficiency and power: both would be maximized according to two opposite universal strategies (infinitely slow, or infinitely fast control speed). In [4] we discussed the optimization of thermal machines in the fast driving regime.

2.2 Framework and fast-driving regime

The *fast-driving regime* is characterized by driving time scales which are much faster than the thermal relaxation $\tau_{\rm rel}$. In [4] we introduce it and discuss it in the context of Markovian dynamics for systems whose Hamiltonian commutes at different times. This encompasses a variety of models of interest in stochastic thermodynamics, from chemical networks to molecular motors and more in general any dynamical system described by stochastic master equations [126].
As schematically depicted in Fig. 1.1, we consider a D-dimensional quantum system $\mathcal{S}$ (the working medium or working fluid of the model) that is weakly coupled to N thermal baths characterized by inverse temperatures β_α, for $\alpha = 1, \dots M$. As in Chapter 1, we assume $\mathcal{S}$ to be externally controlled through a set of n time-dependent control parameters collectively represented by a real

vector function

$$\vec{\lambda}_t \ , \quad 0 \leq t \leq \tau \ , \tag{2.1}$$

where τ is the total duration of the driving. In the following, we denote the function $\vec{\lambda}_t$ as the protocol or the driving. In our analysis $\vec{\lambda}_t$ acts as a modulator both for the local Hamiltonian of the system $H_{\vec{\lambda}_t}$ as well as for the interactions with the thermal baths which, adopting the standard Markovian approximations [51] (cf. Sec. 1.1), induce a first-order dynamics expressed by a Lindbladian super-operator $\mathcal{D}_{\alpha,\vec{\lambda}_t}$. We hence assign the temporal evolution of the system in terms of the following master equation for the reduced density matrix ρ_t of $\mathcal{S}$,

$$\dot{\rho}_t = \mathcal{L}_{\vec{\lambda}_t}[\rho_t] \equiv -i\left[H_{\vec{\lambda}_t}, \rho_t\right] + \sum_{\alpha=1}^{M} \mathcal{D}_{\alpha,\vec{\lambda}_t}[\rho_t] \, , \tag{2.2}$$

where $\mathcal{L}_{\vec{\lambda}_t}$ is the (time-dependent) Lindbladian generator of the dynamics. For simplicity, we assume that the Hamiltonian commutes at all times, i.e. that $[H_{\vec{\lambda}_1}, H_{\vec{\lambda}_2}] = 0$ for all $\vec{\lambda}_1, \vec{\lambda}_2$, in order to suppress non-adiabatic transitions. In such regime in fact, $\mathcal{D}_{\alpha,\vec{\lambda}_t}$ locally depends on time through $\vec{\lambda}_t$, and the master equation can be rigorously derived [1] (see details in [4] and references therein). We describe the possibility of deciding which bath is coupled to $\mathcal{S}$ at any given time through the dependence of the dissipators on $\vec{\lambda}_t$. If only bath α is coupled to $\mathcal{S}$, and if we fix the control parameters $\vec{\lambda}_t = \vec{\lambda}$, we expect $\mathcal{S}$ to thermalize by evolving towards the Gibbs density operator

$$\omega_{\alpha;\vec{\lambda}} \equiv \exp[-\beta_\alpha H_{\vec{\lambda}}]/Z_{\alpha;\vec{\lambda}} \, , \tag{2.3}$$

$Z_{\alpha;\vec{\lambda}} \equiv \mathrm{Tr}[\exp[-\beta_\alpha H_{\vec{\lambda}}]]$ being the partition function. The instantaneous heat flux flowing out of bath α can then be computed as [13, 80]

$$J_\alpha(t) \equiv \mathrm{Tr}\left[H_{\vec{\lambda}_t} \mathcal{D}_{\alpha,\vec{\lambda}_t}[\rho_t]\right] . \tag{2.4}$$

Again, within the above framework, we are interested in performing thermodynamic cycles, i.e. in performing a periodic driving $\vec{\lambda}_t$, with period τ, such that the variation of internal energy $U_t \equiv \mathrm{Tr}[H_{\vec{\lambda}_t}\rho_t]$ of the working fluid is zero after each cycle. In this regime, the first law of thermodynamics guarantees us that all the work extracted from the system is only provided by

[1] Moreover, in [4] we also include a discussion of the non-commuting case, the regimes in which Eq. 2.2 is still justified, and how our results can be affected by quantum rotations in the Hamiltonian control.

the heat baths, and not by some particular state preparation of $\mathcal{S}$. The periodicity of U_t requires both $\vec{\lambda}_t$ and ρ_t to be periodic functions. In general, ρ_t is not a periodic function. However, using the spectral properties of the Lindblad non-pathological master equations, the following property holds (a proof is provided, e.g., by Theorem 2 of Ref. [127]): if $\vec{\lambda}_t$ is a τ-periodic function, then the solution of Eq. (2.2) asymptotically converges toward a "limiting cycle" solution $\rho^{(\mathrm{lc})}_{[\vec{\lambda}]}(t)$, which is independent of the initial condition of the system, and which is periodic with the same period τ of the controls, i.e. $\rho^{(\mathrm{lc})}_{[\vec{\lambda}]}(t+\tau) = \rho^{(\mathrm{lc})}_{[\vec{\lambda}]}(t)$ for all τ. The subscript in $\rho^{(\mathrm{lc})}_{[\vec{\lambda}]}(t)$ emphasizes that the limiting cycle is a functional of the whole protocol, i.e. it depends on the control parameters along the whole cycle. In this asymptotic regime, the internal energy U_t becomes a periodic function, providing us with a thermodynamic cycle. From now on, we therefore focus solely on this regime.
We now wish to identify the optimal choice of $\vec{\lambda}_t$ that allows us to maximize the extracted power from a heat engine, or the cooling power of a refrigerator, averaged over a cycle. Both these quantities can be expressed as linear combinations of time integrals of the currents, defined in Eq. (2.4). Therefore, given an arbitrary collection c_α of real coefficients, we define the Generalized Average Power (GAP), which is a functional of the whole protocol, as

$$\begin{aligned} P_{\mathbf{c}}[\vec{\lambda}] \equiv & \frac{1}{\tau}\sum_{\alpha=1}^{M}\int_0^\tau c_\alpha J_\alpha(t)\,\mathrm{d}t, \\ = & \frac{1}{\tau}\int_0^\tau \mathrm{Tr}\left[H_{\vec{\lambda}_t}\sum_{\alpha=1}^{M} c_\alpha \mathcal{D}_{\alpha,\vec{\lambda}_t}\left[\rho^{(\mathrm{lc})}_{[\vec{\lambda}]_t}\right]\right]\mathrm{d}t. \end{aligned} \tag{2.5}$$

For instance, if we choose $c_\alpha = 1$ for all α, Eq. (2.5) represents the average of the total extracted heat flux, which coincides with the *average extracted power* for periodically driven heat engines; if instead $c_\alpha = \delta_{\alpha,M}$, with $\alpha = M$ labelling the coldest bath and δ representing the Kronecker delta, Eq. (2.5) represents the average *cooling power*, which measures the performance of a refrigerator; if $c_\alpha = -1$ for all α, Eq. (2.5) represents the average *dissipated heat flux*, which measures the performance of a *heater*, and so on.
Finding the optimal value of $\vec{\lambda}_t$ that maximize the functional (2.5) is, in general, a formidable task. Nonetheless, as we shall see, an explicit solution to the problem can be obtained when studying the performance in the limit of fast modulation of the driving parameters. In fact, until now, our framework is general and does not assume simplifications on the dynamics of the working fluid $\mathcal{S}$. The *fast driving regime* is characterized by driving the system with a protocol $\vec{\lambda}_t$ whose period τ is much shorter than the typical relaxation times

induced by the baths. Therefore, we may expect that the limiting cycle state of $\mathcal{S}$ "does not have time" to thermalize with the bath, so it might actually converge to a fixed, time-independent out-of-equilibrium state. This is precisely what happens.

More specifically, let us denote with $\Gamma_{[\vec{\lambda}]}$ the maximum rate which characterizes the master equation (2.6) along the cycle, that is the rate characterizing the fastest possible relaxation to the steady state (see technical details in [4]). Formally, we can expand $\rho^{(\mathrm{lc})}_{[\vec{\lambda}]}$ in a power series in $\Gamma_{[\vec{\lambda}]}\tau \ll 1$. As we prove in the technical material of [4], it turns out that the leading order term $\rho^{(0)}_{[\vec{\lambda}]}$ is indeed time-independent. A closed expression for such term can be obtained by making use of a projection technique that allows us to replace the dynamical generator $\mathcal{L}_{\vec{\lambda}_t}$ with the superoperator $\mathcal{G}_{\vec{\lambda}_t}$ which has the important property of being invertible on the (D^2-1)-dimensional linear subspace of traceless linear operators acting on $\mathcal{S}$. Specifically, Eq. (2.2) can be rewritten in the more convenient form

$$\dot{\tilde{\rho}}_t = \mathcal{G}_{\vec{\lambda}_t}\left[\tilde{\omega}_{\vec{\lambda}_t} - \tilde{\rho}_t\right], \tag{2.6}$$

where $\omega_{\vec{\lambda}_t}$ is the (unique) fixed point of $\mathcal{L}_{\vec{\lambda}_t}$ and where, for all density matrices ρ of $\mathcal{S}$, we define

$$\tilde{\rho} \equiv \rho - \mathbb{1}/d\,, \tag{2.7}$$

its traceless component. Equipped with this notation, in [4] we prove that

$$\tilde{\rho}^{(0)}_{[\vec{\lambda}]} \equiv \left(\int_{I_{[\vec{\lambda}]}} \mathcal{G}_{\vec{\lambda}_t}\, dt\right)^{-1}\left[\int_{I_{[\vec{\lambda}]}} \mathcal{G}_{\vec{\lambda}_t}[\tilde{\omega}_{\vec{\lambda}_t}]\, dt\right], \tag{2.8}$$

where $I_{[\vec{\lambda}]}$ denotes the time interval of one cycle of duration τ [2]. Using the approximation $\rho^{(\mathrm{lc})}_{[\vec{\lambda}]} \approx \rho^{(0)}_{[\vec{\lambda}]}$, we can write the GAP in Eq. (2.5) in the fast driving regime as

$$P_{\mathbf{c}}[\vec{\lambda}] = \frac{1}{\tau}\int_{I_{[\vec{\lambda}]}} \mathrm{Tr}\left[H_{\vec{\lambda}_t}\sum_\alpha c_\alpha \mathcal{D}_{\alpha,\vec{\lambda}_t}\left[\rho^{(0)}_{[\vec{\lambda}]}\right]\right] \mathrm{d}t, \tag{2.10}$$

[2] To understand non-rigorously the reason behind the limiting solution (2.8), notice that, in the limit of infinitely-fast driving, the steady state $\tilde{\rho}^{(0)}_{[\vec{\lambda}]}$ should satisfy the stability equation

$$\int \mathrm{d}t\ \dot{\tilde{\rho}}_t = 0 \quad \Rightarrow \quad \int \mathrm{d}t\ \mathcal{G}_{\vec{\lambda}_t}\left[\tilde{\omega}_{\vec{\lambda}_t} - \tilde{\rho}^{(0)}_{[\vec{\lambda}]}\right] = 0\,, \tag{2.9}$$

which implies (2.8) after inverting $\int_{I_{[\vec{\lambda}]}} \mathcal{G}_{\vec{\lambda}_t}\mathrm{d}t$.

which is guaranteed to be valid up to linear corrections in the expansion parameter $\Gamma_{[\vec{\lambda}]}\tau$ (however, it should be stressed that, by direct evaluation, the GAP of the optimal protocol turns out to be valid up to second order corrections in $\Gamma_{[\vec{\lambda}]}\tau$ in all the analytical and numerical applications we used).
Equations (2.8) and (2.10) are the main starting point of our analysis: they allow us to express the GAP as an explicit functional of the protocol $\vec{\lambda}_t$ without requiring us to solve the master equation.

2.3 Fast-driving optimization

Among all possible control strategies and protocols, in [4] we provided a universal proof that the power is optimized by "generalized Otto cycles", i.e. by performing sudden variations (quenches) of the control parameters among a finite number of fixed values. The generality of the proof is guaranteed by the fact that it holds for *any* Hamiltonian describing the working fluid, the baths, and the coupling. Furthermore, it holds regardless of the number of baths, and regardless of the specific form of the time dependent dissipators in the Lindblad master equation, that can depend on an arbitrary number of external controls subject to arbitrary constraints. In addition, it holds for the maximization of any linear combination of the heat currents, which includes the extracted power of a heat engine, the cooling power of a refrigerator, the dissipated heat by a heater, and so on.
The optimal protocol, i.e. the generalized Otto cycle, is characterized by L infinitesimal time intervals, connected by an identical number of quenches, in which the control parameters are held constant. We prove that, in general, $L \leq D$, where D is the dimension of the Hilbert space of the working fluid. This bound also places a constraint on the number of thermal baths that are necessary to maximize the power.

2.3.1 Divide et impera

Instead of performing a direct constrained functional optimization of the GAP (see Eq. (2.10)) with respect to $\vec{\lambda}_t$, we employ an iterative procedure that eventually leads to the identification of the "generalized Otto cycle" as the optimal one. The main idea of the proof is the following: given any assigned periodic protocol which respects the constraint $\vec{\lambda}_t$, we prove that it is possible to "cut away" parts of it to build a new, shorter, cycle which delivers a higher or equal GAP than the starting one. By reiterating this process over and over, we end up with the generalized Otto cycle. We therefore denote this procedure as *cut-and-choose*.

In order to detail the *cut-and-choose* procedure, let us first formally introduce the notion of cyclic sub-protocols. Given an arbitrary cyclic protocol $\vec{\lambda}_t$ of period τ and fundamental period $I_{[\vec{\lambda}]} = [0, \tau]$, consider a subset I_A of $I_{[\vec{\lambda}]}$ of non-zero measure τ_A. A cyclic sub-protocol $\vec{\lambda}_t^A$ of $\vec{\lambda}_t$ with period τ_A and fundamental period $I_{[\vec{\lambda}^A]} \equiv [0, \tau_A]$ is hence obtained by rigidly joining the various parts which compose the restriction of $\vec{\lambda}_t$ on I_A. This procedure may introduce localized discontinuities, i.e. quenches, within the protocol – see Fig. 2.1 for a 1-dimensional example.
Since Eq. (2.8) holds for any periodic protocol in the fast driving regime, by repeating $\vec{\lambda}_t^A$ many times, the state of $\mathcal{S}$ will tend to a new asymptotic constant state $\rho^{(0)}_{[\vec{\lambda}^A]}$ whose traceless component reads accordingly

$$\tilde{\rho}^{(0)}_{[\vec{\lambda}^A]} = \left(\int_{I_{[\vec{\lambda}_A]}} \mathcal{G}_{\vec{\lambda}_t^A} \mathrm{d}t \right)^{-1} \left[\int_{I_{[\vec{\lambda}_A]}} \mathcal{G}_{\vec{\lambda}_t^A} [\tilde{\omega}_{\vec{\lambda}_t^A}] \, \mathrm{d}t \right] . \tag{2.11}$$

It goes without mentioning that analogous conclusions can be drawn also for the sub-protocol $\vec{\lambda}_t^B$ that is obtained by considering the restriction of $\vec{\lambda}_t$ to the complement I_B of I_A, i.e. the set $I_B = I_{[\vec{\lambda}]}/I_A$ of measure $\tau_B = \tau - \tau_A$: once more, under iterated application of such driving, the state of S will tend to a constant asymptotic state $\rho^{(0)}_{[\vec{\lambda}^B]}$ given by Eq. (2.11) by simply replacing everywhere the index A with B; see Fig. 2.1 for an example.
Assume next that the states $\rho^{(0)}_{[\vec{\lambda}^A]}$ and $\rho^{(0)}_{[\vec{\lambda}^B]}$ introduced above coincide and are equal to $\rho^{(0)}_{[\vec{\lambda}]}$, i.e.

$$\rho^{(0)}_{[\vec{\lambda}^A]} = \rho^{(0)}_{[\vec{\lambda}^B]} = \rho^{(0)}_{[\vec{\lambda}]} \, . \tag{2.12}$$

Equation (2.12) is a rather strong requirement which in general is not met by generic choices of I_A and I_B: still, as we shall discuss in the next section, the possibility of identifying sub-protocols fulfilling this property is always granted. For the moment we hence assume that Eq. (2.12) is satisfied. The GAPs $P_{\mathbf{c}}[\vec{\lambda}^A]$ and $P_{\mathbf{c}}[\vec{\lambda}^B]$ delivered respectively by the sub-protocols $\vec{\lambda}^A$ and $\vec{\lambda}^B$ can be computed using Eq. (2.10). Assuming Eq. (2.12) is fulfilled, we notice that the integrands entering $P_{\mathbf{c}}[\vec{\lambda}]$, $P_{\mathbf{c}}[\vec{\lambda}^A]$ and $P_{\mathbf{c}}[\vec{\lambda}^B]$ are all the same. Therefore, exploiting the linearity of the integral respect to the its integration domain (i.e. time), and recalling that $\tau = \tau_A + \tau_B$, we have that

$$P_{\mathbf{c}}[\vec{\lambda}] = \frac{\tau_A P_{\mathbf{c}}[\vec{\lambda}^A] + \tau_B P_{\mathbf{c}}[\vec{\lambda}^B]}{\tau_A + \tau_B}. \tag{2.13}$$

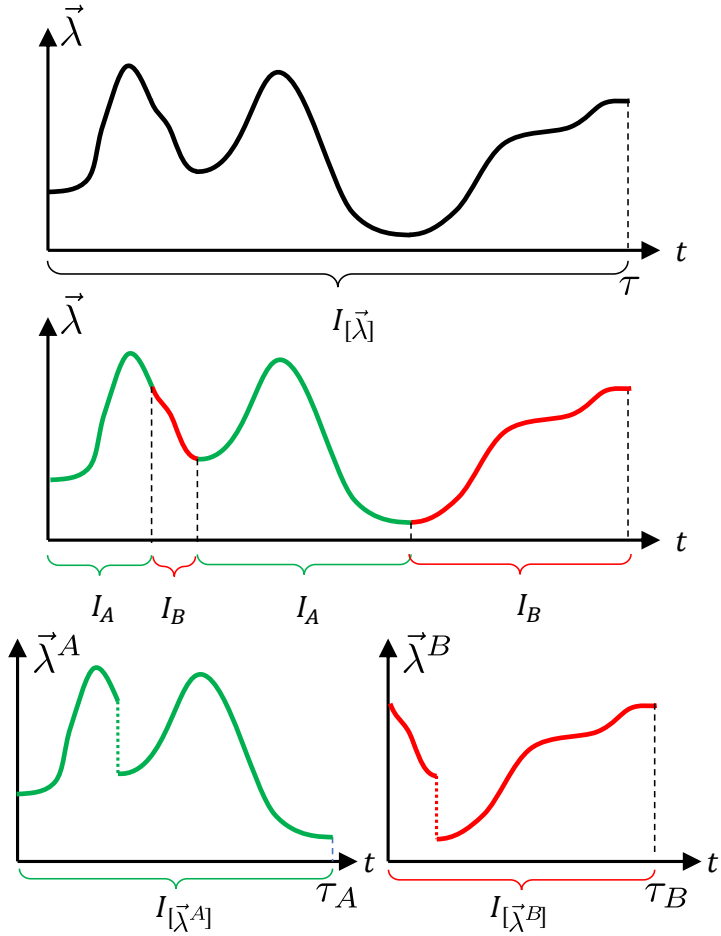

Figure 2.1: Schematic representation of the *cut-and-choose* procedure. Upper panel: representation of an arbitrary protocol $\vec{\lambda}_t$ defined on the time interval $I_{[\vec{\lambda}]}$ of duration τ. Central panel: we partition $I_{[\vec{\lambda}]}$ into two disjoing subsets I_A and I_B. Lowe panel: we define two new sub-protocols $\vec{\lambda}_t^A$ and $\vec{\lambda}_t^B$ by restricting $\vec{\lambda}_t$ respectively to I_A and I_B. This process may introduce discontinuities in the controls, denoted as quenches.

The above equation establishes that the GAP of the original protocol $\vec{\lambda}_t$ can be expressed as a non-trivial convex combination of the GAPs of the sub-protocols $\vec{\lambda}_t^A$ and $\vec{\lambda}_t^B$: therefore it must be smaller or equal to the maximum of those two quantities, i.e.

$$P_{\mathbf{c}}[\vec{\lambda}] \leq P_{\mathbf{c}}[\vec{\lambda}^A], \tag{2.14}$$

where, without loss of generality we assumed $P_{\mathbf{c}}[\vec{\lambda}^B] \leq P_{\mathbf{c}}[\vec{\lambda}^A]$. Inequality

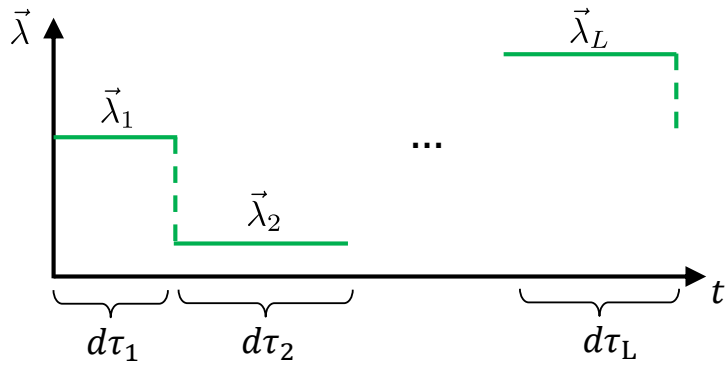

Figure 2.2: Representation of a generalized Otto cycle, which results from a large number of re-iterations of the *cut-and-choose* procedure depicted in Fig. 2.1. L (infinitesimally short) constant controls are alternated with quenches. Finite upper bounds can be obtained for the maximum value of L that is needed for the optimization.

(2.14) implies that given a generic periodic protocol $\vec{\lambda}_t$, it is possible to construct a *shorter* one $\vec{\lambda}^A$ that delivers a larger or equal GAP. This is the reason for the name *cut-and-choose* procedure. We can now re-iterate the *cut-and-choose* procedure starting from $\vec{\lambda}^A_t$, thus obtaining another (even shorter) protocol $\vec{\lambda}^{AA}_t$ that produces a greater or equal GAP, and so on and so forth. After many iterations of the *cut-and-choose* procedure, we end up with a protocol that cannot be further optimized via this technique. This protocol is characterized by an infinitesimal domain I of duration $d\tau$, divided into L segments of length $d\tau_i$. Without loss of generality, we can assume that the $d\tau_i$s are short enough such that the controls $\vec{\lambda}_t$ take on a constant value $\vec{\lambda}_i \equiv \vec{\lambda}_{t_i}$ during each time interval $d\tau_i$. This is a generalized Otto cycle; see Fig. 2.2 for a schematic representation.
Using Eq. (2.10), the associated GAP of such protocol can hence be expressed as

$$P_{\mathfrak{c}}[\{\vec{\lambda}_i, \mu_i\}] = \sum_{j=1}^{L} \mu_j \mathrm{Tr}\left[H_{\vec{\lambda}_j} \sum_{\alpha} c_\alpha \mathcal{D}_{\alpha,\vec{\lambda}_j}\left[\rho^{(0)}_{[\{\vec{\lambda}_i,\mu_i\}]}\right]\right], \tag{2.15}$$

where $\mu_i = d\tau_i/d\tau$ represents the percentage of the total protocol time spent at each point $\vec{\lambda}_i$, and $\rho^{(0)}_{[\{\vec{\lambda}_i,\mu_i\}]}$ is the time-independent limiting cyclic state whose traceless component is, from Eq. (2.8),

$$\tilde{\rho}^{(0)}_{[\{\vec{\lambda}_i,\mu_i\}]} \equiv \left(\sum_{j=1}^{L} \mu_j \mathcal{G}_{\vec{\lambda}_j}\right)^{-1} \left[\sum_{j=1}^{L} \mu_j \mathcal{G}_{\vec{\lambda}_j}[\tilde{\omega}_{\vec{\lambda}_j}]\right] . \tag{2.16}$$

Crucially, we are able to place a finite upper bound to the number L of time intervals of the optimal generalized Otto cycle. In [4] we prove that, to maximize the GAP in general, it is sufficient to consider L to be at most equal to the degrees of freedom of the density matrix plus one. Since in the commuting case only the diagonal component of ρ_t plays a role in determining the heat currents, we find that

$$L \leq D \, . \tag{2.17}$$

Furthermore, if the dissipator of each bath is characterized by a single (control-dependent) timescale and $N = 2$, then $L = 2$ regardless of the dimensionality of the system; in this case, the optimal protocol reduces to a conventional infinitesimal Otto cycle. For applications, we can thus directly optimize Eq. (2.15) over the values of the controls $\vec{\lambda}_i$ and of the time fractions μ_i which are model-specific.

In order to gain further physical insight into our result, let us consider the paradigmatic case in which our system $\mathcal{S}$ can only be coupled to one bath at the time. Mathematically, this assumption can be described by a specific control parameter, say α_t, whose value is the index of the bath we are coupled to, $\alpha = 1, \ldots, M$. Therefore, $\mathcal{S}$ must be coupled only to a single bath in each time interval $d\tau_i$. In this scenario, it is interesting to notice that our bound on the number of time intervals poses a limit to the maximum number of thermal baths necessary to maximize the GAP: indeed, at most L baths will be used. Therefore, for low dimensional working fluids, the maximum number of thermal baths necessary to maximize the GAP is strongly limited. However, we also explicitly show in [4] that three thermal baths at different temperatures can outperform two thermal baths when the working fluid is a qutrit. This result is in contrast with the maximization of the efficiency, which is always obtained by coupling $\mathcal{S}$ only to the hottest and coldest bath available.

2.3.2 A geometric interpretation of Eq. (2.12)

The argument presented in the previous section relies on the assumption (2.12) that one can identify two new sub-protocols $\vec{\lambda}_t^A$ and $\vec{\lambda}_t^B$ that preserve the asymptotic state $\rho^{(0)}_{[\vec{\lambda}]}$ of the original protocol $\vec{\lambda}(t)$. We provide an explicit proof that such condition can always be fulfilled by translating it into a geometric problem.

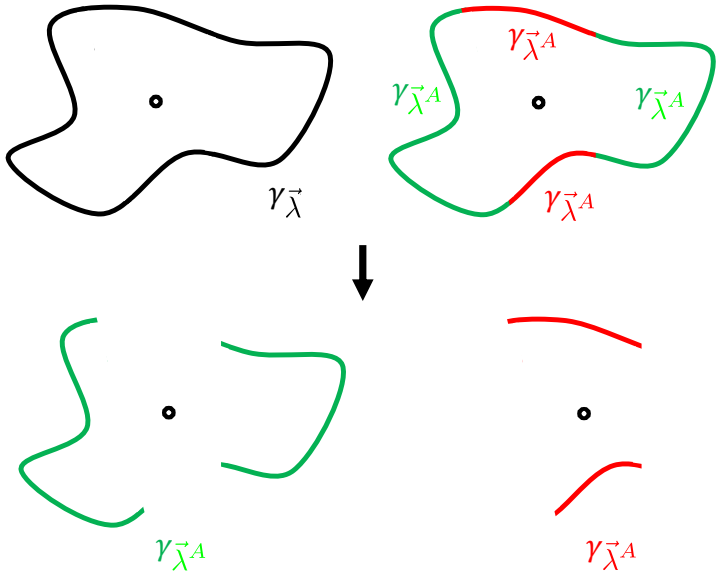

Figure 2.3: Upper left: schematic representation of $\gamma_{[\vec{\lambda}]}$ and of its center of mass. Upper right: schematic representation of a partition of $\gamma_{[\vec{\lambda}]}$ into $\gamma_{[\vec{\lambda}^A]}$ and $\gamma_{[\vec{\lambda}^B]}$. Lower panels: schematic representation of the sub-curves $\gamma_{[\vec{\lambda}^A]}$ and $\gamma_{[\vec{\lambda}^B]}$ which preserve the center of mass of the original curve $\gamma_{[\vec{\lambda}]}$.

For this purpose, let us define the curves $\gamma_{[\vec{\lambda}]} \equiv \{v_{\vec{\lambda}(t)} | t \in I_{[\vec{\lambda}]}\}$, $\gamma_{[\vec{\lambda}^A]} \equiv \{v_{\vec{\lambda}^A_t} | t \in I_{[\vec{\lambda}^A]}\}$, and $\gamma_{[\vec{\lambda}^B]} \equiv \{v_{\vec{\lambda}^B_t} | t \in I_{[\vec{\lambda}^B]}\}$ generated by the functions

$$v_{\vec{\lambda}_t} \equiv \mathcal{G}_{\vec{\lambda}_t}\left[\tilde{\omega}_{\vec{\lambda}(t)} - \tilde{\rho}^{(0)}_{[\vec{\lambda}]}\right] , \tag{2.18}$$

$$v_{\vec{\lambda}^{A,B}_t} \equiv \mathcal{G}_{\vec{\lambda}^{A,B}_t}\left[\tilde{\omega}_{\vec{\lambda}^{A,B}_t} - \tilde{\rho}^{(0)}_{[\vec{\lambda}]}\right] . \tag{2.19}$$

That is, the above simply represent the trajectories given by the instantaneous time-derivatives (velocities $v_{\vec{\lambda}_t}$) associated to $\tilde{\rho}^{(0)}_{[\vec{\lambda}_t]}$. Since the domains $I_{[\vec{\lambda}^A]}$ and $I_{[\vec{\lambda}^B]}$ are complementary and provide a decomposition of $I_{[\vec{\lambda}]}$, $\gamma_{[\vec{\lambda}^A]}$ and $\gamma_{[\vec{\lambda}^B]}$ are disjoint, and their union coincides with $\gamma_{[\vec{\lambda}]}$ (see upper panels of Fig. 2.3 for a schematic representation). It is important to notice that the functions in Eq. (2.19), thus also the curves $\gamma_{[\vec{\lambda}]}$, $\gamma_{[\vec{\lambda}^A]}$, and $\gamma_{[\vec{\lambda}^B]}$, belong to the special subspace of traceless Hermitian operators of $\mathcal{S}$ which is locally isomorphic to $\mathbb{R}^{D^2-1}$.By exploiting the fact that the Hamiltonian commutes at all times, we can further reduce the number of degrees of freedom to be locally isomorphic to $\mathbb{R}^{D-1}$. This is due to the fact that the heat currents can be written solely in terms of the diagonal part of ρ_t, which in turn satisfies a closed equation of

motion.
Since $\tilde{\rho}^{(0)}_{[\vec{\lambda}]}$ satisfies Eq. (2.8), the curve $\gamma_{[\vec{\lambda}]}$ has a null "center of mas" $O_{[\vec{\lambda}]}$ (represented by the black dot in Fig. 2.3), i.e.

$$O_{[\vec{\lambda}]} \equiv \int_{I_{[\vec{\lambda}]}} v_{\vec{\lambda}(t)}\, \mathrm{d}t = 0\ . \tag{2.20}$$

Using the linearity of the integral respect to its integration domain, it is easy to verify that the sum of the "centers of mass" $O_{[\vec{\lambda}^A]} \equiv \int_{I_{[\vec{\lambda}^A]}} v_{\vec{\lambda}^A_t}\, \mathrm{d}t$ with $O_{[\vec{\lambda}^B]} \equiv \int_{I_{[\vec{\lambda}^B]}} v_{\vec{\lambda}^B_t}\, \mathrm{d}t$ is null, i.e.

$$O_{[\vec{\lambda}^A]} + O_{[\vec{\lambda}^B]} = O_{[\vec{\lambda}]} = 0\ . \tag{2.21}$$

We claim that a necessary and sufficient condition for Eq. (2.12) to hold is that the curve $\gamma_{[\vec{\lambda}^A]}$ (and hence due to Eq. (2.21), also $\gamma_{[\vec{\lambda}^B]}$) must have a null center of mass too. Indeed, exploiting the invertibility of $\int_{I_{[\vec{\lambda}^A]}} \mathcal{G}_{\vec{\lambda}^A_t} \mathrm{d}t$, one can observe that setting $O_{[\vec{\lambda}^A]} = 0$ is fully equivalent to having

$$\tilde{\rho}^{(0)}_{[\vec{\lambda}]} = \left(\int_{I_{[\vec{\lambda}^A]}} \mathcal{G}_{\vec{\lambda}^A_t} \mathrm{d}t \right)^{-1} \left[\int_{I_{[\vec{\lambda}^A]}} \mathcal{G}_{\vec{\lambda}^A_t} [\tilde{\omega}_{\vec{\lambda}^A_t}]\, \mathrm{d}t \right] = \tilde{\rho}^{(0)}_{[\vec{\lambda}^A]}, \tag{2.22}$$

where, in the last step, we used Eq. (2.11) to recognize $\tilde{\rho}^{(0)}_{[\vec{\lambda}^A]}$. An analogous conclusion holds also for the sub-protocol $\vec{\lambda}^B_t$ thanks to Eq. (2.21).
This is the geometric reformulation of Eq. (2.12) we were looking for: our partitioning technique works if, starting from a generic curve $\gamma_{[\vec{\lambda}]}$ in $\mathbb{R}^{D-1}$ having a null center of mass, we are able to split it into two sub-curves $\gamma_{[\vec{\lambda}^A]}$ and $\gamma_{[\vec{\lambda}^B]}$ such that these still have a null center of mass (this concept is schematically represented in Fig. 2.3). In [4] we prove that it is indeed possible assuming that the original protocol $\vec{\lambda}_t$ possesses some weak notion of regularity. The main idea is that, given an arbitrary curve in $\mathbb{R}^{D-1}$ with zero center of mass, it is always possible to identify a null convex combination of at most D points lying on the curve.
Reiterating this *cut-and-choose* procedure many times may lead to a piecewise continuous curves with a large number of discontinuities. Crucially, in [4] we show that it is always possible to end up with a curve characterized by at most D discontinuities. This result gives rise to the bounds on L.

2.4 Applications

In Ref. [4], we applied the main results presented above to specific scenarios. For example, when all observables of the working fluid share the same (control-dependent) thermalization time, we further proved that the number of quenches needed in an optimal protocol is $L = 2$, that is, the optimal protocol is a standard infinitesimal Otto cycle.
Moreover, in such models, assuming to have total control over the Hamiltonian of the working fluid, we identify the optimal modulation of the control parameters, which consists of producing a highly-degenerate many-body spectrum characterized by a single energy gap. This protocol allows us to compute the maximum achievable power using a working medium made up of N interacting qubits. We showed that the power of such heat engine goes beyond its counterpart based on N non-interacting qubits, displaying a many-body advantage. The value of the maximum power has a supra-extensive transient regime in N, and in the $N \to \infty$ limit we find that it is linear in the temperature difference ΔT between the hottest and the coldest bath, while the non interacting case exhibits the more common quadratic ΔT^2 dependence. In addition, the interacting case displays an efficiency at maximum power which asymptotically approaches Carnot efficiency (for $N \to \infty$). Surprisingly, we find that in the refrigerator case, many-body interactions do not provide significant advantage over non-interacting qubits.
Additionally, we studied a qutrit system as a testing ground for our general results. We numerically show that while the common $L = 2$ case is optimal for typical thermalization models used to describe Bosonic and Fermionic baths, the generalized Otto cycle (characterized in this case by $L = 3$ quenches) outperforms the $L = 2$ case for some particular forms of the master equation. This implies that our bound on L (2.17) is, in general, tight. Furthermore, as opposed to the maximum efficiency, we show that the power can be enhanced by the presence of more than two thermal baths at different temperatures.
All the detailed results and simulations can be found in [4]. In the following, we briefly present the relevant case of systems that have a (point-dependent) single relaxation timescale.

2.4.1 Simple thermalization case

In this section we discuss a simplified model of thermalization where the super-operator $\mathcal{G}_{\vec{\lambda}_t}$ of Eq. (2.6) is purely multiplicative, leading to a master equation of the form

$$\dot{\tilde{\rho}}_t = \Gamma_{\vec{\lambda}_t}(\tilde{\omega}_{\vec{\lambda}_t} - \tilde{\rho}_t), \tag{2.23}$$

with $\Gamma_{\vec{\lambda}_t} > 0$ a scalar number which defines the rate of thermalization of all the observables of the system. Furthermore, we assume that the model allows $\mathcal{S}$ to be coupled to a single bath at the time. As discussed in the final part of Sec. 2.3.1, we formally introduce a single control parameter, denoted with α_t, indexing the bath we are coupled to at time t. Notice that, for all values of $\vec{\lambda}_t$, the equilibrium states $\omega_{\vec{\lambda}_t}$ always correspond to the Gibbs distribution of bath α_t, i.e. $\omega_{\alpha_t;\vec{\lambda}_t}$ as in Eq. (2.3). As discussed in the first part of the paper, the maximum GAP is given by Eqs. (2.15) and (2.16) which, using Eq. (2.23), can be rewritten as

$$P_{\mathbf{c}}[\{\vec{\lambda}_i, \mu_i\}] = \frac{\sum_{i,j=1}^{L} c_{\alpha_i} \mu_i \mu_j \Gamma_{\vec{\lambda}_i} \Gamma_{\vec{\lambda}_j} P_{i\leftarrow j}}{\sum_{i=1}^{L} \mu_i \Gamma_{\vec{\lambda}_i}} , \tag{2.24}$$

$$\tilde{\rho}^{(0)}_{[\{\vec{\lambda}_i,\mu_i\}]} = \frac{\sum_{i=1}^{L} \mu_i \Gamma_{\vec{\lambda}_i} \tilde{\omega}_{\vec{\lambda}_i}}{\sum_{i=1}^{L} \mu_i \Gamma_{\vec{\lambda}_i}} , \tag{2.25}$$

where α_i is the constant value of $\alpha(t)$ during the interval $d\tau_i$ (we used the assumption that S can be coupled to a single bath at the time to remove the sum over α in Eq. (2.24)), and where

$$P_{i\leftarrow j} \equiv \mathrm{Tr}\left[H_{\vec{\lambda}_i}\left(\tilde{\omega}_{\vec{\lambda}_i} - \tilde{\omega}_{\vec{\lambda}_j}\right)\right] . \tag{2.26}$$

Notice that $P_{i\leftarrow i} = 0$, while

$$P_{j\leftarrow k} + P_{k\leftarrow j} \leq 0 , \tag{2.27}$$

when $\mathcal{S}$ is coupled to the same temperature during the time intervals $d\tau_j$ and $d\tau_k$. This is given by the fact that $P_{j\leftarrow k} + P_{k\leftarrow j}$ is equal to $P_{\mathbf{c}}[\{\vec{\lambda}_i, \mu_i\}]$ with $c_\alpha = 1\ \forall i$, and $\mu_i = 0\ \forall i \neq j, k$, which physically represents the average power extracted from a heat engine operating between equal temperatures (and therefore cannot be positive).

Using these properties, we show that with the only assumption of the dynamics being described by Eq. (2.23), it is possible to greatly simplify the optimization of the GAP of thermal machines. We consider *positive* GAPs, i.e. generalized average powers consisting of a positive linear combination of the heat currents extracted from the different thermal baths (formally we assume that $c_\alpha \geq 0\ \forall\alpha$). This hypothesis includes both the average power extracted from a heat engine ($c_\alpha = 1\ \forall\alpha$), and the cooling power of a refrigerator ($c_\alpha = \delta_{\alpha,N}$, with $\alpha = N$ labelling the coldest bath). We prove that [4], in order to maximize a positive GAP, it is sufficient to consider a protocol with at most one time interval per temperature; therefore $L \leq N$. Moreover, if more

than one heat current is neglected in the definition of the GAP, it is possible to further reduce the number of intervals. Specifically, given $\kappa \leq N$ the number of distinct temperatures of the baths for which $c_\alpha \neq 0$, we prove that

$$L \leq \min(N, \kappa + 1) \, . \tag{2.28}$$

This implies that a refrigerator ($\kappa = 1$) is always characterized by $L = 2$, regardless of the number of baths, while a heat engine ($\kappa = N$) by $L \leq N$. In the following, for simplicity, we focus on the refrigerator and heat engine case with two thermal baths at our disposal. As a consequence, $L = 2$. Under this hypothesis, we find that:

(i) The optimal durations of the time intervals and the resulting maximum GAPs can be determined as a function of the control parameters.

(ii) If the thermalization rates are a function of the bath α_t, but only weakly depend on the specific value of the other control parameters, i.e. $\Gamma_{\vec{\lambda}_i} = \Gamma_{\alpha_i}$, and if we assume to have total control over the system Hamiltonian, we can fully carry out the maximization of the GAP, finding that the optimal control strategies involve degenerate spectra of the Hamiltonian of the working fluid.

Moreover, under the toy model hypothesis of (ii), we compare the GAP of a heat engine and of a refrigerator delivered by N non-interacting qubits [$\mathrm{GAP}_{\mathrm{NI}}(N)$], with the GAP of N interacting qubits [$\mathrm{GAP}_{\mathrm{I}}(N)$], finding that there is a many-body advantage in the engine case (only!). All the details of these results can be found in [4]. For completeness, we provide the analytical expressions of the maximum power of a standard refrigerator and an engine in the following.

Refrigerator Let us consider two inverse temperatures β_1 and β_2 such that $\beta_1 < \beta_2$. The average cooling power of a refrigerator, $P_{[\mathrm{R}]}$, is described by the GAP with $c_2 = 1$ on the cold bath while $c_1 = 0$. Since $L = 2$, Eq. (2.24) reduces to

$$P_{[\mathrm{R}]} = \frac{\mu_1 \mu_2 \Gamma_{\vec{\lambda}_1} \Gamma_{\vec{\lambda}_2} P_{2\leftarrow 1}}{\mu_1 \Gamma_{\vec{\lambda}_1} + \mu_2 \Gamma_{\vec{\lambda}_2}} \tag{2.29}$$

where $\mu_2 = 1 - \mu_1$. We can thus explicitly maximize the above expression over the choice of the time fraction μ_1, leading to

$$P_{[\mathrm{R}]}^{(\max)} = \frac{\mathrm{Tr}\left[H_{\vec{\lambda}_2}\left(\tilde{\omega}_{\vec{\lambda}_2} - \tilde{\omega}_{\vec{\lambda}_1}\right)\right]}{\left(\sqrt{\Gamma_{\vec{\lambda}_1}^{-1}} + \sqrt{\Gamma_{\vec{\lambda}_2}^{-1}}\right)^2}, \tag{2.30}$$

which is obtained for $\mu_1 = \sqrt{\Gamma_{\vec{\lambda}_2}}/(\sqrt{\Gamma_{\vec{\lambda}_2}} + \sqrt{\Gamma_{\vec{\lambda}_1}})$. Notably, the expression of the maximum cooling power in Eq. (2.30) only requires a maximization over $\vec{\lambda}_1$ and $\vec{\lambda}_2$, which in general is model dependent.

Engine Let us consider the same setting $\beta_1 < \beta_2$. The average extracted power of a heat engine, $P_{[\mathrm{E}]}$, is described by the GAP with $c_1 = c_2 = 1$. Since $L = 2$, Eq. (2.24) reduces to

$$P_{[\mathrm{E}]} = \frac{\mu_1\mu_2\Gamma_{\vec{\lambda}_1}\Gamma_{\vec{\lambda}_2}\left(P_{1\leftarrow 2} + P_{2\leftarrow 1}\right)}{\mu_1\Gamma_{\vec{\lambda}_1} + \mu_2\Gamma_{\vec{\lambda}_2}} \tag{2.31}$$

It follows that the optimization over the time fraction μ_1 is identical to that of the refrigerator, see Eq. (2.29), leading to

$$P_{[\mathrm{E}]}^{(\max)} = \frac{\mathrm{Tr}\left[\left(H_{\vec{\lambda}_1} - H_{\vec{\lambda}_2}\right)\left(\tilde{\omega}_{\vec{\lambda}_1} - \tilde{\omega}_{\vec{\lambda}_2}\right)\right]}{\left(\sqrt{\Gamma_{\vec{\lambda}_1}^{-1}} + \sqrt{\Gamma_{\vec{\lambda}_2}^{-1}}\right)^2}, \tag{2.32}$$

which is obtained for $\mu_1 = \sqrt{\Gamma_{\vec{\lambda}_2}}/(\sqrt{\Gamma_{\vec{\lambda}_2}}+\sqrt{\Gamma_{\vec{\lambda}_1}})$. Also in this case, Eq. (2.32) only requires a model-dependent maximization over $\vec{\lambda}_1$ and $\vec{\lambda}_2$.

2.5 Discussion

In Ref. [4] we exhaustively discussed the optimization of thermal machines in the fast driving regime for commuting Hamiltonians. We proved in full generality that the optimal protocols are universally given by generalized Otto cycles, which are composed by a certain number L of infinitesimal time intervals where the control is fixed. We then bounded L from above in terms of the dimension of the Hilbert space of the working fluid. The proof holds regardless of the specific choice of the control-dependent dissipators, of possible constraints on the control parameters, and regardless of the specific form of the Hamiltonian of the working fluid. Moreover, the bound is shown to be tight in general. We showed that the standard fast Otto cycle, characterized by two quenches ($L = 2$) is optimal in a vast class of systems. Assuming full control over the system, we explicitly found the optimal driving strategy, which involves producing highly degenerate states, revealing an interesting connection with the results of [122] and our [1], as well as being optimal for thermometry (see Chap. 3). In [4] we also applied this optimal strategy to compare the performance of a refrigerator and of a heat engine based on N interacting and non-interacting qubits. In the refrigerator case, we found that

the non-interacting qubits perform almost as well as the interacting ones; it is therefore reasonable to consider constructing a refrigerator operating in parallel many simple independent units. Conversely, in the heat engine case we found a many-body advantage resulting in the enhancement of both the maximum power, and of the efficiency at maximum power, which approaches Carnot efficiency in the limit of many qubits.
Besides their theoretical relevance, from an operational point of view, the results derived in this work hugely simplify the numerical procedure of finding optimal protocols, due to the intrinsic simplicity of the generalized Otto cycle. This is somewhat analogous to what happens to control optimizations in the *slow driving* regime (cf. Chap. 1), in which the driving is much slower than the dissipative dynamics induced by the baths. These results show that, by exploiting the concept of time scale separation, we can simplify the characterization of the power generation in thermal machines. This simplification can be exploited both for analytical and numerical treatments, such as the ones showed in [4] and more. As future directions, it is interesting to assess the role of coherence in the non-commuting case, and to understand for which classes of systems the fast driving regime is optimal for power extraction. Furthermore, by providing strict bounds on optimal protocols, our results can be used as benchmarks to assess if effects beyond the Markovian regime and weak coupling approximation can indeed enhance, or decrease, the performance of thermal machines (see for example our work [9], not included in this thesis). At last, it seems natural to investigate the properties of the fast-driving regime respect to other thermodynamic figures of merit, such as the efficiency at maximum power, or work fluctuations. We partially developed this goal in the technical material of Ref. [12], which is not included in this thesis.

Chapter 3

Engineering good thermometers

This chapter is based on Ref. [8]:
" P. Abiuso, P.A. Erdman, M. Ronen, G. Haack, F. Noé and M. Perarnau-Llobet, *Discovery of Optimal Thermometers with Spin Networks aided by Machine-Learning*, arXiv preprint 2211.01934, (2022)."

3.1 Background: equilibrium thermometry

A fundamental operational task in thermodynamics is that of measuring temperature, in order to correctly assess a system at thermal equilibrium. Such problem enters in the general theoretical study of parameter estimation, which gave rise to the modern field of quantum metrology [128, 129]. The typical task in metrology is that of estimating one or more parameters $\vec{\theta}$ which are encoded in the state of a *probe* $\rho_{\vec{\theta}}$, via a given measurement scheme. It turns out that this problem is intimately related to the geometry of the set of states itself [130], via the so called *Fisher information* metric (cf. Chap. 4 and App. C). Focusing on the estimation of temperature, we define the *thermal sensitivity* of a system at thermal equilibrium, as the variation in the mean energy of such a system per temperature change unit. Such a figure of merit is also termed *heat capacity* of the system. More precisely, given a bath at some temperature $T = \beta^{-1}$ (we use natural units in which $k_B = 1$), we set about by letting a probe (the thermometer) ρ, fully thermalize with a thermal reservoir, resulting in the Gibbs state for the probe

$$\rho \longrightarrow \omega_\beta(H) := \frac{e^{-\beta H}}{\mathrm{Tr}[e^{-\beta H}]} \,. \tag{3.1}$$

The heat capacity is then given by

$$\mathcal{C}(H,\beta) := \frac{\mathrm{d}}{\mathrm{d}T}\mathrm{Tr}[H\rho_\beta(H)] = -\beta^2\frac{\mathrm{d}}{\mathrm{d}\beta}\mathrm{Tr}[H\rho_\beta(H)]\ . \tag{3.2}$$

Besides its relevance in temperature estimation (see below), the heat capacity $\mathcal{C}$ is a fundamental thermodynamic property of physical systems: it has been proven proportional to the maximum power a system can output in Carnot-like engines [1, 39], and its critical behaviour is studied in phase transitions.
In the context of thermometry, $\mathcal{C}$ quantifies the maximum precision one can get in estimating the temperature T via measuring the probe ρ_β. That is, given an estimator $\hat{T}$ [1] of the temperature, the Cramér-Rao bound [76], specified to the case of temperature estimation [40, 41], bounds the precision of any $\hat{T}$ as

$$\frac{\langle(\hat{T}-T)^2\rangle}{T^2} \geq (\nu\mathcal{C})^{-1}\ , \tag{3.3}$$

where ν is the number of i.i.d. repetitions of the experiment. Moreover the bound can be saturated by performing energy measurements. Notice that such bound is valid only in the above mentioned typical scenario, with uncorrelated repeated measurements of the fully thermalized probe. Other metrological schemes can involve partially thermalized probes [40, 41].
An *optimal thermometer*, is therefore, in the above setting, a probe exhibiting the maximal value of $\mathcal{C}$. In such case the problem of optimal thermometry reduces to identifying, under physical constraints, systems having large $\mathcal{C}$. Such a problem has been addressed in terms of the Hamiltonian spectrum of the probe in [42] that is, assuming the only constraints to be the (finite) Hilbert space dimension of the thermal probe and a fully controllable Hamiltonian spectrum otherwise. The resulting optimal spectrum consist in an effective two-level system, with a single ground state and an exponential degeneracy of the excited level. At the same time, no physical system construction on how to realize the spectrum of such optimal Hamiltonian was given. It would be therefore desirable to characterize optimal thermal probes under additional, physically motivated assumptions. In particular, checking whether the optimal probes can be built with realistic interacting systems that use only *(i)* 2-body interactions, and possibly only *(ii)* short-ranged interactions. In Ref. [8] we tackled this problem answering both questions affirmatively.

[1] An estimator $\hat{T}(x_1,\ldots,x_\nu)$ is a function of the ν stochastic outputs of the measurement chosen for the estimation protocol.

3.2 Spectrum-based thermometry optimization

Before going to the main results of Ref. [8], it is useful to understand the maximization of $\mathcal{C}$ that has been performed in [42] in terms of the Hamiltonian spectrum, as well as additional considerations that we make below. In terms of the eigensystem $\{E_i, |E_i\rangle\}$ of the probe's Hamiltonian H the state populations of $\omega_\beta(H)$ read

$$p_i \equiv Z_\beta^{-1} e^{-\beta E_i} \ , \quad \text{with} \quad Z_\beta = \sum_i e^{-\beta E_i} \ . \tag{3.4}$$

so that $\omega_\beta = \sum_i p_i \, |E_i\rangle \langle E_i|$. In such basis it is easy to verify from (3.2) that the heat capacity is proportional to the energy variance of the thermal state, that is

$$\mathcal{C}(H,\beta) = \beta^2 \Delta_\beta^2 H \ , \tag{3.5}$$

$$\Delta_\beta^2 H = Z_\beta^{-1} \sum_{i=1}^{D} E_i^2 e^{-\beta E_i} - \left(Z_\beta^{-1} \sum_{i=1}^{D} E_i e^{-\beta E_i} \right)^2 \ . \tag{3.6}$$

Such expression clarifies that the heat capacity only depends on the spectrum of the Hamiltonian H. Moreover, there is a clear scale-symmetry as $\mathcal{C}(\lambda H, \lambda^{-1}\beta) = \mathcal{C}(H,\beta)$. For this reason, we will often consider temperature units in which $\beta = 1$ and simply refer to $\mathcal{C}(H) := \mathcal{C}(H, \beta = 1)$. Finally, here and throughout the work, we notice that a global shift in the energy levels makes no difference in the thermal state nor for the heat capacity $\mathcal{C}(H) = \mathcal{C}(H + c\mathbb{1})$. ($c$ is any real number).

As mentioned above, the maximization of $\mathcal{C}$ of a generic D-dimensional system at thermal equilibrium has been carried out in [42] assuming full-control of the spectrum and its Hamiltonian,

$$\mathcal{C}^{\text{spec}}_{\text{max}}(D) := \max_{H|\dim H = D} \mathcal{C}(H) \tag{3.7}$$

The resulting optimal spectrum consists of a single ground state and a $(D-1)$-degenerate excited state, that is

$$H_{\text{deg}} = 0 \, |0\rangle \langle 0| + \sum_{i=1}^{D-1} E \, |i\rangle \langle i| \ , \tag{3.8}$$

with an *optimal gap* $x = \beta E$ in temperature units that satisfies the transcendental equation $e^x = \frac{(D-1)(x+2)}{(x-2)}$. The corresponding heat capacity is $\mathcal{C}^{\text{spec}}_{\text{max}}(D) = x^2 e^x (D-1)/(D-1+e^x)^2$ [42]. This expression gives in the

asymptotic regime of large probes ($D \to \infty$) $x \simeq \ln D$, hence

$$\mathcal{C}_{\max}^{\text{spec}}(D) \simeq (\ln D)^2/4 \, . \tag{3.9}$$

Consider now the case in which the system is composed by N constituents. For reasons that will become clearer later we consider units with minimal dimension, i.e. a system made of N (quantum) bits. It follows that

$$D = 2^N \, , \tag{3.10}$$

thus the theoretical bound of [42] translates in this case to

$$\mathcal{C}_{\max}^{\text{spec}}(2^N) \simeq \frac{N^2(\ln 2)^2}{4} \, , \quad \beta E \simeq N \ln 2 \, , \tag{3.11}$$

for $N \gg 1$, i.e., yielding a quadratic scaling in the number of particles. This "Heisenberg-like" scaling of the thermal sensitivity is to be confronted with the typical extensive behaviour of the heat capacity (i.e. linear in N).
Without any restriction on the possible interactions among the N spins it is always possible to generate the Hamiltonian (3.8) and saturate the theoretical maximum value $\mathcal{C}_{\max}^{\text{spec}}$ of the heat capacity (see e.g. [131], where the authors make use of arbitrary N-body interactions).
The question we address in [8] is what happens if we restrict ourselves to physically motivated Hamiltonians, which cannot have arbitrary spectrum in general. To solve this problem, it is useful to understand the origin of the N^2 scaling, that is what features of the ideal Hamiltonian (3.8) are relevant. The peculiarity of the degenerate model (3.8) can be appreciated in terms of the thermal ground state probability

$$p_0 = \frac{1}{1 + (D-1)e^{-\beta E}} \, , \quad D = 2^N \, . \tag{3.12}$$

At low energies, the value of p_0 is ~ 0, meaning that in the thermal state, the population is spread evenly in the degenerate excited subspace until $E \sim (N \ln 2)/\beta$, where it "suddenly" climbs to ~ 1. This transition has width $\sim 1/\beta$, where the system experiences the peak in heat capacity, while for smaller (resp. larger) values of E, the whole population collapses to the excited subspace (resp. ground state), with no energy variance. As the optimal gap is linear in N, the resulting variance scales quadratically. It should be noticed that a similar scaling is obtained also when substituting an effective dimension that scales exponentially $D' \propto \theta^N$ for any $\theta > 1$. In such case one would obtain $\mathcal{C}_{\max}^{\text{spec}}(\theta^N) \simeq (N \ln \theta)^2/4$. That being, the exponential (in N) degeneracy of the first excited level is the first main ingredient for a system

exhibiting such quadratic scaling of the heat capacity. Thanks to the following lemma, which we prove in Appendix B, such a condition is also sufficient, i.e., if one can generate a single ground state and an exponentially degenerate first excited state, albeit not necessarily $(2^N - 1)$-degenerate, the obtainable maximal heat capacity scales quadratically in N.

Lemma 1. *Consider two Hamiltonians, H_1 and H_2, such that H_1 has 1 ground state and a k_1-degenerate excited state ($k_1 + 1$ levels in total), while H_2 has the same spectrum and additional excited states above (totaling $1 + k_1 + k_2$ levels),*

$$H_1 = 0\,|0\rangle\langle 0| + \sum_{i=1}^{k_1} E\,|i\rangle\langle i| \ , \tag{3.13}$$

$$H_2 = 0\,|0\rangle\langle 0| + \sum_{i=1}^{k_1} E\,|i\rangle\langle i| + \sum_{\alpha=k_1+1}^{k_1+k_2} E_\alpha\,|\alpha\rangle\langle\alpha| \ , \tag{3.14}$$

with $0 \leq E \leq E_\alpha \ \forall\alpha$. Assuming control over the first excited gap E, the maximum heat capacity obtainable with H_2 is always larger than the maximum heat capacity obtainable with H_1.

$$\max_E \mathcal{C}(H_1) \leq \max_{E\leq E_\alpha} \mathcal{C}(H_2) \ . \tag{3.15}$$

That is, additional excess levels above the k_1-degeneracy of H_1 can only increase the maximal heat capacity. Notice that the control over E, while keeping $E_\alpha \geq E \ \forall\alpha$, can be easily obtained by rescaling all the parameters of H_1 or H_2 globally. The proof of the lemma is given in the technical material of [8] and in App. B.

What to look for in an ideal thermometer. Following from the above discussion we know what is the desired spectrum in any constrained maximisation of the heat capacity, that is, *a single ground state and the largest possible degeneracy of the first excited level*, regardless of the value of larger levels. As the system size grows, any model featuring an exponential degeneracy of the first excited level and a tunable gap, will show the desired Heisenberg-like scaling of the heat capacity.

3.3 Optimal spin thermal probes

According to the above discussion, in Ref. [8] we tackled the problem of maximising the heat capacity of realistic systems, that is, *engineering the best thermometer*. To address the question, we chose as a platform a generic system

of spins, such as those currently programmable in quantum annealers, whose thermodynamics is starting to be studied [43–45], and which represent flexible physical devices with a high degree of control. Motivated by real-life feasibility, we require the interactions among the spins to be two-body typed. In the case of a classical system, this means that the general N-spins Hamiltonian we will consider is of the kind

$$H = \sum_{i}^{N} h_i \sigma_i^z + \sum_{i<j}^{N} J_{ij} \sigma_i^z \sigma_j^z \ . \tag{3.16}$$

We considered also the case of fully quantum-mechanical spin Hamiltonians. However, numerical evidence showed that no relevant advantage is given by adding off-diagonal interactions involving σ^x and σ^y terms to the Hamiltonian (3.16). Moreover, let us mention that although (3.16) allows, in principle, to connect any two spins with an interaction of strength J_{ij}, Hamiltonians with short-range interactions are strongly desirable for such systems to be scalable in size.
In the following, we showcase the main results of [8]. There, we demonstrated that it is possible to design a thermal probe ("Adam model") consisting of N interacting spins that approximates the maximum value $\mathcal{C}_{\max}^{\text{spec}}$ of the thermal sensitivity (3.11). Moreover, restricting ourselves to short-range interactions, we showed that it is still possible ("Abel model") to obtain a heat capacity exhibiting the same scaling as (3.11) with a prefactor that can be made arbitrarily close to the maximum.

3.3.1 Adam model

We maximised numerically $\mathcal{C}(H)$ using Eq. (3.6) and constraining H to be of the form (3.16). More precisely (see details in [8]), using techniques commonly employed in the machine learning community, the numerical routine starts from a random choice of the h_i and J_{ij} parameters in (3.16), and reiteratively improves them. At each iteration, we compute the gradient of $\mathcal{C}(H)$ with respect to h_i and J_{ij} using backpropagation [132], and we use the Adam optimizer (Adaptive Moment Estimation) [133] to update the parameters using the gradient. The Adam method is similar in spirit to gradient descent, but includes various improvements and is often found to converge better. After repeating the optimization for different total numbers of spins N, a recurrent pattern emerged, which we dub "Adam model" and describe here.

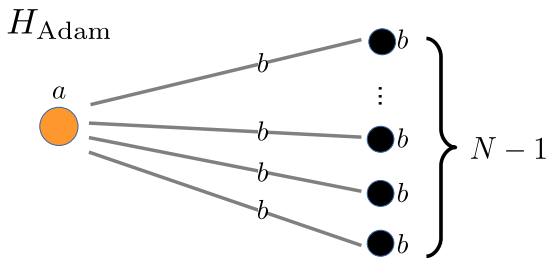

Figure 3.1: Schematic representation of the Adam model (cf. Hamiltonian (3.17).

Let us consider the following system Hamiltonian describing particular coupling choices for N spins (labeled from 1 to N) in (3.16)

$$H_{\text{Adam}[N]}(a,b) := a\,\sigma_1^z + b\sum_{i=2}^{N}\sigma_i^z\,(\mathbb{1}+\sigma_1^z)\ , \tag{3.17}$$

where a and b are two coefficients. A representation is given in Fig. 3.1. The resulting spectrum has 2 main classes of eigenstates. First, $\binom{N-1}{k}$-degenerate states with energy

$$E_k = a + 2b(k-(N-1-k)), \quad \text{for } k=0,\dots,N-1\ , \tag{3.18}$$

corresponding to the first spin being up $\sigma_1^z = 1$, and k spins up among the remaining $N-1$ ones. If instead the first spin is down $\sigma_1^z = -1$, we get a 2^{N-1} times degenerate excited state at energy

$$E_{\text{deg}} = -a\ . \tag{3.19}$$

That is, the simple topology and choice of the couplings in (3.17) makes the first spin σ_1^z being able to "switch on and off" the effective magnetic field on the remaining spins, generating an exponential degeneracy of the E_{deg} level. Moreover, it is easy to see that by choosing

$$b>0 \quad \text{and} \quad b(N-3)\le a < b(N-1) \tag{3.20}$$

one ensures

$$E_0 < E_{\text{deg}} \le E_k \quad \text{for } k=1,\dots,N-1\ . \tag{3.21}$$

That is, a single ground state, and 2^{N-1}-degenerate first excited level. By saturating $b(N-3)=a$ one gets $E_{\text{deg}} = E_1$ totalling a $2^{N-1}+N-1$

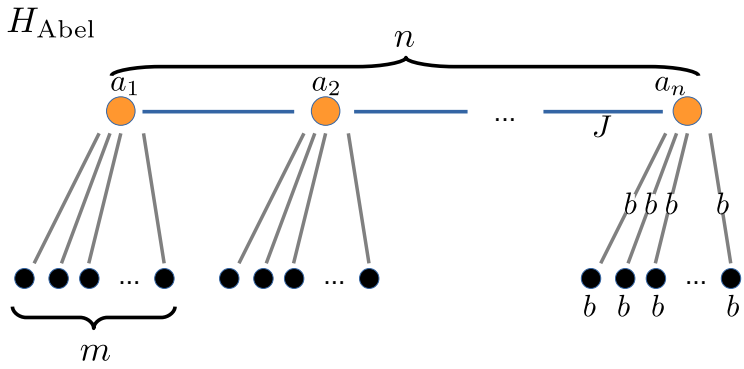

Figure 3.2: Schematic representation of the Abel model (cf. Hamiltonian (3.24). In the Abel model the total number of spins is $N = n(m+1)$.

degeneracy for the first excited [2]. Notice that Lemma 1 ensures that such model can achieve at least the heat capacity $\mathcal{C}^{\mathrm{spec}}_{\mathrm{max}}(2^{N-1})$, that is

$$\mathcal{C}^{\mathrm{spec}}_{\mathrm{max}}(2^{N-1}) \leq \mathcal{C}^{\mathrm{Adam}[N]}_{\mathrm{max}} \leq \mathcal{C}^{\mathrm{spec}}_{\mathrm{max}}(2^{N}) \ , \tag{3.22}$$

which in the asymptotic limit

$$\mathcal{C}^{\mathrm{Adam}[N]}_{\mathrm{max}} \gtrsim \frac{(N-1)^2(\ln 2)^2}{4} \tag{3.23}$$

becomes indistinguishable from the theoretical bound $\mathcal{C}^{\mathrm{spec}}_{\mathrm{max}}(2^{N})$ (3.11) (cf. Fig. 3.4).

3.3.2 Abel Model

As the Adam model arises from an unconstrained numerical optimization of $\mathcal{C}$ for Hamiltonians of the form (3.16) (cf. above Sec. 3.3.1), we conjecture it to be the global optimum for such class. At the same time, in physical implementations, the star-shaped connectivity of (3.17) (cf. Fig. 3.1) cannot be scaled to any number of particles. For this reason, we generalized the Adam model to a similar configuration that only has short-range interactions, the Abel model. Specifically, inspired by the Adam model, we consider $N = n(m+1)$ spins with Hamiltonian (as from Figure 3.2)

[2] We conjecture that $2^{N-1} + N - 1$ is the maximum obtainable degeneracy of the first excited level, in Hamiltonians of the form (3.16) with a single ground state.

$$H_{\mathrm{Abel}[N]}(a_\alpha, J, b) := \sum_\alpha a_\alpha \sigma^z_\alpha - J \sum_\alpha \sigma^z_\alpha \sigma^z_{\alpha+1} + b \sum_{\alpha,i} (\sigma^z_\alpha + \mathbb{1}) \sigma^z_{\alpha,i} \ , \tag{3.24}$$

where α is the index identifying the central spin of each Adam-like sub-unit, while (α, i) selects the ith spin in each sub-unit, i.e.

$$\alpha = 1, \dots, n \ , \quad i = 1, \dots, m \ . \tag{3.25}$$

This model guarantees a $2^{n_\downarrow m}$ degeneracy for each energy level with $n_\downarrow$ α-spins. In particular, when all the α-spins are down, i.e. $\sigma^z_\alpha = -1 \ \forall \alpha$, it obtains a 2^{mn} degeneracy. At the same time, one needs to make such configuration to be the first excited state, with a single ground state. One way to do this is by making the couplings J strong enough to force all the n α-spins to be the same. In such case one is left with two configurations that is $\sigma_\alpha = 1 \ \forall \alpha$, or $\sigma_\alpha = -1 \quad \forall \alpha$. In the first case, one has $E = \sum_\alpha a_\alpha + 2b \sum_{\alpha,i} \sigma_{\alpha,i}$ while in the second case $E = -\sum_\alpha a_\alpha$ with degeneracy 2^{nm}. Calling then $\sum_\alpha a_\alpha := h$, we have

$$2^{mn} \text{ levels with } E = h + 2b \sum_{\alpha,i} \sigma_{\alpha,i} \ , - (n-1)J \tag{3.26}$$

$$2^{mn} \text{ levels with } E = -h - (n-1)J \ , \tag{3.27}$$

and $(2^n - 2)2^{mn}$ remaining neglected levels. Notice that Eq.s (3.26) and (3.27) effectively recreate the same spectrum of the Adam model, with $N' = nm$ (instead of $N' = N - 1$). Thanks to the Lemma 1, we know that such model, by choosing correctly the couplings, can get a heat capacity at least as large of that of a system having 1 ground state and $2^{mn} + mn$ degeneracy of the first excited. This is done by choosing

$$-h = h - 2bmn + 4b \to b(mn - 2) = h \ . \tag{3.28}$$

With such choice one gets $\mathcal{C}_{\max} \sim \frac{(\ln(2^{mn}+mn))^2}{4} \simeq \frac{1}{4}(\ln 2)^2 (mn)^2$ which is essentially quadratic in $N = n(m+1)$, i.e.

$$\mathcal{C}^{\mathrm{Abel}[N]}_{\max} \gtrsim \left(\frac{m \ln 2}{2(m+1)} \right)^2 N^2 \ . \tag{3.29}$$

Remark 1. It is clear from (3.29) how larger values of m increase the obtainable heat capacity. For $m = N - 1$ the Abel model coincides with the Adam model ($n = 1$, cf. Fig. 3.1). For short range interactions there will be a maximum m allowed. We notice that for $m = 3$, the Abel model could

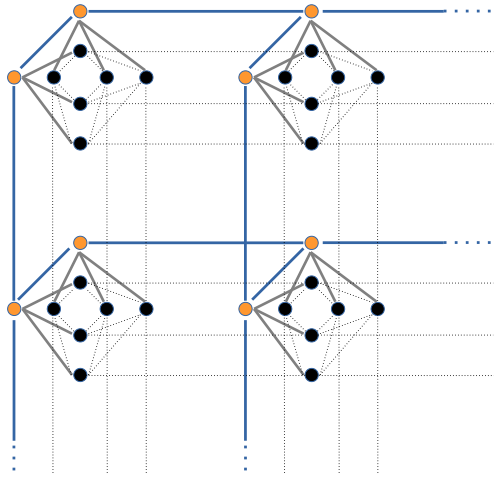

Figure 3.3: Embedding of the Abel model for $m = 3$ (see Fig. 3.1) into the Chimera graph, which is used by D-Wave Systems [134]. Dotted lines are unused couplings of the Chimera architecture.

be programmed on current D-Wave annealers, as the topology of the interactions of (cf. Fig. 3.1 and Fig. 3.3) can be embedded, for $m = 3$, in the Chimera graph, corresponding to the D-Wave processors [134]. This means that, as from (3.29), a programmable spin network in the Chimera graph can reach at least $\sim \frac{m^2}{(m+1)^2} = 9/16$ of the overall bound $\mathcal{C}^{\text{spec}}_{\max}(2^N)$ (3.11). Moreover, surprisingly, numerical optimization of the $\mathcal{C}$ for the Chimera model results in the Abel model with $m = 3$ [8].

Remark 2. We point out an unexpected practical feature of both the Adam and Abel model. In both models, when performing the temperature estimation via energy measurements, *it is enough to measure a single spin.* In fact, in the regime of large N, the only relevant energy levels contributing to the Gibbs state are the ground level and the first excited (levels above are exponentially suppressed in the statistics), which are indicated by the value of σ^z_1 for the Adam model (3.17), or *any* of the σ^z_α in the Abel model (3.24).

3.4 Spin probes comparison

As a results-summary, in Figure 3.4 we compare the maximum values of the heat capacity $\mathcal{C}$ for different models of spin Hamiltonians, as the number N of spins grows. The Adam and Abel models show a quadratic scaling in N

that eventually surpasses all standard models – such as the Ising model in 1D (and 2D), as well as a model of uniform “all-to-all" interactions – as these show instead the standard thermodynamic extensive scaling (i.e., linear in N) of the heat capacity. We briefly describe each of the compared models in the following paragraphs.

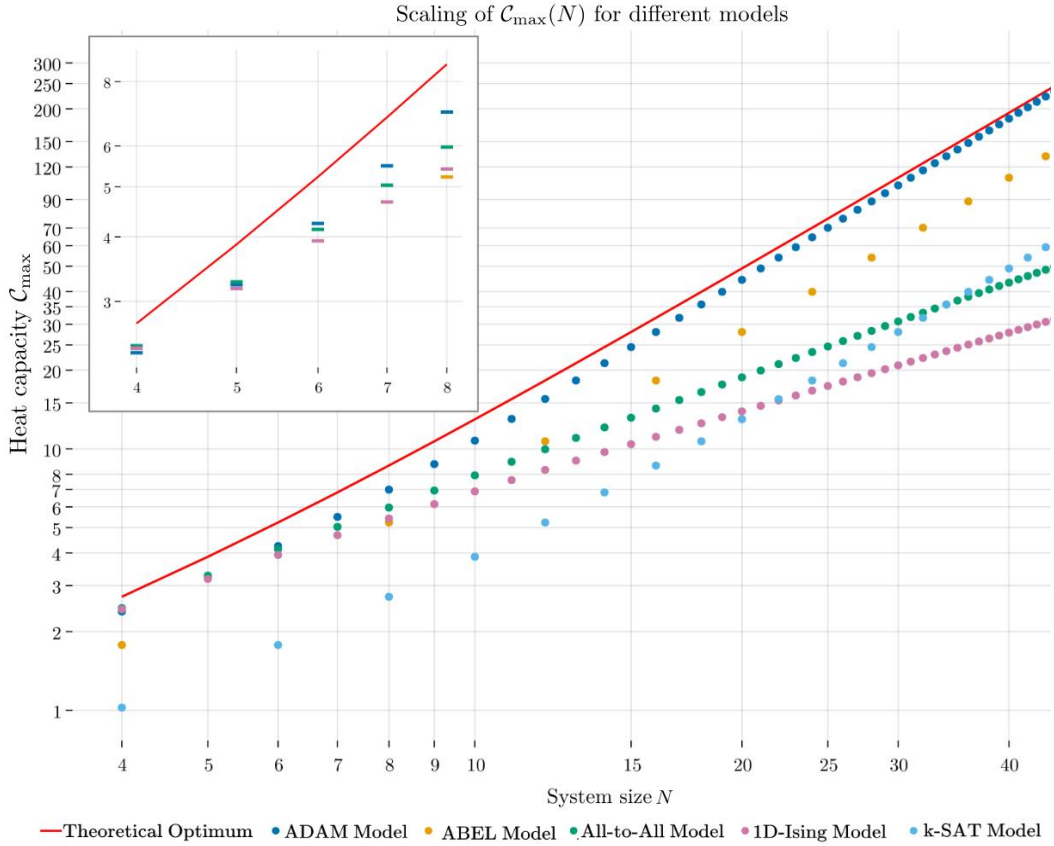

Figure 3.4: Maximum heat capacity $\mathcal{C}_{\max}$ and its scaling for N spins employed in different models of interaction. Both the Adam model and Abel model show a Heisenberg scaling $\propto N^2$ of the optimal sensitivity, therefore outperforming standard models that have an extensive scaling of $\mathcal{C}$. For $N \geq 6$ the Adam model is numerically optimal among all possible spin Hamiltonians (3.16), and asymptotically saturates the mathematical bound 3.11.

Ising Lattices. The 1D Ising model stands as the most superficial candidate for a spin graph thermometry probe. For N spins it is defined by the Hamiltonian

$$H_{1D}(\vec{h},\vec{J}) := -\sum_{i=1}^{N} h_i\sigma_i^z - \sum_{i=1}^{N} J_i\sigma_i^z\sigma_{i+1}^z, \tag{3.30}$$

where we chose periodic boundary conditions $\sigma_{N+1} \equiv \sigma_1$. The heat capacity in such model can be efifciently computed from with standard techniques [135]. Numerically maximizing it leads consistently to homogeneous interactions $J_{i,j} = J$ and local fields $h_i = h$. As expected, an Ising chain probe will at most achieve a linear scaling of the heat capacity in N, as seen in Figure 3.4. Notice that a 2-dimensional Ising model can achieve a slightly higher scaling i.e. $\mathcal{C} \propto N \ln N$ [136, 137].

All-to-All Model We considered also a model that is completely symmetric under permutations, with all-to-all interactions, i.e.

$$H_{\text{All}}(h,J) := -h\sum_{i=1}^{N}\sigma_i^z - J\sum_{i<j}\sigma_i^z\sigma_j^z, \tag{3.31}$$

describing a complete graph with homogeneous interactions $J > 0$ and local fields $h > 0$. Taking the systems' symmetries in to account we get the following $\binom{N}{k}$-degenerate eigenenergies for $k \leq N$ up-spins:

$$E_k = h(N-2k) + \frac{J}{2}\left[4k(N-k) - N(N-1)\right]. \tag{3.32}$$

This expression makes it easy to visualize the spectrum and the corresponding degeneracies. The All-to-All-models' spectrum shows pairs of sufficiently proximate excited states, although not exponentially degenerate [8]. In fact, for small $N \leq 5$ the All-to-All model appears to have the highest heat capacity among the considered models (see inset in Fig. 3.4) and consistently emerged from numerical optimisation in the same regime.

3.4.1 k-SAT model and the exponential degeneracy

Finally, we notice that in Refs. [50, 138] an Hamiltonian replicating a global AND-operation between M logical bits (represented by M spins) was introduced, with the aid of M additional ancillary spins. Such Hamiltonian was proposed as the basic element to build general models to solve k-SAT problems. We will thus refer to it as the *k-SAT model*. A logical AND identifies a single string (without loss of generality, the string given by $111\ldots1$, M

Model	1st excited deg.	Asymptotic Max. $\mathcal{C}$	Short-range?
k-Sat [50]	$2^{\frac{N}{2}}-1$	$\sim \frac{(\ln 2)^2}{4}\frac{N^2}{4}$	✗
Adam	$2^{N-1}+N-1$	$\sim \frac{(\ln 2)^2}{4}(N-1)^2$	✗
Abel	$2^{\frac{mN}{m+1}}+\frac{mN}{m+1}$	$\sim \frac{(\ln 2)^2}{4}\frac{m^2N^2}{(m+1)^2}$	✓

Table 3.1: Models recreating an effective spectrum with a single ground state and an expontially degenerate first excited level, plus additional levels above. Both Adam and Abel perform better than the k-Sat model, in terms degeneracy of the first excited level. This is reflected in the value of the maximal heat capacity $\mathcal{C}$ (cf. Fig. 3.4). Moreover the Abel model has the additional feature of being short-ranged.

times) with an energy E_{AND} different from the energy $E_{\overline{\text{AND}}}$ associated to all the other 2^M-1 logical strings. Formally, this spectrum coincides with the ideal two-level degenerate model (3.8). At the same time, the construction uses a total $N=2M$ of spins. This means that the Hamiltonian proposed in [50, 138] exhibits the desired single ground state and exponential degeneracy of the first excited level, thus achieving a quadratic scaling of the maximum $\mathcal{C}$ (cf. Fig. 3.4).However, the effective dimension of the logical spins' system is $D'=2^{N/2}$, while considering the ancillary spins the whole space has dimension $D=2^N$ (the neglected levels correspond to energies that can be made arbitrarily high, see [50]). In other words, an overhead of $M=N/2$ spins was used to achieve a spectrum with single ground state and $\sim 2^{N/2}$ degeneracy of the first excited level. The models proposed in our work (Adam and Abel) achieve the same while using a much smaller overhead, that is, a 1-spin overhead for the Adam model and a $N/(m+1)$-spins overhead for the Abel model. In table 3.1 we compare these quadratic models.

3.5 Discussion

In Ref. [8] we took on the problem of maximising the heat capacity $\mathcal{C}$ of physically feasible systems. From the operational point of view, this corresponds to engineering the best probe for temperature estimation, in the context of equilibrium thermometry. We showed that it is possible to engineer simple probes based on interacting spins that can approximate asymptotically (i.e. in the limit of large size) the maximum mathematical of $\mathcal{C}$ based on the sole Hilbert space dimension, which was found in [42]. That is, we showed that such thermal sensitivity can essentially be achieved via 2-body interactions (Adam model, Sec. 3.3.1). Additionally, an arbitrary good approximation

can be achieved also when requiring short range interactions (Abel model, Sec. 3.3.2).
In terms of Hamiltonian spectrum engineering, by generalizing some of the results of [42], we showed that the essential requirement for an optimal thermal probe made of N constituents, is the presence of a single ground state and the exponential (in N) degeneracy of the first excited level, independently of higher spurious levels. This effective two-level spectrum appears not only in problems of heat capacity maximisation, but also in thermometry out of equilibrium [139], Otto cycles optimization [4, 122], protein folding modelling [46, 47], adiabatic Grover's search [48, 49], and energy based boolean functions [50]. We thus speculate that our work might have application in these subjects.
In particular our Abel model can be programmed on current quantum annealers. An experimental demonstration is therefore at hand. At the same time, it is important to assess possible noise-sources and noise-tolerance (which is partially assessed in [8]) of our results, as well as taking into account the thermalization timescale, which depends on the specifics of the experiment.
Another interesting direction would be that of engineering "critical" probes to study artificial phase-transitions, or meant to measure different parameters (such as the magnetic field). Finally, similar problems could be studied on different physical platforms, such as cold atom arrays.

Part II

Tasks in Information Theory: geometry and optical proposals

The following part of the thesis is devoted to the analysis and optimisation of operational tasks that belong to the field of Quantum Information Theory, specifically in the areas of Quantum Markovianity, Quantum Nonlocality, and entanglement witnessing.

First, in Chapter 4, we study memory effects in open systems dynamics. The regime in which a system monotonically loses information to a (typically much bigger) environment is dubbed *Markovian regime*. Markovian dynamics, or Markov models, have been studied for more than a century and are typically introduced as those probabilistic dynamics whose update in time does not depend on the complete history of the state's evolution (in this sense is *memoryless*), but rather only on the immediately previous time-step, i.e. $P(i_k, t_k|j_{k-1}, t_{k-1}; j_{k-2}, t_{k-2}; \dots) = P(i_k, t_k|j_{k-1}, t_{k-1})$. While in the classical case the set of Markovian evolutions is well defined by the above equation, such conditional probabilities are not well defined in quantum scenarios, which makes the definition of Markovian dynamics for quantum open systems non-trivial [53, 54, 140]. A condition that is equivalent, in the classical case, at the level of single-time tomography [53], and is canonically extended to the quantum scenario, is that of *stochastic-divisibility*. That is, Markovian are those physical processes in which each intermediate evolution between two times $t_2 > t_1$ satisfies the same mathematical properties of the total evolution from time 0 to t, which is dubbed "stochastic". Markovian models are ubiquitous in the mathematical modelling of physical processes, bio-processes, economics etc.. At the same time, non-Markovian effects are of interest from the theoretical point of view, both for understanding when the Markovian approximation breaks down, and to explore its (possibly beneficial) effects in different operational tasks.
The theoretical study of Markovian and non-Markovian dynamics has therefore two fundamental objectives: *i)* the full characterization of Markovian evolutions from the mathematical point of view, and *ii)* the collection of practical tasks that can be boosted by non-Markovian effects during the evolution. We contribute to both purposes, by analysing the behaviour of the *Fisher Information* [56], which is a distance-based quantifier of information whose mathematical and operational properties have been intensively studied both in mathematics and physics. We characterise completely the relation between the variation of Fisher Information in time, the underlying geometry of the set of states, and information revivals in the several operational tasks to which the Fisher Information is associated (state discrimination, error correction, metrology, etc.) [7].

Secondly, in Chapters 5 and 6, we move to the field of Bell nonlocality, or simply nonlocality. Broadly speaking, nonlocality studies the gap between the statistical structure of local hidden variable models and quantum mechanics. That is, given the probabilistic outputs of an experiment whose devices are left uncharacterized, nonlocality studies if the resulting probability distribution is compatible with an underlying classicality of the unobserved sources and measurements, or not, with minimal assumptions regarding the causal separation among different parts of the experiment. It was first formalised by Bell [58] in the simple scenario of two non-communicating parties, Alice and Bob, performing experiments locally in their laboratories, with the possibility of previously distributing shares of a common random source, which might be classical or quantum [59]. Nonlocality experiments are of great interest both for their relevance in the foundations of quantum mechanics and for their technological applications in communication theory, cryptography, randomness generation, entanglement certification, and others. At the same time, experiments in this area are typically quite challenging, as they rely on the strongest forms of entangled states and measurements, and are easily fragile to noise.

Accordingly, we consider optimization of theoretical nonlocal tasks as bringing them closer to technological feasibility, particularly using quantum optics. More precisely, we study variations and relaxations of the original Bell scenario that make an experimental realisation of nonlocal experiments suitable for simple *photonic* platforms.

In particular, in Chap. 5 we design an experiment that is able to demonstrate the nonlocality of a simple class of entangled states by using only standard passive optics. The maximally entangled state $|\psi^+\rangle := \frac{|01\rangle+|10\rangle}{\sqrt{2}}$ can be encoded in the Fock basis on two different modes, in particular with a single photon being in a superposition of being sent in two different directions. This is perhaps the simplest entangled state that can be created easily in a laboratory, by heralding a single photon and sending it to a balanced beamsplitter (i.e. with transmissivity = reflectivity = 50%). This state cannot violate Bell locality when using measurement apparatuses consisting only of simple passive optics (cf. Chap. 5). Nevertheless, we study its nonlocal properties in the *network scenario*. In fact, the field of nonlocality has been generalized to the case of general causal networks of parties sharing unobserved classical/quantum sources, which show a range of new phenomena [60] with respect to the standard Bell scenario. In particular we provide an experimental proposal in which single photon entangled states can violate locality in the *triangle network* without using additional photon-field excitations in the measurement apparatus.

In Chapter. 6 instead, we study a different relaxation of the Bell scenario, in

which parties A and B are allowed to input verified quantum states in their laboratories, therefore assuming additional level of trust in their experiment with respect to the standard Bell experiment. In fact, the Bell scenario is sometimes referred to as the Device-Independent (DI) scenario, as it does not assume any particular form of the devices and sources used by A and B. Allowing for trusted quantum inputs leads to the so called Measurement-Device-Independent (MDI) scenario [64], which offers the possibility of certifying entanglement of any entangled state [65], something impossible in the standard Bell scenario. Among our results we use this framework to design an experimentally friendly MDI entanglement witness for all entangled gaussian states, which can be realised using only standard linear optics and homodyne detection.

Chapter 4

The geometry of (non-)Markovianity and information flow

This chapter is based on Ref. [7] :
"P. Abiuso, M. Scandi, D. De Santis and J. Surace, *Characterizing (non-)Markovianity through Fisher Information*, arXiv preprint 2204.04072, (2022)."

4.1 Background: Markovianity

Quantum Mechanics postulates that the evolution of an isolated system is unitary. At the same time, isolated systems are difficult to confine in a laboratory, and even more difficult to encounter in nature. Virtually any system evolves interacting with external degrees of freedom, that typically belong to a much larger *environment*. For these reasons it is of paramount importance for physicists to study and understand *open systems dynamics* [51, 141].
The information contained in an evolving open system may undergo two possible dynamical regimes, dubbed Markovian or non-Markovian. Markovian evolutions are characterized by *memoryless* environments, where at each time the dynamical trajectory can be represented by physical transformations that solely depend on the immediately previous time step [51, 141]. It follows that in this regime the information contained in the system undergoes a monotonic degradation, being dispersed into the environment irreversibly. On the other hand, non-Markovian dynamics are distinguished by memory effects, meaning that at any time the evolution might in general depend on all the previous time steps of the trajectory. This property allows information to flow back into the system. This regime attracted interest both at a theoretical level, as well as for its operational consequences in various quantum information

scenarios, such as for quantum key distribution [142], metrology [143], entanglement generation [144], quantum communication [145], quantum teleportation [146], quantum thermodynamics (cf. our Ref. [9], which we did not include in this thesis).

4.1.1 Mathematical description

Consider first the case of finite-dimensional classical dynamics. In such case the state of a system is defined by a probability vector $\boldsymbol{p}$, having components p_i that represent the probability of the system being in the microstate i. It follows that the set of *states* is given by the symplex $\mathcal{S}(\mathbb{R}^N)$ defined as

$$\mathcal{S}(\mathbb{R}^N) := \{\boldsymbol{p} \in \mathbb{R}^N | p_i \geq 0 \wedge \sum_i p_i = 1\} \ . \tag{4.1}$$

The dynamics can be represented as a linear [1] map acting on the set of states. More precisely, a stochastic map (or *channel*) $T^{(t,0)}$ is a linear operator that evolves $\boldsymbol{p}$, from a time 0 to t,

$$T^{(t,0)}[\boldsymbol{p}(0)] = \boldsymbol{p}(t) \ . \tag{4.2}$$

It follows, from the requirement of mapping any positive normalized vector to a positive normalized vector, that for every t,

$$\sum_i T^{(t,0)}_{ij} = 1 \ , \quad T^{(t,0)}_{ij} \geq 0 \ \ \forall \ i,j \ . \tag{4.3}$$

In fact, the elements $T^{(t,0)}_{ij}$ can be interpreted as the conditional probability of ending up in microstate i at time t starting from j at time 0, i.e. $T^{(t,0)}_{ij} = P(i,t|j,0)$.
We call *evolution* a family of stochastic maps parametrized by time, i.e. $T := \{T^{(t,0)} | 0 \leq t\}$. Assuming the channel to be continuous and differentiable in time, as well as invertible [2], one can define the intermediate channel $T^{(t,s)}$ between two increasing times s and t as

$$T^{(t,s)} \equiv T^{(t,0)} \circ {T^{(s,0)}}^{-1} \ , \quad t \geq s \geq 0 \ . \tag{4.4}$$

[1] The linearity of the map follows from consistency with the physical possibility of preparing convex mixtures of states.

[2] Invertibility of the maps can can be always physically satisfied by introducing undetectable ε-noises to the dynamics.

Channels with these properties constitute the class of *smooth evolutions*. A smooth evolution T is called *stochastic-divisible* or *Markovian* [53, 141] if for any partition of the interval $[0, t]$ it can be split into intermediate channels, all of which are stochastic

$$T^{(t,0)} = T^{(t,t_{k-1})} \circ T^{(t_{k-1},t_{k-2})} \circ \cdots \circ T^{(t_1,0)} \ . \tag{4.5}$$

More precisely, Markovian evolutions are multi-time processes $\{t_0, t_1, \dots\}$ where the evolution of the state only depends on the latest previous sampling of it, i.e.

$$P(i_k, t_k | j_{k-1}, t_{k-1}; j_{k-2}, t_{k-2}; \dots) = P(i_k, t_k | j_{k-1}, t_{k-1}) \ , \tag{4.6}$$

from which the dynamics is said to be memoryless. This condition is equivalent to Eq. (4.5) when restricting the tomography of the dynamics to two-times probabilities [3], such as the case of quantum dynamics, in which single-shot measurements inevitably interfere with the dynamics. It should be therefore kept in mind that we will identify Markovian evolutions with stochastic-divisible evolutions.

The case of quantum dynamics. In the case of (finite-dimensional) quantum dynamics, a state is represented by a density matrix ρ, that is, a positive definite matrix acting on a complex Hilbert space $\mathbb{C}^N$, with unit trace. As such, one could define physical evolutions as those linear superoperators that map normalised density matrices to normalised density matrices, similarly to (4.3) above. These form the class of Positive Trace Preserving (PTP) superoperators [51, 54]. It turns out, however, that when considering multipartite systems with trivial dynamics on ancillary degrees of freedom, there exist quantum (entangled) states ρ_{AB} and PTP maps $\mathcal{T}$ which, when considered together with the trivial dynamics of the ancillary degrees of freedom, stop being positive. That is, $\mathcal{T}$ is PTP, but $\mathcal{T} \otimes \mathbb{1}$ is not PTP ($\mathbb{1}$ is the identity operator) and it exists a bipartite state $\rho_{AB} \geq 0$ such that $\mathcal{T}_A \otimes \mathbb{1}_B[\rho_{AB}]$ is not positive definite. For this reason, physical maps for open quantum systems are defined to be those maps that are PTP and for which any extension $\mathcal{T} \otimes \mathbb{1}$ with trivial dynamics on ancillary degrees of freedom is also PTP. These superoperators form the class of *Completely Positive Trace Preserving* (CPTP) maps [51]. It can be proven that CPTP maps coincide with those evolutions arising from unitary interactions with an environment and partial trace on the system of interest [147]. This mathematical phenomenon property has

[3] More precisely, Eq. (4.6) implies Eq. (4.5). Viceversa, when multi-time probabilities are not accessible beyond the two-point conditional $P(i, t_k | j, t_l)$, any dynamics satisfying (4.5) can be represented by a Markovian process (4.6). For a detailed analysis, see [53].

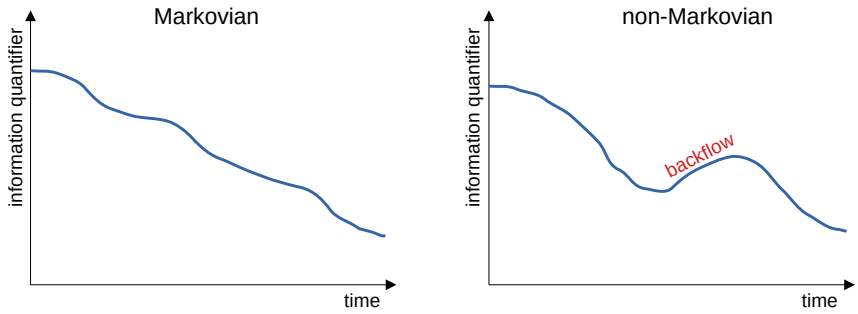

Figure 4.1: Pictorial representation of Markovian and non-Markovian dynamics, and the witnessing of non-Markovian effects. The typical approach consists in finding an *information quantifier* that decreases under stochastic maps (CPTP maps in the quantum case). According to Eq. (4.5), such quantifier decreases monotonically if the evolution is Markovian. A non-monotonic behaviour of such quantity, i.e. a *backflow* is therefore a witness of non-Markovianity.

no counterpart in the classical domain, where it exists no gap between PTP and CPTP maps. Accordingly, a quantum evolution is a family of CPTP maps parametrized by time $\mathcal{T} := \{\mathcal{T}^{(t,0)} | 0 \leq t\}$ acting on density matrices, giving rise to the evolution $\mathcal{T}^{(t,0)}[\rho(0)] = \rho(t)$.
Although the definition of Markovianity in quantum dynamics is debated [140], in analogy to Eq. (4.5), in this manuscript we adopt the canonical definition of quantum Markovianity as CP-divisibility [53], i.e., $\mathcal{T}$ is said to be Markovian if for any partition of the interval $[0, t]$, it can be written as the composition of intermediate maps that are CPTP.

4.1.2 Non-Markovianity detection

The complete characterization of non-Markovianity passes through the classification of the possible information backflow phenomena that these evolutions can offer. This task is at the core of the non-Markovian witnessing problem [53, 54]. More specifically, one needs to find information quantifiers and specific initializations of the system (which may include ancillas) that are able to signal the non-Markovian nature of the evolution via revivals of an otherwise monotonically decaying information, see Fig. 4.1. Different quantifiers have been considered in this context, such as for state discrimination [55, 148], channel capacity [149], volume of accessible states [150] and correlations [151, 152]. This approach allows to understand how we can gain benefits from non-Markovian evolutions and, at the same time, it is committed

to test the correspondence between its phenomenological and mathematical definition.
The most typical quantifier canonically used in the community is the trace distance. In the classical case [4], it is given by

$$D_{\mathrm{Tr}}(\boldsymbol{p},\boldsymbol{q}) = |\boldsymbol{p}-\boldsymbol{q}| = \sum_i |p_i - q_i| \ , \tag{4.7}$$

where $\boldsymbol{p}$ and $\boldsymbol{q}$ are two classical probability distributions. Such distance quantifies the distinguishability between the states $\boldsymbol{p}$ and $\boldsymbol{q}$, as the error probability of distinguishing the two in a single-shot measurement is lower bounded by $P_{\mathrm{err}} = \frac{2-|\boldsymbol{p}-\boldsymbol{q}|}{4}$ [52]. Moreover, this quantity decreases under physical maps, leading to a monotonic decrease for Markovian evolutions [53]. Hence, one consequence associated to the loss of information due to Markovianity is a continuous decrease in the ability of an agent to discriminate between any two states.
Still, the effects linked to Markovianity go well beyond what can be quantified by the trace distance alone. In the following we study an alternative quantifier: the Fisher information distance.

4.2 Background: Fisher metric and Fisher Information

The Fisher distance can be defined from its value between two infinitesimally close points. That is, given a small perturbation $|\boldsymbol{d}| \ll 1$ it reads

$$D^2_{\mathrm{Fish}}(\boldsymbol{p},\boldsymbol{p}+\boldsymbol{d}) \simeq \langle \boldsymbol{d},\boldsymbol{d}\rangle_{\boldsymbol{p}} := \frac{1}{2}\sum_i \frac{d_i^2}{p_i} \ , \tag{4.8}$$

where we define the Fisher scalar product $\langle \boldsymbol{a},\boldsymbol{b}\rangle_{\boldsymbol{p}} := \sum(a_i b_i)/(2p_i)$. Notice that such metric coincides with the differential of the square root of $\boldsymbol{p}$. That is, at leading order,

$$D^2_{\mathrm{Fish}}(\boldsymbol{p},\boldsymbol{p}+\boldsymbol{d}) = 2|\sqrt{\boldsymbol{p}} - \sqrt{\boldsymbol{p}+\boldsymbol{d}}|^2 \ . \tag{4.9}$$

This means that the Fisher metric maps the set of states from the symplex $\mathcal{S}(\mathbb{R}^N)$ to a sphere sector, via the change of variable $\mathbf{y} = \sqrt{\mathbf{p}}$ [56,130]. Accordingly, integrating Eq. (C.3) over the space of states one can obtain the general expression of the distance $D_{\mathrm{Fish}}(\boldsymbol{p},\boldsymbol{q}) = \sqrt{2}\arccos\left(\sqrt{\boldsymbol{p}}\cdot\sqrt{\boldsymbol{q}}\right)$ [56, 153].

[4] Similarly, in the quantum case, $D_{\mathrm{Tr}}(\rho,\sigma) := \mathrm{Tr}[|\rho-\sigma|]$.

The Fisher distance between two infinitesimally close points is sometimes called *Fisher Information* [5]. Besides its interesting mathematical properties, this quantity has numerous operational interpretations: in metrology it is used to derive the Cramér-Rao bound [76, 154], a fundamental limit on the precision with which a parameter can be estimated; it quantifies the asymptotic distinguishability between multiple copies of two states (Chernoff bound [155]); moreover, it also arises as the infinitesimal expansion of the relative entropy [156] (in fact, it can be shown that any f-divergence locally behaves as the Fisher information [157]).
The relation between Fisher metric and Markovianity has been previously partially analyzed [158]. In fact, a key characterisation of the Fisher metric is given by the Chentsov's theorem: this says that the Fisher information is the unique Riemannian metric that contracts under the action of all stochastic maps [56, 57, 159].
This is the starting point of our work in Ref. [7]. In particular, we study whether this strong relation between stochasticity and contractivity of the Fisher information can be reversed. That is, is it true that a map is Markovian *if and only if* it contracts monotonically the Fisher information? What are the operational consequences of the contraction/dilation of such metric?

Remark In the following we will concentrate mostly on classical dynamics for simplicity. At the same time, all the discussion and results can be generalized to the quantum scenario. This requires, besides, the identification of Markovian and non-Markovian dynamics on quantum states, also the generalisation of the Fisher distance. By generalizing Chentsov's theorem, Petz [160] found that, whereas the Fisher information distance is uniquely defined for classical systems, the presence of non commutative observables allows only for the definition of a family of Fisher information in the quantum case. In Appendix C we quickly review the different Quantum Fisher distances and the resulting Quantum Fisher informations. For the following discussion and results, it is enough to stress that although not uniquely defined, when restricted to the semiclassical limit case of quantum states that commute, all Quantum Fisher distances collapse onto the same classical Fisher distance.

[5] More precisely, the Fisher Information is the parametric derivative of the Fisher distance, when the state $\boldsymbol{p}_\theta$ is parametrized by some variable θ [130]. That is, the Fisher Information is equal to

$$\lim_{\delta\theta\to 0} \frac{D^2_{\mathrm{Fish}}(\boldsymbol{p}_\theta, \boldsymbol{p}_{\theta+\delta\theta})}{\delta\theta^2} \,.$$

4.3 Rate matrix decomposition of Markovian evolutions

From its definition (4.4), it is easy to verify that $T^{(t,s)}$ still satisfies $\sum_i T^{(t,s)}_{ij} = 1\ \forall j$, which can be equivalently rewritten as

$$T^{(t,s)}_{jj} = 1 - \sum_{i\neq j} T^{(t,s)}_{ij} \quad \forall j \ , \tag{4.10}$$

and corresponds to the requirement that the dynamics preserves the normalisation. In the limit of infinitesimal time-steps, a smooth evolution $T^{(t,0)}$ is generated by the rate matrix

$$R^{(t)} \equiv \lim_{\delta t\to 0} \frac{T^{(t+\delta t,t)} - \mathbb{1}}{\delta t} \ , \tag{4.11}$$

such that $\frac{\mathrm{d}}{\mathrm{d}t}T^{(t,0)} = R^{(t)} \circ T^{(t,0)}$. Thanks to the condition in Eq. (4.10) (see also Eq. (4.13a) below), one can always decompose $R^{(t)}$ as

$$R^{(t)} = \sum_{i\neq j} a^{(t)}_{i\leftarrow j} \left(|i\rangle\langle j| - |j\rangle\langle j|\right) \ , \tag{4.12}$$

where $a^{(t)}_{i\leftarrow j}$ are real coefficients called *rates*.
Since the composition of two Markovian evolutions is again Markovian, one can check the stochastic-divisibility of a channel by studying maps of the form $T^{(t+\delta t,t)} \approx \mathbb{1} + \delta t\, R^{(t)}$ for all possible times. is, we require $T^{(t+\delta t,t)}$ to be stochastic, imposing on the rate matrix, via Eq. (4.3),

$$\sum_i (\delta_{ij} + \delta t R^{(t)}_{i,j}) = 1 + \delta t \sum_i R^{(t)}_{i,j} = 1 \qquad \forall j, \tag{4.13a}$$

$$\delta_{ij} + \delta t R^{(t)}_{i,j} \geq 0 \qquad \forall\ i,j \ \ , \tag{4.13b}$$

where δ_{ij} denotes the Kronecker delta. The first conditions implies that $\sum_i R^{(t)}_{i,j} = 0$, which leads to the canonical decomposition (4.12). If we now focus on the second condition in Eq. (4.13) we see that $R^{(t)}_{i,j} \geq 0$ whenever $i \neq j$. In the parametrisation above this means that $a^{(t)}_{i\leftarrow j} \geq 0$. It follows that the rate matrices generating stochastic evolutions form a cone, whose elements are matrices of the form (4.12) having all the rates $a_{i\leftarrow j}$ positive. Since a smooth evolution T is Markovian if and only if $T^{(t+\delta t,t)}$ is stochastic $\forall t$, we can give the following definition:

Definition 1. A smooth evolution is Markovian if and only if for all times t the rates $a^{(t)}_{i\leftarrow j}$ are positive for all pairs of microstates $i \neq j$.

Quantum master equations As for classical dynamics, it is possible to derive a canonical decomposition for the local time-generator of a CPTP map. That is, given a smooth CPTP-divisible family $\mathcal{T}^{(t,0)}$ of superoperators one can define

$$\mathcal{L}^{(t)} := \lim_{\delta t\to 0} \frac{\mathcal{T}^{(t+\delta t,t)} - \mathbb{1}}{\delta t} \ , \quad \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{T}^{(t,0)} = \mathcal{L}^{(t)} \circ \mathcal{T}^{(t,0)} \ . \tag{4.14}$$

The "quantum rate matrix" $\mathcal{L}^{(t)}$ is typically called *lindbladian*, as its characterization was given by Lindblad [161], and separately by Gorini-Kossakowski-Sudarshan [162]. The resulting canonical decomposition is of the form

$$\mathcal{L}[\rho] = \sum_i \gamma_i \left(A_i \rho A_i^\dagger - \frac{1}{2}\{A_i^\dagger A_i, \rho\} \right) \ , \tag{4.15}$$

where the operators A_i are traceless and orthonormal $\mathrm{Tr}[A_i A_j^\dagger] = \delta_{ij}$, while the coefficients γ_i are real and correspond to the rates of the quantum master equation.
In the most general time-dependent case, the operators $A_i(t)$ are also time-dependent, as well as the rates $\gamma_i(t)$. Albeit less trivial, it was proven [163] that, as in the classical case (cf. Def. 1), the Markovianity (defined as CPTP divisibility) of smooth quantum evolutions becomes equivalent to the positivity of the rates of the master equation at all times $\gamma_i(t) \geq 0$.

4.4 Contractivity of the Fisher metric and the detection of non-Markovianity

In the following we present the main results obtained in Ref. [7]. Notice that for simplicity we only state the main results and partially explain the proofs for the case of classical dynamics. The full proofs can be found in [7], as well as the generalizations to quantum dynamics (cf. also App. C).
From now on we will omit time-dependency when no confusion can arise. The

Fisher distance between any two points decreases under the action of stochastic maps. This directly implies that under Markovian dynamics the Fisher metric contracts continuously. Indeed, simple algebra yields, for infinitesimal $\boldsymbol{d}$,

$$
\begin{aligned}
2\,\frac{\mathrm{d}}{\mathrm{d}t} D^2_{\mathrm{Fish}}(\boldsymbol{p}, \boldsymbol{p}+\boldsymbol{d}) &= \sum_i \left(\frac{2 d_i \dot{d}_i}{p_i} - \frac{d_i^2}{p_i^2}\,\dot{p}_i \right) = \\
&= \sum_{i,j} \left(\frac{2 d_i (a_{i\leftarrow j} d_j - a_{j\leftarrow i} d_i)}{p_i} - \frac{d_i^2}{p_i^2} (a_{i\leftarrow j} p_j - a_{j\leftarrow i} p_i) \right) = \\
&= \sum_{i,j} \left(\frac{2 d_i d_j a_{i\leftarrow j}}{p_i} - \frac{d_i^2}{p_i} a_{j\leftarrow i} - \frac{d_i^2}{p_i^2} p_j a_{i\leftarrow j} \right) = \\
&= \sum_{i,j} a_{i\leftarrow j} \left(\frac{2 d_i d_j}{p_i} - \frac{d_j^2}{p_j} - \frac{d_i^2}{p_i^2} p_j \right) = \\
&= -\sum_{i\neq j} a_{i\leftarrow j} \left(\frac{d_i}{p_i} - \frac{d_j}{p_j} \right)^2 p_j \leq 0 \,. \qquad (4.16)
\end{aligned}
$$

This can be rewritten as

$$
\frac{\mathrm{d}}{\mathrm{d}t} D^2_{\mathrm{Fish}}(\boldsymbol{p}, \boldsymbol{p}+\boldsymbol{d}) = -\sum_{i\neq j} a_{i\leftarrow j} I_{i\leftarrow j} \leq 0 \,, \qquad (4.17)
$$

where we implicitly defined

$$
I_{i\leftarrow j} := \frac{1}{2} \left(\frac{d_i}{p_i} - \frac{d_j}{p_j} \right)^2 p_j \qquad (4.18)
$$

as the *Fisher information flow* associated to the rate $a_{i\leftarrow j}$. These are positive objects, so that the contraction of the Fisher metric is directly associated to the positivity of the rates $a_{i\leftarrow j}$.

Now suppose that for some time t there is a negative rate $a^{(t)}_{\tilde{i}\leftarrow\tilde{j}} < 0$ (i.e., the evolution is non-Markovian). Is this sufficient to reverse the contraction of the Fisher information? The positive answer is given by the following

Theorem 1. *A smooth evolution $T^{(t,0)}$ is Markovian if and only if it induces a decrease in Fisher distance between any two points in $\mathcal{S}(\mathbb{R}^N)$ at all times, i.e.*

$$
\begin{aligned}
T^{(t,0)} \text{ is Markovian } &\Leftrightarrow D_{\mathrm{Fish}}(T^{(t+\delta t,t)}[\boldsymbol{p}], T^{(t+\delta t,t)}[\boldsymbol{q}]) \leq D_{\mathrm{Fish}}(\boldsymbol{p}, \boldsymbol{q}) \\
&\forall t,\ \forall \boldsymbol{p}, \boldsymbol{q} \in \mathcal{S}(\mathbb{R}^N) \,. \qquad (4.19)
\end{aligned}
$$

Note that applying an intermediate map $T^{(t,s)}$ to all the points in the state space $\mathcal{S}(\mathbb{R}^N)$ is in general not physically defined, as only points in the image of $T^{(s,0)}$ are guaranteed to be physical states after applying $T^{(t,s)}$ (cf. Eq. (4.4)). Still, by considering the infinitesimal evolution $T^{(t+\delta t,t)} \simeq \mathbb{1} + \delta t R^{(t)}$, each point *in the interior* of $\mathcal{S}(\mathbb{R}^N)$ is guaranteed to be a physical state after evolving for a δt small enough. This allows $\frac{\mathrm{d}}{\mathrm{d}t} D_{\mathrm{Fish}}(\boldsymbol{p},\boldsymbol{q})|_t$ to be well defined at all times for all points in the interior of $\mathcal{S}(\mathbb{R}^N)$.

Proof. As the $\Rightarrow$ implication in Eq. (4.19) is trivial from the monotonicity property of the Fisher distance, only the proof of the opposite $\Leftarrow$ is needed, that is, that the contraction at all points and times of the Fisher distance implies Markovianity of the evolution. For that, suppose that the first instance of non-Markovianity happens between time t and $t+\delta t$. In order to prove the statement it is sufficient to consider the Fisher distance between any two infinitesimally close points $\boldsymbol{p}$ and $\boldsymbol{p}+\boldsymbol{d}$, with $|\boldsymbol{d}| \ll 1$, evolving according to $T^{(t+\delta t,t)}$. Then, for any negative rate $a^{(t)}_{\tilde{i}\leftarrow\tilde{j}} < 0$, one can find a point $\boldsymbol{p}$ and a perturbation $\boldsymbol{d}$ such that $\frac{\mathrm{d}}{\mathrm{d}t} D^2_{\mathrm{Fish}}(\boldsymbol{p},\boldsymbol{p}+\boldsymbol{d}) > 0$. In fact, assume without loss of generality that $a_{1\leftarrow 2} < 0$ and consider $\boldsymbol{p}$ and $\boldsymbol{d}$ of the form

$$\boldsymbol{p} = \begin{pmatrix} \mathcal{O}(\varepsilon) \\ 1-\varepsilon \\ \mathcal{O}(\varepsilon) \\ \vdots \\ \mathcal{O}(\varepsilon)\,, \end{pmatrix}, \quad \boldsymbol{d} = \begin{pmatrix} \mathcal{O}(\varepsilon) \\ \mathcal{O}(\varepsilon) \\ \mathcal{O}(\varepsilon^2) \\ \vdots \\ \mathcal{O}(\varepsilon^2) \end{pmatrix}, \tag{4.20}$$

where ε is an arbitrary small number and we assume that the vectors $\boldsymbol{p}$ and $\boldsymbol{p}+\boldsymbol{d}$ are properly normalised. Notice that it is always possible to choose $\boldsymbol{p}$ and $\boldsymbol{p}+\boldsymbol{d}$ in the interior of $\mathcal{S}(\mathbb{R}^N)$ (i.e., with strictly positive components). Inserting this expression in Eq. (4.17) we find that the only term of order $\mathcal{O}(1)$ comes from setting $i=1, j=2$ in the sum (4.17). That is, at leading order,

$$\frac{\mathrm{d}}{\mathrm{d}t} D^2_{\mathrm{Fish}}(\boldsymbol{p},\boldsymbol{p}+\boldsymbol{d}) = -a_{1\leftarrow 2}\,\frac{d_1^2}{p_1^2} + \mathcal{O}(\varepsilon)\,. \tag{4.21}$$

Since $a_{1\leftarrow 2} < 0$ we can always find a ε small enough so that this quantity is strictly positive. Hence, for any non-Markovian dynamics there always exists $\boldsymbol{p}$ and $\boldsymbol{d}$ such that $D^2_{\mathrm{Fish}}(\boldsymbol{p},\boldsymbol{p}+\boldsymbol{d})$ locally increases, proving the Theorem. □

Theorem 1 can be seen as a completion of Chentsov's Theorem, as it implies that not only the Fisher information decreases under Markovian evolutions, but also that an evolution contracting the Fisher distance between any two points has to be Markovian. The proof generalizes to the case of quantum

dynamics in the canonical CP-divisibility framework [53], by considering a copy of the system on which the dynamics acts trivially, i.e., the evolution is given by $\mathcal{T}^{(t,0)}\otimes\mathbb{1}_N$ and the set of states is considered on the global bipartition (see details in [7]).
Interestingly, a similar theorem cannot hold for the trace-distance as one can explicitly construct non-Markovian evolutions that monotonically contract $D_{\mathrm{Tr}}(\boldsymbol{p},\boldsymbol{q})$ for any two points ([7, 53]). This corroborates the interpretation of the Fisher metric as the canonical distance whose contractivity identifies stochastic-divisible maps. Still, the argument that lead to Thm. 1 has a shortcoming: even if non-Markovianity implies the dilation of the Fisher information, this is not sufficient to produce an operational witness. In fact, assume that at time t there is a local dilation of Fisher distance for two points close to $\tilde{\boldsymbol{p}}$. In order to observe it, one would need to initialise the system in the state $\boldsymbol{p}(0)=(T^{(t,0)})^{-1}\tilde{\boldsymbol{p}}$. If $\tilde{\boldsymbol{p}}$ is outside the image of $T^{(t,0)}$, this cannot be achieved physically. That is, the drawback of this approach is that the Fisher metric is point-dependent, and the witnessing point might be excluded by the dynamics $T^{(t,0)}$. On the other hand, the trace-distance is translational invariant, i.e., $D_{\mathrm{Tr}}(\boldsymbol{p},\boldsymbol{q})=|\boldsymbol{d}|$ where $\boldsymbol{d}=\boldsymbol{p}-\boldsymbol{q}$. Then, as soon as any two points $\boldsymbol{p}$ and $\boldsymbol{q}$ are increasing their trace-distance, in order to present an operational witness it is sufficient to consider a point $\boldsymbol{r}$ in the interior of the image of $T^{(t,0)}$ and ε small enough for $\boldsymbol{r}_\varepsilon=\boldsymbol{r}+\varepsilon(\boldsymbol{p}-\boldsymbol{q})$ to be in the image as well. Then, $D_{\mathrm{Tr}}(\boldsymbol{r},\boldsymbol{r}_\varepsilon)=\varepsilon D_{\mathrm{Tr}}(\boldsymbol{p},\boldsymbol{q})$ and this increases by assumption. Moreover, if one adds a finite number of ancillary degrees of freedom on which the dynamics acts trivially, one can always find such two points for any non-Markovian evolution (see [55] and [7]). A similar property does not hold for the Fisher distance, as it lacks translation invariance. More specifically,

Theorem 2. *No finite number n of copies of the channel $T^{(t,0)}$ nor ancillary degrees of freedom of any dimension M is enough to witness all non-Markovian evolutions via revivals of the Fisher distance between two initially prepared states.*

Specifically, given n copies of the system, and an ancilla with arbitrary dimension M, the state space will be $\mathcal{S}(\mathbb{R}^{N^{\otimes n}}\otimes\mathbb{R}^M)$ and the dynamics acting on it $\bar{T}^{(t,0)}=T^{(t,0)^{\otimes n}}\otimes\mathbb{1}_M$.

Proof. We constructively provide, for any n and M, a counterexample in which all the states in the image of $\bar{T}^{(t,0)}$ continue decreasing their Fisher distance between time t and $t+\delta t$, even if $T^{(t+\delta t,t)}\simeq\mathbb{1}+\delta t R^{(t)}$ is non-stochastic. Here we provide the proof for the single copy case, deferring the multiple cases one to the paper [7], which has essentially the same treatment. The map we consider is then given by $\bar{T}^{(t,0)}=T^{(t,0)}\otimes\mathbb{1}_M$ and the rate matrix by $\bar{R}^{(t)}=\frac{\mathrm{d}}{\mathrm{d}t}\bar{T}^{(t,0)}=R^{(t)}\otimes\mathbb{1}_M$. Suppose now that there is a unique negative rate $a_{\tilde{i}\leftarrow\tilde{j}}$

and that the image of $T^{(t,0)}$ is contained to a small ball around an appropriate vector $\boldsymbol{\pi}$ (e.g., by a map of the form $T^{(t,0)}[\boldsymbol{p}] = \boldsymbol{\pi}(1-\varepsilon)+\varepsilon\boldsymbol{p}$). Attach at time 0 an arbitrary ancilla, so that the initial state is given by $\boldsymbol{p}(0) \in \mathcal{S}(\mathbb{R}^N \otimes \mathbb{R}^M)$, and define $\boldsymbol{w}$ to be its reduced marginal on $\mathbb{R}^M$, whereas the dynamics is given by $\bar{T}^{(t,0)} = T^{(t,0)} \otimes \mathbb{1}_M$. Then, the state at time t will be ε-close to

$$\boldsymbol{p}(t) \sim \boldsymbol{\pi} \otimes \boldsymbol{w} + \mathcal{O}(\varepsilon)\,. \tag{4.22}$$

Notice also that the rate matrix $\bar{R}^{(t)} = \frac{\mathrm{d}}{\mathrm{d}t}\bar{T}^{(t,0)} = R^{(t)} \otimes \mathbb{1}_M$ has the following coordinate expression

$$[\bar{R}]_{ij,\alpha\beta} = R_{ij}\delta_{\alpha\beta} \quad i,j \in \{1,\dots,N\} \quad \alpha,\beta \in \{1,\dots,M\}\ , \tag{4.23}$$

so that the rates are simply given by $a_{i\alpha\leftarrow j\beta} = a_{i\leftarrow j}\delta_{\alpha\beta}$. In this scenario, the evolution of the Fisher distance (as expressed in Eq. (4.17)) becomes

$$-\sum_{i\neq j,\alpha} a_{i\leftarrow j}\left(\frac{d_{i\alpha}}{p_{i\alpha}} - \frac{d_{j\alpha}}{p_{j\alpha}}\right)^2 p_{j\alpha} + \mathcal{O}(\varepsilon) \quad \text{with } p_{j\alpha} = \pi_j w_\alpha\,. \tag{4.24}$$

Again, consider the case in which at time t a single rate becomes negative, for definiteness say $a_{1\leftarrow 2} < 0$. Then, it is sufficient that $a_{2\leftarrow 1}\pi_1 > |a_{1\leftarrow 2}|\pi_2$ to see that the sum in Eq. (4.24) is strictly negative in the limit $\varepsilon \to 0$. Hence, even if the dynamics is non-Markovian, there is no increase in Fisher distance on the image of $T^{(t,0)}$, proving Theorem 2 for $n = 1$. □

In the same way, even using multiple copies of the channel does not help finding a witness. The proof easily generalises to $n \geq 2$. Moreover, it should be noticed that the condition of finite copies in Thm. 2 cannot be dropped: in fact, in the limit of infinite copies, one can perform full tomography of the evolution, allowing to reconstruct its action also on points outside of the image of $T^{(t,0)}$.

Despite the above "no-go" Theorem, we can introduce an operational non-Markovianity witness which does not require additional copies of the channel, but only some specific post-processing of the states. More specifically, the following technical theorem holds

Theorem 3. *For any state $\boldsymbol{p}$ and perturbation $\boldsymbol{d}$ on $\mathcal{S}(\mathbb{R}^N \otimes \mathbb{R}^M)$ (where we admit an M-dimensional ancilla), it is possible to implement a class of transformations $F_{\boldsymbol{d}}$ depending on $\boldsymbol{d}$ and on $T^{(t,0)}$ that witness non-Markovianity at*

time t. That is, if $T^{(t+\delta t,t)}$ is stochastic, for any choice of $\boldsymbol{d}$, one has that

$$D_{\text{Fish}}(F_{\boldsymbol{d}} \circ T^{(t+\delta t,0)}[\boldsymbol{p}], F_{\boldsymbol{d}} \circ T^{(t+\delta t,0)}[\boldsymbol{p}+\boldsymbol{d}]) \\ \leq D_{\text{Fish}}(F_{\boldsymbol{d}} \circ T^{(t,0)}[\boldsymbol{p}], F_{\boldsymbol{d}} \circ T^{(t,0)}[\boldsymbol{p}+\boldsymbol{d}]) \quad (4.25)$$

whereas in the presence of non-Markovianity (i.e., for $T^{(t+\delta t,t)}$ not stochastic) there exists at least one $\boldsymbol{d}$ for which the inequality is reversed (i.e., the Fisher distance of the post-processed states increases). Moreover, for classical systems $M = 2$ is enough to witness in this way all non-Markovian evolutions (while for quantum systems one needs in general $M = N + 1$ [6]).

The detailed proof is technical and provided in [7]. Theorem 3 ensures that any break of stochastic-divisibility in the interval $[t, t+\delta t]$ can be operationally witnessed via backflow of Fisher information between states that undergo the transformation $F_{\boldsymbol{d}}$ before being measured. That is, from the operational point of view, the above theorem implies that in such case if one is interested in the Fisher distance between the post-processed states, via $F_{\boldsymbol{d}}$ [7] of $\boldsymbol{p}$ and $\boldsymbol{q}$ being as large as possible, if "non-Markovianity occurs" at time t, it is better to wait an additional time δt before measuring the system. The specific form of $F_{\boldsymbol{d}}$ [7] is *ad hoc* to prove the Theorem, which should be considered as a proof of principle of the possibility of witnessing non-Markovianity through post-processing.

4.5 Bayesian retrodiction and the meaning of information backflow

In most of the literature about non-Markovianity backflows of information are considered by studying states $\boldsymbol{p}(t)$ at time t [53, 54], while the question about backflows of information *about the initial state* $\boldsymbol{p}(0)$ remains instead largely unexplored. Even if one can argue that the invertibility of the dynamics preserves the information about the initial conditions (as these can be recovered as $\boldsymbol{p}(0) = {T^{(t,0)}}^{-1}[\boldsymbol{p}(t)]$), actually retrieving the initial state from $\boldsymbol{p}(t)$ requires full tomography both of the state and of the channel, and a post-processing of such data which cannot be performed physically in a single-shot scenario.

[6] It should be noticed that in the quantum case it is never possible to detect all non-Markovian evolutions by using ancillas that have dimension less than N, as the CPTP condition for superoperators needs (at least) an N-dimensional ancilla to be detected [164].

[7] Notice that $F_{\boldsymbol{d}}$ might be thought as an unavoidable part of the measurement apparatus.

Conversely, we consider here a Bayesian inversion of $\boldsymbol{p}(t)$ that allows us to compare it with the initial state through a physically implementable transformation. In particular, define the *prior* as a vector $\boldsymbol{\pi}$ representing our knowledge about the system at time 0. Given an evolution $T^{(t,0)}$, the Bayes-recovery map is defined as

$$\hat{T}_t = J_{\boldsymbol{\pi}} \circ T_t^{\intercal} \circ J_{T[\boldsymbol{\pi}]}^{-1} \tag{4.26}$$

where we use the shorthand notation $T_t := T^{(t,0)}$, and we introduced the map $J_{\boldsymbol{p}}$, corresponding to the diagonal operator that multiplies each component of a vector by the corresponding component of $\boldsymbol{p}$, i.e. $[J_{\boldsymbol{p}}]_{ij} = \delta_{ij} p_j$.
The map $\hat{T}_t$ is stochastic (i.e., *physically implementable*) and represents a recovery of the state via statistical retrodiction [165,166]. Indeed if one identifies the coordinates of the prior with the corresponding probability, i.e.,

$$\pi_i \equiv P(i,0) \ , \tag{4.27}$$

and the components of the maps with the corresponding transitions, i.e., $(T_t)_{ij}$ $\equiv P(i,t|j,0)$, it is easy to verify that

$$(\hat{T}_t)_{ij} \equiv P(i,0|j,t) \tag{4.28}$$

satisfies the Bayes rule. It should also be noticed that $\hat{T}_t$ perfectly recovers the prior at all times ($\boldsymbol{\pi} = \hat{T}_t \circ T_t[\boldsymbol{\pi}]$) (cf. Fig. 4.2).
This channel allows us to study how much information is stored in the evolved state $\boldsymbol{p}(t)$ about its initial conditions. In particular, we can compare the distance between $\boldsymbol{p}(0)$ and the retrodicted state

$$\hat{\boldsymbol{p}}(t) := \hat{T}_t[\boldsymbol{p}(t)] = \hat{T}_t \circ T_t[\boldsymbol{p}(0)] \ . \tag{4.29}$$

We also assume that the prior contains some knowledge on the initial conditions, so that we can write $\boldsymbol{p}(0) = \boldsymbol{\pi} + \boldsymbol{d}$, for some small $|\boldsymbol{d}| \ll 1$. Moreover, since $\boldsymbol{\pi}$ is perfectly recovered, we also have that $\hat{\boldsymbol{p}}(t) = \boldsymbol{\pi} + \hat{\boldsymbol{d}}(t)$. Then, the Fisher information between $\boldsymbol{p}(0)$ and $\hat{\boldsymbol{p}}(t)$ reads

$$D^2_{\text{Fish}}(\boldsymbol{p}(0), \hat{\boldsymbol{p}}(t)) \simeq \left\langle \boldsymbol{d} - \hat{\boldsymbol{d}}(t), \boldsymbol{d} - \hat{\boldsymbol{d}}(t) \right\rangle_{\boldsymbol{\pi}} \ . \tag{4.30}$$

Interestingly, this object is directly connected to the Fisher information at time t:

Theorem 4. *The contractivity of the Fisher information at time t is in one-to-one correspondence with the expansivity of Eq.* (4.30)*. That is,*

$$\frac{\mathrm{d}}{\mathrm{d}t} D^2_{\mathrm{Fish}}(\boldsymbol{p}(0), \hat{\boldsymbol{p}}(t)) = \frac{\mathrm{d}}{\mathrm{d}t} \left\langle \boldsymbol{d} - \hat{\boldsymbol{d}}(t), \boldsymbol{d} - \hat{\boldsymbol{d}}(t) \right\rangle_{\boldsymbol{\pi}} \geq 0 \tag{4.31}$$

if and only if the Fisher information contracts for any two points in the vicinity of $T_t[\pi]$.

Proof. The main ingredient in the proof of this theorem is given by the following identities

$$\langle \boldsymbol{d}, \hat{T}_t T_t[\boldsymbol{d}] \rangle_{\boldsymbol{\pi}} = \langle T_t[\boldsymbol{d}], T_t[\boldsymbol{d}] \rangle_{T_t[\boldsymbol{\pi}]} = \langle \hat{T}_t T_t[\boldsymbol{d}], \boldsymbol{d} \rangle_{\boldsymbol{\pi}}, \tag{4.32}$$

which can verified by directly substituting $\hat{T}_t$ with its definition in Eq. (4.26). One can read from these equalities the following two facts: first, $\hat{T}_t$ can be used to put in relation the Fisher information at time 0 and at time t; secondly, $\hat{T}_t T_t$ is self-adjoint with respect to $\langle \bullet, \bullet \rangle_{\boldsymbol{\pi}}$. This allows to rewrite Eq. (4.31) as

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\langle \boldsymbol{d}, (\mathbb{1} - \hat{T}_t T_t)^2 [\boldsymbol{d}] \right\rangle_{\boldsymbol{\pi}} = -2 \left\langle \boldsymbol{d}, (\mathbb{1} - \hat{T}_t T_t) \frac{\mathrm{d}}{\mathrm{d}t} \hat{T}_t T_t [\boldsymbol{d}] \right\rangle_{\boldsymbol{\pi}}. \tag{4.33}$$

First notice that $(\mathbb{1} - \hat{T}_t T_t)$ is positive definite. In fact, this can be seen from

$$\langle \boldsymbol{d}, \hat{T}_t T_t[\boldsymbol{d}] \rangle_{\boldsymbol{\pi}} = \langle T_t[\boldsymbol{d}], T_t[\boldsymbol{d}] \rangle_{T_t[\boldsymbol{\pi}]} \leq \langle \boldsymbol{d}, \boldsymbol{d} \rangle_{\boldsymbol{\pi}}\ , \tag{4.34}$$

where the last inequality follows from the contractivity of the Fisher information. Moreover, we also have that $\langle \boldsymbol{d}, \frac{\mathrm{d}}{\mathrm{d}t} \hat{T}_t T_t[\boldsymbol{d}] \rangle_{\boldsymbol{\pi}} = \frac{\mathrm{d}}{\mathrm{d}t} \langle T_t[\boldsymbol{d}], T_t[\boldsymbol{d}] \rangle_{T_t[\boldsymbol{\pi}]}$, so that $-\frac{\mathrm{d}}{\mathrm{d}t} \hat{T}_t T_t$ is positive if and only if the Fisher contracts monotonically. If this is the case, we can use the fact that the product of two positive operators has positive spectrum. On the other hand, if the Fisher metric is expanding at time t, there exists an eigenvector $\tilde{\boldsymbol{d}}$ of $-\frac{\mathrm{d}}{\mathrm{d}t} \hat{T}_t T_t$ with negative eigenvalue $\lambda < 0$, so that

$$-2 \left\langle \tilde{\boldsymbol{d}}, (\mathbb{1} - \hat{T} T) \frac{\mathrm{d}}{\mathrm{d}t} \hat{T} T [\tilde{\boldsymbol{d}}] \right\rangle_{\boldsymbol{\pi}} = 2\lambda \left\langle \tilde{\boldsymbol{d}}, (\mathbb{1} - \hat{T} T)[\tilde{\boldsymbol{d}}] \right\rangle_{\boldsymbol{\pi}} < 0, \tag{4.35}$$

proving the claim. □

This theorem tells us that the ability of an agent of retrieving the initial state of the dynamics decreases under Markovian evolution. Moreover, in the case in which there are backflows in the Fisher information, non-Markovianity helps obtaining more information about the initial state (see Fig. 4.2).

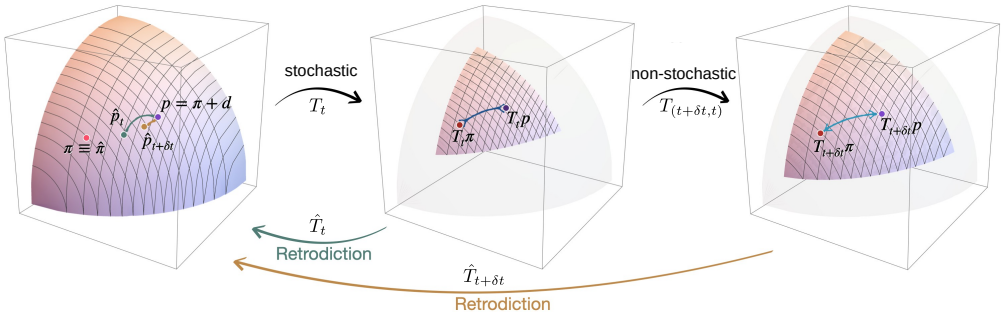

Figure 4.2: Representation of Theorem 4. Given a prior $\boldsymbol{\pi}$, the Bayes retrodiction map $\hat{T}_t$ is a physical (i.e. stochastic) map that recovers perfectly $\boldsymbol{\pi}$ when applied to $T_t[\boldsymbol{\pi}]$ (via Bayes rule), while in general other close-by points are not retrodicted perfectly. We prove that, when the evolution is Markovian (i.e. stochastic-divisible), the retrodicted state $\hat{\boldsymbol{p}} := \hat{T}_t T_t[\boldsymbol{p}]$ continuously gets further from the original value $\boldsymbol{p}$ (green retrodiction in the picture). If stochastic-divisibility is broken, it exist a state whose retrodiction improves, that is $\hat{\boldsymbol{p}}$ gets closer to $\boldsymbol{p}$, according to the Fisher distance (orange retrodiction).

4.6 Discussion

In Ref. [7] we characterized the relation between Markovianity (intended as stochastic divisibility), Fisher metric contractivity and information flow, both from the mathematical and operational point of view. We explored such connection to its very limits, and the resulting picture can be seen in Fig. 4.3: we showed that monotonous contractivity of the Fisher metric on the whole set of states and at all times, is mathematically equivalent to Markovianity (Thm. 1). As known, when the metric dilates locally inside the image of the evolution $\mathrm{Im}(T^{(t,0)})$, a backflow of Fisher information can be operationally witnessed. At the same time, non-Markovian evolutions might in general show Fisher metric dilations only outside the image of the evolution itself, regardless of the number of copies of the channel and ancillary degrees of freedom available (Thm. 2). To witness operationally non-Markovianity in such cases, one needs post-processing to be appended to the dynamics (Thm. 3). Finally, we showed that dilations of the Fisher metric between evolving states can be mapped to a backflow of information about the initial states by applying Bayesian retrodiction (Thm. 4).

Moreover, the results were presented for classical dynamics in order to keep the exposition clean. Despite that, we also show in [7] that the generalizations to the quantum dynamics scenario of our results can be easily worked out constructively, by using properties of the Quantum Fisher metric on diagonal subspaces (cf. Appendix C). At the same time, the non-uniqueness of the

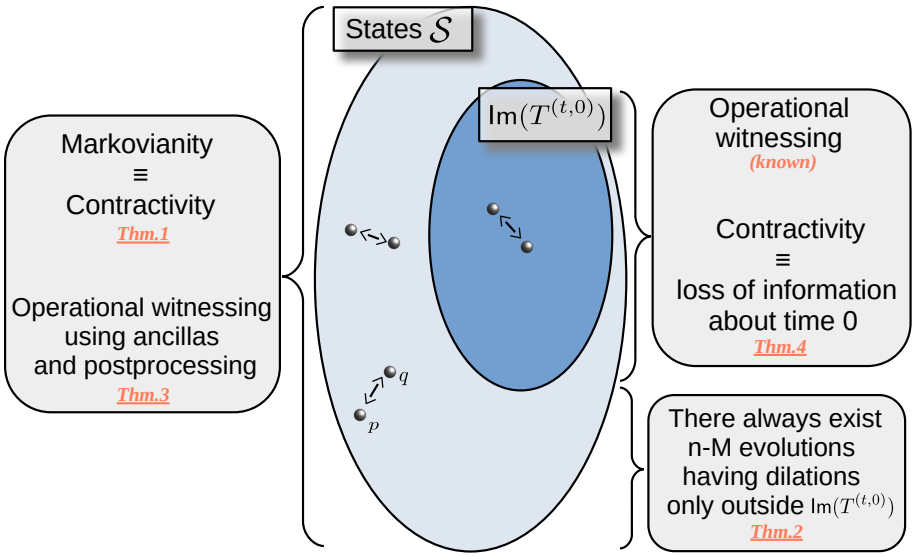

Figure 4.3: The main quantity analyzed in our work [7] is the rate of contraction/dilation of the Fisher metric, i.e., $\frac{\mathrm{d}}{\mathrm{d}t} D_{\mathrm{Fish}}(\boldsymbol{p}, \boldsymbol{q})$, with $|\boldsymbol{p} - \boldsymbol{q}| \ll 1$, when both $\boldsymbol{p}$ and $\boldsymbol{q}$ evolve according to the local intermediate map $T^{(t+\delta t,t)}$. We characterize the mathematical and operational meaning of the negativity/positivity of such rate, both inside and outside the image of the evolution $T^{(t,0)}$.

Quantum Fisher distance offers a rich landscape of possibilities when defining retrodiction, or simply in the computation of the rate of change of Fisher distance outside diagonal subspaces. We will investigate these issues in forthcoming work [66].
Finally, it would be of natural interest to investigate if this picture can be enlarged to the analysis of Markovianity and non-Markovianity based on multitime correlators [140], which is being studied in recent years.

Chapter 5

Network nonlocality with passive optics and single-photons

This chapter is based on Ref. [5]:
"P. Abiuso, T. Kriváchy, E.-C. Boghiu, M.-O. Renou, A. Pozas-Kerstjens and A. Acín, *Single-photon nonlocality in quantum networks*, Phys. Rev. Research **4**, L012041 (2022)"

5.1 Background: Bell nonlocality

Local hidden variables models cannot account for all the predictions of quantum theory. This was formalized in 1964 by J. S. Bell [58], and is now commonly termed *Bell nonlocality*, or simply *nonlocality* [59]. Nonlocality is a quantum property with no classical analogue displayed in the so-called Bell tests, defined by the statistics obtained when performing appropriate local measurements on a well-chosen entangled state. More precisely, the original *bipartite Bell scenario* consists in two parties, Alice and Bob (or A and B), performing an experiment in different, non-communicating laboratories. Alice (Bob) might choose different measurements, labelled by x (y), leading to outputs a (b), and the statistical description of their experiment's results is therefore given by the probability distribution $p(a,b|x,y)$. The absence of communication between the two is formalised as the *non-signalling* property of Alice's (Bob's) output being independent of Bob's (Alice's) input:

$$\sum_b p(a,b|x,y) := p(a|x,y) = p(a|x) \quad \text{(and viceversa)}. \tag{5.1}$$

The non-signalling property can be in principle guaranteed by making the measurement of the two parties space-like separated, and claiming consistency with special relativity. At the same time, such property does not imply the absence of correlations between A and B. For example, they could be allowed to share a classical source of randomness, which can be represented without loss of generality, as a real variable λ. More precisely, assuming no direct causal dependence of Alice from Bob or viceversa, the most general probability distribution arising from a Bell experiment involving only classical sources, is [59]

$$p_{\mathcal{L}}(a,b|x,y) = \int \mathrm{d}\lambda\ \mu(\lambda) p(a|x,\lambda) p(b|y,\lambda)\ . \tag{5.2}$$

Conditional probabilities of the form above form the so called *local set* $\mathcal{L}$. Similarly, if Alice and Bob can share a quantum state as a common source, the resulting non-signalling distribution is given according to the Born-rule

$$p_{\mathcal{Q}}(a,b|x,y) = \mathrm{Tr}[(M_A^{a|x} \otimes M_B^{b|y})\rho_{AB}]\ . \tag{5.3}$$

Perhaps surprisingly, Bell theorem [58] showed that the set $\mathcal{Q}$ strictly contains $\mathcal{L}$. That is, the correlations achievable between non-communicating parties sharing quantum states can be outside the local set, i.e. they can be *nonlocal*. This rules out any local hidden variable theory as a possible fundamental description of reality. In order to violate locality, Alice and Bob need to share an entangled state [59]. When the cardinality of inputs $\{x,y\}$ and outputs $\{a,b\}$ is finite, the local set $\mathcal{L}$ can be described as a finite-dimensional *polytope*, that is, a convex set with a finite number of vertices, which correspond to deterministic strategies employed by the parties [59]. The facets of the polytope correspond to hyperplanes which in turn define linear inequalities for any probability to be inside the local set. These are called *Bell inequalities*. The most famous Bell inequality was found in the simplest scenario of parties having binary inputs $x,y = 0,1$ and binary outputs $a,b = \pm 1$. Clauser, Horne, Shimony and Holt (CHSH) [167], showed that in any Bell experiment with classical sources, the following bound holds

$$\left|\langle A_0 B_0\rangle + \langle A_0 B_1\rangle + \langle A_1 B_0\rangle - \langle A_1 B_1\rangle\right| \leq 2\ , \tag{5.4}$$

where the correlators introduced $\langle A_x B_y\rangle := \sum_{a,b} ab\ p(a,b|x,y)$ are the average values of the outputs' product. Moreover using a maximally entangled state and appropriate measurements, they showed that in quantum mechanics can achieve a value of $2\sqrt{2}$ for the same quantity. Later, Tsirelson [168] proved that this is also the maximum value obtainable with quantum mechanics.

Beyond the fundamental motivation, nonlocality is relevant from an applied point of view, and violating Bell inequalities is the goal of many experiments manipulating quantum sources. In fact, correlations with no classical analogue are the main resource, e.g., for the security of device-independent protocols for quantum random number generation [169,170] and quantum key distribution [171–173], as well as self-testing of quantum states [174].

5.2 Background: Single-photon nonlocality

Bell tests have been performed in many different systems, from massive particles [175] to photons [176,177], and using many different degrees of freedom, such as electronic levels, polarization, orbital angular momentum or time bins. In most of these realizations the relevant degrees of freedom used to encode the entanglement are transmitted to each distant observer by a separate physical particle. We are interested in the question of whether *single-particle quantum states* can display nonlocal correlations with no classical analogue. In particular, we consider the question in the context of single-photon entanglement, that is, the state

$$|\psi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} + |10\rangle_{AB}), \tag{5.5}$$

obtained when sending a single photon into a balanced beamsplitter. Here $|01\rangle_{AB}$ (resp. $|10\rangle_{AB}$) represents the situation in which the photon is sent to the right party B (resp. the left party A). The resulting state therefore consists of only one photon and entanglement is encoded in the two optical spatial modes.

Is the state (5.5) nonlocal? This question has been intensively debated in the quantum foundations and quantum optics community, e.g. [178–188]. In principle, a positive answer is provided by the following simple argument [180–182]: the two optical modes can be transferred to the population of two energy levels of two distant massive particles. Single-photon entanglement is therefore mapped into two-particle entanglement and a Bell test can now be implemented. The question is much subtler when considering only optical means. To obtain a nonlocal behavior, the two observers need to use local *active measurements* involving local oscillators creating extra local photons [178, 179, 184, 186]: without these active measurements, measuring the information content of the state (5.5) would allow the observers to deduce if they received the photon sent by the source, destroying the indeterminacy in the photon path, i.e. the coherences in (5.5). Then, the statistics become classically simulable. One is therefore tempted to conclude that the observation of

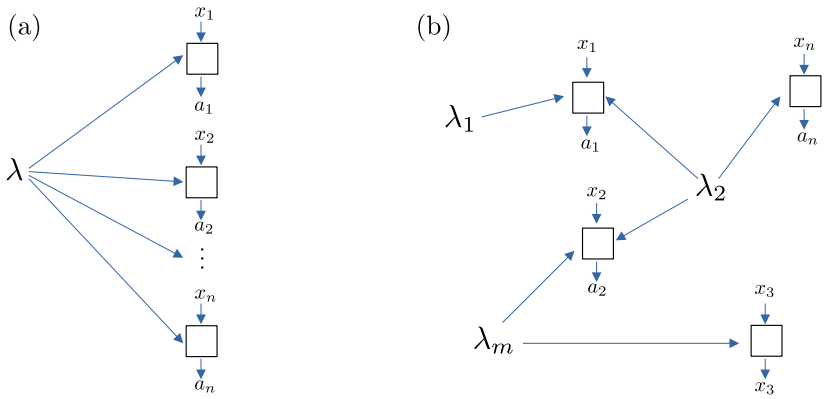

Figure 5.1: (a) Bell n-local experiment. (b) Example of network-local experiment.

nonlocal effects in the single-photon entangled state by passive optical means, that is, phase shifters, beamsplitters and photodetectors, is impossible.
The main result of our work [5] is to show that this is not the case and one can indeed reveal the nonlocality of state (5.5) with only *passive measurements*. To do so, we go beyond standard Bell tests and consider setups defined by causal networks. We show that three copies of single-photon entangled states placed in a *triangle causal network* (cf. Fig. 6.4) can exhibit non-classical correlations.

5.3 Network nonlocality and the Triangle Network

The standard Bell scenario can be extended to experiments involving multiple parties. A probabilistic output of $n > 2$ is then said *Bell n-local* if it can be decomposed as

$$p_{n-\mathcal{L}}(a_1, \ldots, a_n | x_1, \ldots, x_n) = \int \mathrm{d}\lambda \; \mu(\lambda) p(a_1|x_1, \lambda) \ldots p(a_n|x_n, \lambda) \; . \tag{5.6}$$

Such definition corresponds to all the n parties sharing a common (classical) source λ (see Fig. 5.1(a)).
More recently, theoreticians realized that n-local correlations are a particular case of a classical *causal network*, in which the non-observable source λ influences all n parties. Starting in the early 2010s, seminal works [61, 189–192] considered more complex causal structures involving several independent sources, each being distributed to a subset of the parties involved in the scenario. It is therefore possible to define *network-local* correlations those that

arise from parties sharing classical sources according to a structure defined by a network [60]. That is, for a network with m sources and n parties,

$$p_{\text{network}-\mathcal{L}}(a_1,\ldots,a_n|x_1,\ldots,x_n) = \int \mathrm{d}\lambda_1\ldots\mathrm{d}\lambda_m\; p(a_1|x_1,\vec{\lambda}^{(1)})\ldots p(a_n|x_n,\vec{\lambda}^{(n)})\;, \quad (5.7)$$

where $\vec{\lambda}^{(i)}$ is the array of sources that is sent to party (i). For example, for the network in Fig. 5.1(b), one has $\vec{\lambda}^{(1)} = \{\lambda_1,\lambda_2\}$, $\vec{\lambda}^{(2)} = \{\lambda_2,\lambda_m\}$, $\vec{\lambda}^{(3)} = \{\lambda_m\}$.

For a fixed network, the generic study of nonlocality consists in characterizing the set of probability distributions arising from (5.7), and whether such set is strictly larger when using quantum sources, or even theoretical post-quantum sources [60]. It is well understood that these networks offer new possibilities to design quantum experiments with no classical analogue [61, 62, 189, 191–194]. At the same time, the analysis of nonlocality in networks becomes mathematically and computationally challenging, as networks with $m > 1$ sources generally define non-convex sets of correlations, both in the classical and quantum case [60].

A network that has attracted considerable interest is the *triangle network* [61, 62, 195–199]. One of the main features such network is the possibility of exhibiting nonlocality without the need of measurement inputs [61, 62, 199]. In such case the considered Bell-type experiment consists of three observers, A, B and C, receiving states prepared by three sources, see Fig. 6.4. These states are measured producing outcomes a, b and c with probability $p(a,b,c)$, without inputs.

In this case, a classical local description of the experiment compatible with the causal constraints defined by the network has the form

$$p(a,b,c) = \int \mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma\; p_A(a|\beta,\gamma)p_B(b|\gamma,\alpha)p_C(c|\alpha,\beta)\;. \quad (5.8)$$

The causal model therefore consists of classical variables α, β and γ uniformly (without loss of generality) distributed by the sources and local response functions p_X, with $X = A, B, C$, producing the measurement outcomes. In analogy with standard Bell tests, we define probability distribution $p(a,b,c)$ that can be written as Eq. (5.8) as causally classical or, simpler, *local*.

A quantum description of the experiment compatible with the causal network replaces the random variables by quantum states ρ_α, ρ_β and ρ_γ and the local

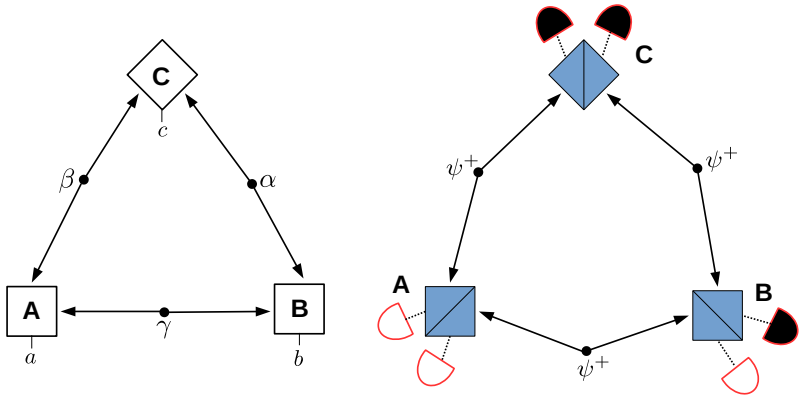

Figure 5.2: (Left) Causal model for the Triangle Network: three independent sources $\{\alpha, \beta, \gamma\}$ prepare correlated states that are distributed among the three parties. Each of them produces an output through a local process acting on the received parts of the states. The form of the states and local processes depend on the theory, say classical or quantum, used to reproduce the correlations in the network. (Right) Schematics of the proposed quantum optical experiment. A, B and C share single-photon entangled states $|\psi^+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$ prepared by the sources. Each party receives two optical modes that are mixed on a beamsplitter, the resulting output modes being measured by photodetectors. In the specific experimental instance depicted here, A does not detect any photon, B has one detector firing, and C has both detectors firing.

response functions by quantum measurements. Therefore, quantum probabilities compatible with the triangle network have the form

$$p(a,b,c) = \mathrm{Tr}\left[(\rho_\alpha \otimes \rho_\beta \otimes \rho_\gamma)(M_A^{(a)} \otimes M_B^{(b)} \otimes M_C^{(c)})\right] , \qquad (5.9)$$

where $M_A^{(a)}$ denote the positive measurement operators defining the Positive-Operator Valued Measure (POVM) for A, $\sum_a M_A^{(a)} = \mathbb{1}_A$, and similarly for B and C. We slightly abuse the notation in Eq. (5.9) by not specifying the tensor products and different Hilbert spaces in which the different operators act, but this is clear from Fig. 6.4. We say that a quantum experiment, defined by states and measurements producing the outcome distribution $p(a,b,c)$ according to Eq. (5.9), is *nonlocal* whenever this distribution cannot be described by a classical model (5.8).

5.4 Witnessing Single-Photon nonlocality in the Triangle network

In Ref. [5] we provided a nonlocal quantum experiment in the triangle network that only uses single-photon entangled states, beamsplitters and photodetectors. Our main idea is to exploit the topology of the network to reintroduce indeterminacy in the photon path, necessary to exploit the coherences of these states when using passive measurements.
The experimental setup proposed is depicted in Fig. 6.4: three parties A, B, C share, for each pair AB, BC, CA, the single photon entangled state $|\psi^+\rangle$, see Eq. (5.5). The initial state is thus

$$|\psi^+\rangle_{A_2B_1} \otimes |\psi^+\rangle_{B_2C_1} \otimes |\psi^+\rangle_{C_2A_1} \equiv |\Psi^+\rangle_{A_1A_2B_1B_2C_1C_2} \ . \tag{5.10}$$

Each party then receives its two optical inputs on modes X_1X_2 ($X = A, B, C$) and mixes them with a beamsplitter, which induces a unitary transformation $\mathcal{B}_{X_1X_2}(t, \phi)$ parametrized by its transmissivity t and phase ϕ. All parties use the same value for t, and the phases are all null for simplicity. After passing through the beamsplitters, the photons end up in photodetectors. For each mode X_i, the operators describing a perfectly efficient photodetection correspond to the projectors onto the vacuum state $D^{\square}_{X_i} = |0\rangle\langle 0|_{X_i}$ (detector off) and the projector on its orthogonal complement $D^{\blacksquare}_{X_i} = \mathbb{1}_{X_i} - |0\rangle\langle 0|_{X_i}$ (detector firing). Indeed, we assume that the detectors do not resolve the number of photons but only their presence. The measurement obtained by mixing two modes with the beamsplitter and the ideal photodetectors can be accordingly expressed as a POVM for each party (here $\mathcal{B}_{X_1X_2} = \mathcal{B}_{X_1X_2}(t, 0)$)

$$\begin{aligned}
{\Pi_t^{(0)}}_{X_1X_2} &= \mathcal{B}^\dagger_{X_1X_2}(D^{\square}_{X_1} \otimes D^{\square}_{X_2})\mathcal{B}_{X_1X_2}, \\
{\Pi_t^{(L)}}_{X_1X_2} &= \mathcal{B}^\dagger_{X_1X_2}(D^{\blacksquare}_{X_1} \otimes D^{\square}_{X_2})\mathcal{B}_{X_1X_2}, \\
{\Pi_t^{(R)}}_{X_1X_2} &= \mathcal{B}^\dagger_{X_1X_2}(D^{\square}_{X_1} \otimes D^{\blacksquare}_{X_2})\mathcal{B}_{X_1X_2}, \\
{\Pi_t^{(2)}}_{X_1X_2} &= \mathcal{B}^\dagger_{X_1X_2}(D^{\blacksquare}_{X_1} \otimes D^{\blacksquare}_{X_2})\mathcal{B}_{X_1X_2},
\end{aligned} \tag{5.11}$$

where the measurement labels stand respectively for no photon counts (0), a count in the left detector (L), a count in the right detector (R), or counts in both detectors (2). The crucial point is that when $t \neq 0$, the L and R measurements actually detect superpositions of photons in the incoming modes (see details in [5]).

The quantum experiment described here results in the output distribution

$$p_t(abc) = \mathrm{Tr}[|\Psi^+\rangle\langle\Psi^+| \, (\Pi_t^{(a)} \otimes \Pi_t^{(b)} \otimes \Pi_t^{(c)})] \\ a, b, c \in \{0, L, R, 2\} \tag{5.12}$$

which depends on the transmissivity t of the beamsplitters used by the parties. The resulting distribution p_t turns out to be nonlocal in a wide range of transmissivities t. More precisely, to prove nonlocality, first we simplified the structure that classical strategies must follow in the triangle network (5.8). Specifically, all the local response functions p_A, p_B, p_C in (5.8) can be assumed to be deterministic, and all the indeterminacy is therefore delegated to the classical sources $\{\alpha, \beta, \gamma\}$, which can all be assumed to be, w.l.o.g, real numbers uniformly distributed in the interval $[0, 1]$. Therefore, any local model is specified by deterministic triangle-local response functions $p_A p_B p_C$ that map all the points of the cube $[0, 1]^3$ to the observed outputs

$$\{\alpha, \beta, \gamma\} \to \{a(\beta, \gamma), b(\gamma, \alpha), c(\alpha, \beta)\} \, . \tag{5.13}$$

Secondly, we were able to identify strict constraints that need to be satisfied by all possible classical causal models simulating the considered experimental output $p_t(abc)$ in the triangle network. In particular, we exploited the cyclic symmetry and null components (due to photon number conservation) of the distribution. For example, all outputs of the form (here χ represents any of L or R) $\{(000), (00\chi), (2\chi\chi), (22\chi)\}$, or any of their permutations, have zero probability, due to the fact that there are initially 3 photons in the network, of which at most 2 can end up in the same photodetector. That is, in each run of the experiment the total number of clicks in the detectors must be 2 or 3. By taking all the relevant properties of p_t into account, one can identify constraints that need to be satisfied by any classical strategy, specified by the response functions (5.13), aiming at reproducing p_t. In fact, while the exact form of the response functions remains in general unknown, some of its marginals can be expressed in terms of the output p_t. These relevant marginals are nothing other than linear constraints on the response functions, parametrized by t. Together with standard normalization and positivity constraints, these define a Linear Program. The feasibility of such Linear Program is, by definition, necessary for the existence of such local response functions. Therefore, when infeasible, no local model exists to simulate our experiment proposal. Results show that the Linear Program is infeasible for $t \in (0.785, 1)$ and $t \in (0, 0.215)$, proving the claims of this section.

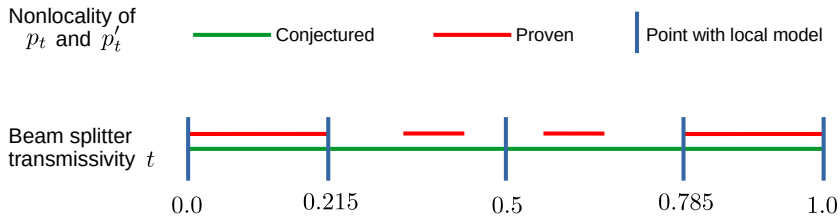

Figure 5.3: Range of nonlocality for the distribution p_t (5.12) resulting from the ideal experiment represented in Figure 6.4. The nonlocality of such distribution is equivalent to the distribution p'_t of Ref. [62], whose nonlocality has been proven to be in the range $t \in (0, 0.215) \cup (0.344, 0.437) \cup (0.563, 0.656) \cup (0.785, 1)$ (see Ref.s [62, 198]), and conjectured in [201] to be over the whole $[0, 1]$ interval except the indicated five points, which have a local model [62].

The techniques we used are similar to those introduced in [62] and generalized in [200]. However, their findings cannot be applied directly to our scenario. The reason behind this is that the works [62, 200] are based on a token-counting approach to some physical "tokens" that are: $i)$ generated from the sources, $ii)$ distributed to the parties in a coherent superposition of different ways, and $iii)$ counted at the output. In our experiment the physical tokens are the photons, which however can be miscounted at the output, as more than one could enter in the same photodetector [1]. For these reasons, we had to extend these techniques so that they could be applied to our setup.
Importantly, as part of the proof, we showed that our distribution is nonlocal if and only if the distribution proposed in [62], which we dub p'_t, is nonlocal as well. After finishing our manuscript, Pozas-Kerstjens *et al.* proved [198] p'_t is nonlocal in the range $t \in (0.563, 0.656)$ and $t \in (0.344, 0.437)$ as well[2]. Nonlocality of p'_t in the intervals $(0.215, 0.5)$ and $(0.5, 0.785)$ has been conjectured already in [201]. Given the above mentioned equivalence between the nonlocality of p_t and p'_t proven in this work, this would imply that the proposed ideal experiment is nonlocal for all transmissivities except $t \in \{0.0, 0.215, 0.5, 0.785, 1.0\}$, which are known to have local models (cf. [5, 62]). The up-to-date summary of nonlocality for our distribution is given in Fig. 5.3.

[1] In [5] we also discuss the idealized case with perfect, number-resolving photodetectors.
[2] Notice that both Ref. [62] and [198] use as main parameter for the distribution $u = \sqrt{t}$.

5.5 Optical realisation and noise analysis

After proving the nonlocality of the outputs of the ideal noiseless experiment, we analyzed the robustness of our results against typical noise errors, by modelling imperfections which occur in experimental realizations of the optical network presented in Fig. 6.4.
In our modelling of the network we focused on three main parameters: the impurity of the generated single-photon entangled state (Q), the transmissivity of the optical channels (T) of the network, and the efficiency of the final photodetectors (ν). We thus obtained a noisy output distribution $p_t^{Q,T,\nu}(a,b,c)$ (see details in [5]). It follows that

$$p_t^{Q=0,T=1,\nu=1}(a,b,c) \equiv p_t(a,b,c) \ , \tag{5.14}$$

that is, with no impurity, and perfect transmission and detection, we recover the idealized experiment.
Inevitably, part of the key properties and symmetries of $p_t(a,b,c)$ disappear as soon as noise is introduced in the network. This makes the analytic approach unworkable in this case. Consequently, in order to estimate the tolerance to the noises introduced above, we resorted to a technique recently introduced in [201]: there, a feed-forward neural network is shaped with the same topology of the causal network under study, and it is then asked to reproduce the target distribution $p_t^{Q,T,\nu}$. Each output of the neural network is thus *literally* an instance of a classical model (which can be therefore described by Eq. (5.8) in our case) trying to reproduce $p_t^{Q,T,\nu}$. For a fixed target distribution, the neural network is trained by minimizing the Euclidean distance from the neural network's local model to the target. When the target distribution is inside the local set, a sufficiently large neural network should be capable of learning it. Instead, a large distance between the machine's best guess and the target is taken as an indication of nonlocality. What it means to be "large" enough can be somewhat arbitrary, since some nonlocal behaviors are extremely close to the local set (as is the case here), and additionally the neural network's model is not guaranteed to converge to the optimal solution as it can get stuck in local minima during training. In order to gain deeper insight into the boundary between locality and nonlocality we examine transitions of the learning algorithm's behavior when adding noise to the target distribution, and retraining the machine independently for each target distribution. The very noisy case is guaranteed to be local and the machine learning results on those give a reference to which we can compare the nonlocal regime. By definition, this technique does not certify nonlocality in an absolute way, but has been shown to be reliable and efficient from the point of view of computational resources [201].

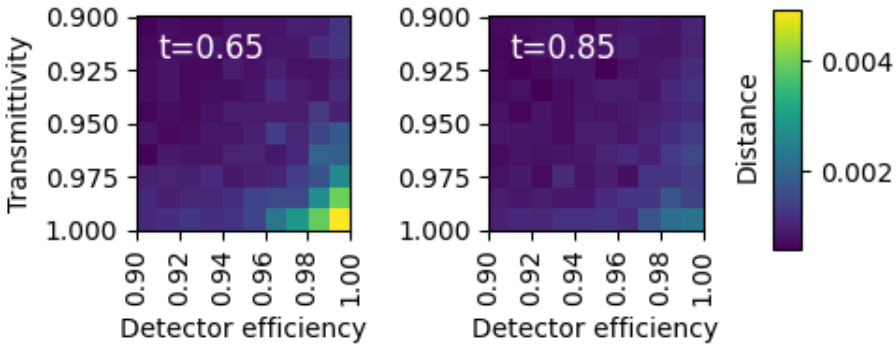

Figure 5.4: Euclidean distance of machine learned local models from the noisy distribution $p_t^{(Q,T,\nu)}$ under an experimentally realistic noise model for $t = 0.65$ (left) and $t = 0.85$ (right), with $Q = 0.006875$ for both (see [5]).

For the noiseless distribution $p_t^{Q=0,T=1,\nu=1}(a,b,c) \equiv p_t(a,b,c)$, the neural network's best guess is distant from the experimental output, corroborating the analytical proof of nonlocality for $t \in (0.785, 1)$. At the same time the neural network hints at the locality of the output distribution for $t = 0.5$ and $t = 1$, which have local strategies (cf. Fig. 5.3). A local model exists as well for $t \sim 0.785$. Moreover, the same machine indicates (seemingly even stronger) nonlocality in the range $t \in (0.5, 0.785)$, in line with the conjecture of [201] and the results of [198] (cf. Fig. 5.3).
The noise robustness is, however, small. The neural network seems to indicate that the points that are "most nonlocal" are $t \sim 0.85$ and $t \sim 0.65$. For these two points we tested the tolerance to the physical noises introduced above, see Fig. 5.4: choosing $Q \simeq 0,7\%$ (cf. [5]), the neural network tries to learn $p_t^{(Q,T,\nu)}$ for different values of the transmissivity T and detector efficiency ν. Results show that nonlocality is more robust for $t = 0.65$, where it is lost when $T \lesssim 95\%$ or $\nu \lesssim 95\%$.

5.6 Discussion

In Ref. [5] we provided an experimental proposal to demonstrate nonlocality of single-photon entangled states. Typical photon-based Bell experiment are usually based on two-photon states in which entanglement is encoded in the polarization degree of freedom. At the same time, heralded preparation of a two-photon maximally entangled state is quite challenging, even more if more than one source has to be prepared contemporaneously. In turn, single-photon entanglement can be easily prepared in a heralded way: an arbitrarily good

approximation to it can be obtained when detecting photons in one of the two modes resulting from the SPDC process and sending the non-measured mode into a balanced beamsplitter [5]. This makes it a relevant technological candidate for experiments and applications of network nonlocality.
Remarkably, the setup we proposed is not only passive in terms of the implemented measurements, but also because it does not require any active choice of measurements. That is, in our setup, there are no classical inputs and observers perform a single measurement on their received shares. These characteristics make the proposal, arguably, the simplest experimental demonstration of the nonlocality of the single-photon entangled state, as well as the first experimental proposal for genuine network nonlocality [62].

Chapter 6

Measurement-device-independent certification of quantum properties

This chapter is based on Ref. [3]:
"P. Abiuso, S. Bäuml, D. Cavalcanti and A. Acín, *Measurement-device-independent entanglement detection for continuous-variable systems*, Phys. Rev. Lett. **126**, 190502 (2021)."

6.1 Background: Entanglement detection, Nonlocality and Device-Independent certifications

Entanglement is the main resource for a broad range of applications in quantum information science, among which are quantum key distribution [171], quantum computation [202], and quantum metrology [203]. It is therefore crucial to develop methods to detect entanglement that are reliable and practical. The most common method to detect entanglement is given by entanglement witnesses [204]. The typical entanglement detection routine consists in finding an observable W_{AB} whose average value is, without loss of generality, positive for some entangled state ρ_{AB}^{ent}, while at the same time negative for any separable state. That is ,

$$\text{Tr}[\rho_{AB}^{\text{ent}} W_{AB}] > 0 \,, \tag{6.1}$$

while

$$\text{Tr}[\rho_{AB} W_{AB}] \leq 0 \quad \forall \rho | \rho = \sum_i p_i \rho_A^{(i)} \otimes \rho_B^{(i)} \,. \tag{6.2}$$

For any entangled ρ_{AB}^{ent}, the existence of an observable W_{AB} satisfying (6.1) and (6.2) is guaranteed by the Hahn-Banach theorem and the fact that the set of separable states is convex [56]. However, to be reliable this technique requires a perfect implementation of the measurements that lead to the observable W_{AB}. Indeed, small calibration errors can lead to false-positive detection of entanglement [205, 206], which can be critical when using the wrongly detected entangled state for quantum information purposes. A possible way of circumventing this problem is to move into the so called device-independent (DI) scenario [59]. The violation of Bell inequalities (cf. 5.1), can in fact be seen as a DI entanglement witness. The reason is that separable states cannot violate any Bell inequality. This can be clearly seen: using a separable state

$$\rho_{AB} = \sum_i p_i \rho_A^{(i)} \otimes \rho_B^{(i)} \ , \tag{6.3}$$

in a Bell experiment, leads to a probabilistic output (cf. (5.3))

$$\begin{aligned} p(a,b|x,y) &= \text{Tr}[(M_A^{a|x} \otimes M_B^{b|y})\rho_{AB}] \\ &= \sum_i p_i \, \text{Tr}[M_A^{a|x} \rho_A^{(i)}] \text{Tr}[M_B^{b|y} \rho_B^{(i)}] \ , \end{aligned} \tag{6.4}$$

which is clearly local, i.e. in the form (5.2), via identifying the shared classical source $\mu(\lambda) := p_\lambda$ and the local response functions $p(a|x,\lambda) := \text{Tr}[M_A^{a|x} \rho_A^{(\lambda)}]$ (similarly for B). In the DI framework measurements and sources do not need to be characterised, since entanglement is detected through the violation of Bell inequalities, which only use the statistics provided by the experiment, without making any assumptions on the real implementation, except the no-signalling assumption (5.1), which can be guaranteed e.g. by isolating the labs of A and B, or performing space-like separated measurements.
The DI scenario is however stringent from an experimental point of view, requiring low levels of noise and high detection efficiencies. This is why other approaches requiring intermediate level of trust on the devices have been developed. In particular, there exist methods that do not require any characterization of the measurement implemented for entanglement detection, known as measurement-device-independent (MDI) [64, 65].

6.2 Measurement-device-independent entanglement witnessing

An entanglement detection scenario where two parties, Alice and Bob, do not assume a particular description of their measurement but use trusted sources

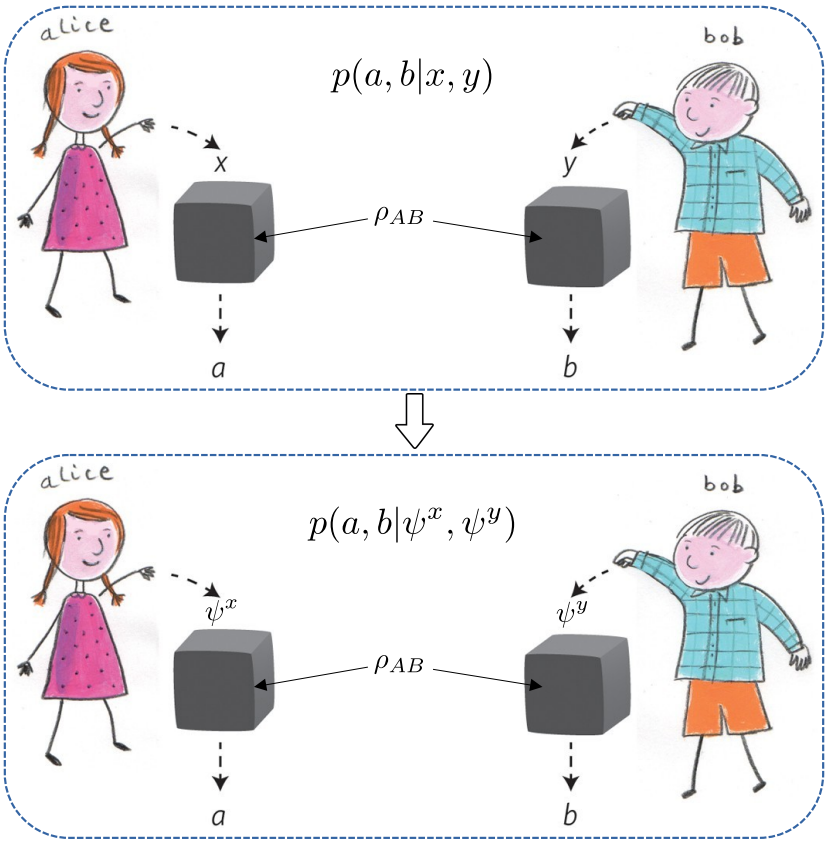

Figure 6.1: Entanglement witnessing in the Device-Independent scenario (upper image) and in the Measurement-Device-Independent scenario (lower image): the latter can be seen as a relaxation of a Bell game in which the inputs are quantum states.

of states was first introduced by Buscemi [65]. Namely, let us consider that Alice and Bob can produce states $\psi^{\mu}_{A'}$ and $\psi^{\nu}_{B'}$ according to some distributions $p_{A'}(\mu)$, $p_{B'}(\nu)$ respectively. Alice and Bob can use these states as inputs to their measurement devices, which return outcomes a and b respectively. Then, assuming only the validity of quantum theory and the locality of the measurements, these outcomes occur with probability

$$p(a,b|\psi^{\mu}_{A'},\psi^{\nu}_{B'}) = \mathrm{Tr}[M^{a}_{AA'} \otimes N^{b}_{BB'}(\psi^{\mu}_{A'} \otimes \rho_{AB} \otimes \psi^{\nu}_{B'})] \,, \tag{6.5}$$

where $M^{a}_{AA'}$ and $N^{b}_{BB'}$ are unknown elements of a Positive-Operator Valued Measure (POVM), and ρ_{AB} is also undetermined. The main goal of Alice and Bob is to determine if ρ_{AB} is entangled based on the knowledge of $p(a,b|\psi^{\mu}_{A'},\psi^{\nu}_{B'})$, as well as the quantum inputs and their distribution $\psi^{\mu}_{A'}$, $\psi^{\nu}_{B'}$, $p_{A'}(\mu)$, $p_{B'}(\nu)$. In this sense, MDI entanglement witnessing can be seen as a modified version of standard Bell nonlocality experiments (cf. Fig. 6.1). Besides the calibration issue discussed before, this scenario is motivated by cryptographic tasks in which Alice and Bob do not trust the provider of the measurement devices they are using [207–210].

For finite dimensional Hilbert spaces, Buscemi has shown that any entangled state ρ_{AB} can be certified in this scenario, but his proof is not constructive [65]. The authors of [64] have shown how to construct MDI entanglement witness from standard entanglement witnesses. At the same time, an intuitive explanation of why it is possible to detect any entangled state in this scenario was given in [211]. Indeed collecting the statistics $p(a,b|\psi^{\mu}_{A'},\psi^{\nu}_{B'})$ from the experiment (6.5), and having the knowledge of the inputs ψ^{μ} and ψ^{ν}, corresponds to testing, on such inputs the POVM defined as

$$\mathcal{M}^{a,b}_{A'B'} := \mathrm{Tr}_{AB}[(M^{a}_{AA'} \otimes N^{b}_{BB'})\rho_{AB}] \,. \tag{6.6}$$

In fact, if follows from its definition that

$$p(a,b|\psi^{\mu}_{A'},\psi^{\nu}_{B'}) = \mathrm{Tr}[\mathcal{M}^{a,b}_{A'B'}(\psi^{\mu}_{A'} \otimes \psi^{\nu}_{B'})] \,. \tag{6.7}$$

Moreover it is easy to see that if ρ_{AB} is separable, the corresponding $\mathcal{M}^{a,b}_{A'B'}$ is also a separable POVM. If instead ρ_{AB} is entangled, it is always possible to choose $M^{a}_{AA'}$ and $N^{b}_{BB'}$ such that the corresponding $\mathcal{M}^{a,b}_{A'B'}$ is entangled as well [211]: for example it is enough to choose measurements including the projection on the maximally entangled state $|\phi^{+}\rangle := \frac{1}{\sqrt{d}}\sum_{i}^{d}|ii\rangle$, as in [64] (here d is the dimension of the local Hilbert space). In fact, given

$M^1_{AA'} = |\psi^+\rangle\langle\psi^+|$ and $N^1_{BB'} = |\psi^+\rangle\langle\psi^+|$, the resulting POVM has an element proportional to the state ρ_{AB} itself

$$\mathcal{M}^{1,1}_{A'B'} = \frac{1}{d^2}\rho^{\mathsf{T}}_{A'B'} \,. \tag{6.8}$$

This means that, if the correct measurements are chosen, the entanglement of $\mathcal{M}^{a,b}_{A'B'}$ can then be detected by choosing a tomographically complete basis $\{\psi^\mu_{A'}, \psi^\nu_{B'}\}$ and using Eq. (6.7) to reconstruct $\mathcal{M}^{a,b}_{A'B'}$.

6.3 MDI entanglement witnessing for continuous variables states

In what follows we present the main results of our work [3]. We first generalise the results of [64, 65] and show that the entanglement of every CV entangled state can in principle be detected in a MDI scenario. We then move to the experimentally relevant case of Gaussian states and operations and show a MDI protocol that is able to certify the entanglement of all two-mode Gaussian entangled states. Moreover, this protocol only requires the production of coherent states and the implementation of homodyne measurements. Notice that a fully DI approach in this context is not well suited, because as we lack a sistematic and practical approach to find useful Bell tests for continuous-variable states. For instance, in the Gaussian regime, which is the most feasible experimentally, DI entanglement detection is impossible because no Bell inequality can be violated [59].
Our approach provides also an interesting connection between MDI entanglement detection and quantum metrology.

6.3.1 Reduction to process tomography

In this section, we show that it is possible to detect the entanglement of any entangled state in a MDI scenario where Alice and Bob use coherent states as trusted inputs (the proof is presented for two-mode bipartite states but can be generalized to more modes, see Supplementary material of [3]). The technique is a CV adaptation of the scheme used in [64].
Suppose Alice and Bob are in possession of trusted sources producing coherent states $|\alpha\rangle_{A'}$ and $|\beta\rangle_{B'}$, respectively, according to some distribution. The shared entangled state is ρ_{AB}. The systems AA' and BB' are then projected onto respective two-mode squeezed vacuum (TMSV) states, i.e. the measurement $\{|\Phi^{(r)}\rangle\langle\Phi^{(r)}|, \mathbb{1} - |\Phi^{(r)}\rangle\langle\Phi^{(r)}|\}$ is performed on both AA' and

BB' (r is the squeezing parameter). Conditioned on α and β, the probability of both measurements obtaining output '1', corresponding to the projector $|\Phi^{(r)}\rangle\langle\Phi^{(r)}| \equiv \Phi^{(r)}$, can be expressed as

$$\begin{aligned} &P_\rho(1,1|\alpha,\beta) \\ &= \mathrm{Tr}\left[\left(\Phi^{(r)}_{AA'} \otimes \Phi^{(r)}_{BB'}\right)\left(|\alpha\rangle\langle\alpha|_{A'} \otimes \rho_{AB} \otimes |\beta\rangle\langle\beta|_{B'}\right)\right] \\ &= \mathrm{Tr}\left[M^{(r)}_{A'B'}\,|\alpha\rangle\langle\alpha|_{A'} \otimes |\beta\rangle\langle\beta|_{B'}\right], \end{aligned} \tag{6.9}$$

where we have defined

$$M^{(r)}_{A'B'} := \mathrm{Tr}_{AB}\left[\left(\Phi^{(r)}_{AA'} \otimes \Phi^{(r)}_{BB'}\right)\left(\rho_{AB} \otimes \mathbb{1}_{A'B'}\right)\right], \tag{6.10}$$

which is a POVM element by construction. Our main observation in this section is that non-separability of the POVM element defined by (6.10) is equivalent to the underlying state being entangled. Namely, we have

Proposition 1. *For any $r > 0$, the POVM element $M^{(r)}_{A'B'}$ defined by eq. (6.10) is entangled if and only if ρ_{AB} is entangled.*

Proof. Let us assume ρ_{AB} is separable, i.e.

$$\rho_{AB} = \sum_\mu p_\mu \rho^\mu_A \otimes \sigma^\mu_B. \tag{6.11}$$

We can then define

$$\begin{aligned} M^{(r)\mu}_{A'} :=& \mathrm{Tr}_A\left[\Phi^{(r)}_{AA'}\left(\rho^\mu_A \otimes \mathbb{1}_{A'}\right)\right], \\ N^{(r)\mu}_{B'} :=& \mathrm{Tr}_B\left[\Phi^{(r)}_{BB'}\left(\sigma^\mu_B \otimes \mathbb{1}_{B'}\right)\right], \end{aligned} \tag{6.12}$$

which are POVM elements by construction. It is easy to see that $M^{(r)}_{A'B'} = \sum_\mu p_\mu M^{(r)\mu}_{A'} \otimes N^{(r)\mu}_{B'}$, which is separable. It remains to be shown that the POVM element defined by Eq. (6.10), which can be rewritten as [1]

$$M^{(r)}_{A'B'} = (1-\lambda^2)\lambda^{\hat{n}_A+\hat{n}_B}\rho^T_{AB}\lambda^{\hat{n}_A+\hat{n}_B} \tag{6.13}$$

[1] Here we used the well-known decomposition of two-mode squeezed states in the Fock basis

$$|\Phi^{(r)}\rangle = \sqrt{1-\tanh^2 r}\sum_{i=0}^{\infty}(\tanh r)^i\,|ii\rangle$$

(where $\lambda = \tanh r$ and $\hat{n}_X = a_X^\dagger a_X$ the number operator on mode X), is entangled for all entangled ρ_{AB}. In fact, suppose there exists an entanglement witness W such that $\mathrm{Tr}[\rho W] < 0$ while $\mathrm{Tr}[\rho' W] \geq 0$ for any ρ' separable [2]. From W we can obtain a Hermitian operator $\tilde{W}$ such that $\mathrm{Tr}[M^{(r)}_{A'B'}\tilde{W}] < 0$, whereas for any separable POVM $\mathrm{Tr}[\sum_\mu p_u (M^\mu_A \otimes N^\mu_B)\tilde{W}] \geq 0$. Consider in fact

$$\tilde{W} = \lambda^{-\hat{n}_A - \hat{n}_B} W^T \lambda^{-\hat{n}_A - \hat{n}_B} \ . \tag{6.14}$$

It is then easy to see that

$$\mathrm{Tr}[M^{(r)}_{A'B'}\tilde{W}] = (1-\lambda^2)\mathrm{Tr}[\rho W] < 0 \ . \tag{6.15}$$

For separable POVMs, on the other hand, it holds

$$\sum_\mu p_\mu \mathrm{Tr}[\tilde{W}(M^\mu_A \otimes N^\mu_B)] = \sum_\mu p_\mu \mathrm{Tr}[W(\tilde{M}^\mu_A \otimes \tilde{N}^\mu_B)], \tag{6.16}$$

where $\tilde{M}^\mu_A = \lambda^{\hat{n}_A}(M^\mu_A)^T\lambda^{\hat{n}_A}$, $\tilde{N}^\mu_B = \lambda^{\hat{n}_B}(N^\mu_B)^T\lambda^{\hat{n}_B}$. The operators $\tilde{M}^\mu_A$ and $\tilde{N}^\mu_B$ are manifestly positive semidefinite, meaning that under a proper renormalization they can be seen as states, thus generating a separable state ρ' such that $\mathrm{Tr}[\rho' W] \geq 0$. This implies

$$\sum_\mu p_\mu \mathrm{Tr}[W(\tilde{M}^\mu_A \otimes \tilde{M}^\mu_B)] \geq 0 \ , \tag{6.17}$$

which finishes the proof. □

We show in [3] that the violation of the derived witness (6.15) scales as $1/N$, where N is the energy scale (number of photons) defined by the original witness, which is to be compared with the $1/d$ scaling found in [64] (d being the Hilbert space dimension of ρ_{AB}).
As a consequence of Proposition 1, Alice and Bob can certify the entanglement of ρ_{AB} in a MDI way, if their output statistics allow them to fully reconstruct the POVM element $M^{(r)}_{A'B'}$. As in the case of discrete variables [211], this can be achieved by means of tomography. In order to perform measurement tomography, the input states have to be chosen from a tomographically complete set [214]. The set of all coherent states form a tomographically complete set via the Glauber-Sudarshan P-representation [215, 216].

[2] Such a witness always exists, as the set of separable states is defined to be closed for operational consistence, see e.g. [212, 213]

Once Alice and Bob have reconstructed $M^{(r)}_{A'B'}$, they can determine whether it is non-separable using an entanglement criterion. We also note that, for a given entangled state, if the witness W is known, all that is necessary is to evaluate $\mathrm{Tr}[M^{(r)}_{A'B'}\tilde{W}]$, which might not require full tomography. In summary, we have the following

Corollary 1. *For every entangled state ρ_{AB}, if $|\alpha\rangle$ and $|\beta\rangle$ are chosen from tomographically complete sets, Alice and Bob can certify the entanglement of ρ_{AB} in a measurement-device-independent way.*

The results presented in this section suffer from practical problems in their realization: firstly, they rely on performing the POVM that projects on the two mode squeezed states defined in Eq. (6.9). A typical scheme for such measurement involves photodetection, which typically has low efficiency and high cost. Secondly, the full tomography could be in general experimentally inefficient. Therefore, the previous proof is mostly a proof-of-principle result. Next, we show that feasible schemes for MDI entanglement detection are possible. In fact, we propose an experimentally-friendly MDI entanglement detection protocol which is based solely on homodyne measurements and can detect all two-mode Gaussian entangled states.

6.3.2 MDI Entanglement Witness for all two-mode Gaussian states

In this section we present a practical method for MDI entanglement certification of Gaussian states that can be implemented using readily available optical components. Our method is inspired by the entanglement witness introduced in a seminal paper by Duan *et al.* [217] (see also Simon [218]). In that work, it was proven that the inequality

$$\langle \mathrm{EW}_\kappa \rangle \equiv \left\langle \Delta^2 \hat{u}_\kappa \right\rangle + \left\langle \Delta^2 \hat{v}_\kappa \right\rangle \geq \frac{\kappa^2 + \kappa^{-2}}{2}\,, \tag{6.18}$$

where $\langle \Delta^2 \hat{O} \rangle$ is the variance of the operator $\hat{O}$, and

$$\hat{u}_\kappa = \left(\kappa \hat{x}_A - \frac{\hat{x}_B}{\kappa}\right)\,, \quad \hat{v}_\kappa = \left(\kappa \hat{p}_A + \frac{\hat{p}_B}{\kappa}\right)\,, \tag{6.19}$$

(i) holds for any two-mode separable state, real number κ, where $\kappa > 0$ without loss of generality, and pairs of orthogonal quadratures of the bosonic

modes A and B [3], while (ii) for any entangled Gaussian state there exist a value of κ and pairs of quadratures such that Eq. (6.18) is violated.
Our main result is an experimentally-friendly method for MDI entanglement detection inspired by the witness (6.18) and given by the following proposition:

Proposition 2. *Let $|\alpha\rangle$ and $|\beta\rangle$ be coherent states prepared by Alice and Bob according to the Gaussian probability distribution*

$$P(\alpha) = \frac{1}{\pi\sigma^2} e^{-|\alpha|^2/\sigma^2} \qquad \alpha \equiv \alpha_x + i\alpha_p \tag{6.20}$$

(for different choices of input distribution, see [3]). Consider the setup in Fig. (6.4) *in which uncharacterized local measurements are applied jointly on these states and half of an unknown state ρ_{AB}, producing as a result two real numbers (a_1, a_2) for Alice and (b_1, b_2) for Bob. For all local measurements and all separable states one has*

$$\langle \mathrm{MDIEW}_\kappa \rangle \equiv \langle U_\kappa^2 \rangle + \langle V_\kappa^2 \rangle \geq \frac{\kappa^2 + \kappa^{-2}}{2} \frac{\sigma^2}{1+\sigma^2}, \tag{6.21}$$

where U_κ and V_κ are

$$\begin{aligned} U_\kappa &\equiv \kappa a_1 - \frac{b_1}{\kappa} - \frac{\kappa\alpha_x - \frac{\beta_x}{\kappa}}{\sqrt{2}}\,, \\ V_\kappa &\equiv \kappa a_2 + \frac{b_2}{\kappa} - \frac{\kappa\alpha_p + \frac{\beta_p}{\kappa}}{\sqrt{2}}. \end{aligned} \tag{6.22}$$

For any two-mode entangled Gaussian state, there exist local measurements acting jointly on the state and the input coherent states violating inequality (6.21).

Looking at its definition, the operational meaning of the witness is clear: Alice and Bob results should be such that their difference and sum, weighted by κ, are as close as possible to the same difference and sum of the quadratures of the coherent states, divided by $\sqrt{2}$. Note also that expectation values in (6.21) are

[3] We use here a notation similar to [63]. In particular the quadratures are defined as

$$\hat{a} = \hat{x} + i\hat{p}\,, \quad \hat{a}^\dagger = \hat{x} - i\hat{p}\,,$$

such that

$$[\hat{x}, \hat{p}] = \frac{1}{4i}[\hat{a} + \hat{a}^\dagger, \hat{a} - \hat{a}^\dagger] = \frac{i}{2}\,.$$

Note for example that the variances of the quadratures will have a factor $1/2$ with respect to the normalization choice of Duan *et al.* [217].

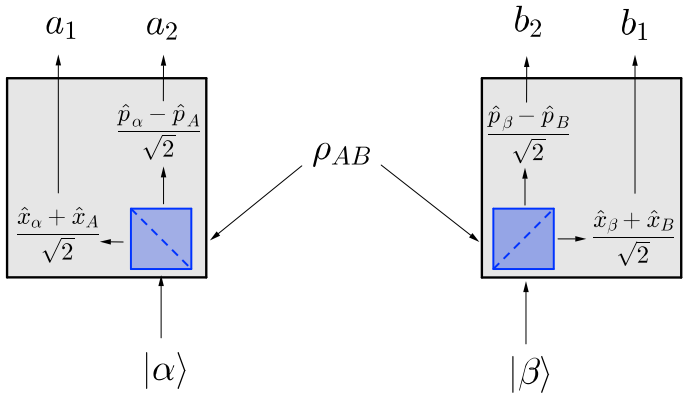

Figure 6.2: Experimental setup for the MDI entanglement detection of a 2-mode state. Fiduciary coherent states are prepared by the parties and measured together with the corresponding subsystems of the unknown state ρ_{AB}. To compute the bound of the entanglement witness (6.21), measurements should be seen as uncharacterized black boxes producing the outputs. To obtain a violation, the following specific measurements are implemented: the coherent states are mixed with the respective modes of ρ_{AB} in a 50:50 beam splitter and homodyne measurements of $\hat{x}$ and $\hat{p}$ are performed on the two outputs.

computed with respect to the quantum state and the distribution of coherent states.

Proof of (6.21). To prove the inequality we need to minimize the value of the witness over all separable states $\rho_{AB} = \sum_i p_i \rho_A^{(i)} \otimes \rho_B^{(i)}$. We can restrict the analysis to product states because the witness is linear on the state. It follows that the output probability factorises

$$p(a_1, a_2, b_1, b_2|\alpha, \beta) = \mathrm{Tr}\left[M^A_{a_1,a_2} |\alpha\rangle\langle\alpha|\right] \mathrm{Tr}\left[M^B_{b_1,b_2} |\beta\rangle\langle\beta|\right], \tag{6.23}$$

and we are left with two independent POVMs on the input states to be optimised. Using $\langle U_\kappa^2\rangle + \langle V_\kappa^2\rangle \geq \langle \Delta^2 U_\kappa\rangle + \langle \Delta^2 V_\kappa\rangle$ and the fact that the distribution is completely factorised between the two sides it follows that the value of the witness is lower bounded by the minimum of

$$\frac{1}{2}\left(\kappa^2 + \frac{1}{\kappa^2}\right)\left(\left\langle \Delta^2\left[\sqrt{2}a_1 - \alpha_x\right]\right\rangle + \left\langle \Delta^2\left[\sqrt{2}a_2 - \alpha_p\right]\right\rangle\right). \tag{6.24}$$

That is: in the absence of correlations, the best Alice and Bob can do is to separately perform the optimal measurements to estimate the input coherent

states.
Minimizing the expression in the second parenthesis in (6.24) looks essentially like a metrology problem. A lower bound, in turn, can be obtained using a multi-parameter Bayesian version of the quantum Cramér-Rao bound [219]. The simultaneous estimation of the position and momentum quadratures has been studied thoroughly and is optimized for coherent states by measuring $\hat{x}$ and $\hat{p}$ on two different modes after a 50:50 beam-splitter [220, 221]. In particular, assuming a Gaussian prior distribution, like in our case, the minimal sum of variances is equal to $\sigma^2/(1+\sigma^2)$ [220], which proves the bound for the entanglement witness. □

To prove the violation claimed in Proposition 2, it suffices to show that for any entangled two-mode Gaussian state there exist local measurements and values of (κ, σ) that lead to it. Consider now the optical setup depicted in Fig. 6.4. Alice and Bob, upon receiving their respective subsystems of ρ_{AB}, first mix them with local coherent states in a balanced beam splitter, and then measure the position and momentum quadratures in the output ports. The output observables are thus

$$\begin{aligned} \hat{A}_1 &= \frac{\hat{x}_\alpha + \hat{x}_A}{\sqrt{2}}\,, \quad \hat{A}_2 = \frac{\hat{p}_\alpha - \hat{p}_A}{\sqrt{2}}\,, \\ \hat{B}_1 &= \frac{\hat{x}_\beta + \hat{x}_B}{\sqrt{2}}\,, \quad \hat{B}_2 = \frac{\hat{p}_\beta - \hat{p}_B}{\sqrt{2}}. \end{aligned} \tag{6.25}$$

The quadratures $\hat{x}_A$, $\hat{p}_A$, $\hat{x}_B$ and $\hat{p}_B$ are those used in the standard device-dependent witness (6.18). Observables (6.25) are used to define the measurement outputs needed for the computation of our MDI witness (6.21). More precisely, consider first the case in which the average values of the state's quadratures are null, $\langle \hat{x}_A \rangle = \langle \hat{p}_A \rangle = \langle \hat{x}_B \rangle = \langle \hat{p}_B \rangle = 0$. Then, by taking (a_1, a_2, b_1, b_2) equal to the statistical output of $(\hat{A}_1, \hat{A}_2, \hat{B}_1, \hat{B}_2)$ respectively, it follows by substituting in (6.22) [4]

$$\begin{aligned} \langle U_\kappa^2 \rangle &= \left\langle \left(\kappa \frac{\hat{x}_\alpha + \hat{x}_A}{\sqrt{2}} - \frac{1}{\kappa} \frac{\hat{x}_\beta + \hat{x}_B}{\sqrt{2}} - \frac{\kappa \alpha_x - \frac{\beta_x}{\kappa}}{\sqrt{2}} \right)^2 \right\rangle \\ &= \frac{1}{2} \left(\kappa^2 \langle \Delta^2 \hat{x}_\alpha \rangle + \frac{\langle \Delta^2 \hat{x}_\beta \rangle}{\kappa^2} + \langle \Delta^2 \hat{u}_\kappa \rangle \right) \\ &= \frac{1}{2} \left(\frac{\kappa^2 + \kappa^{-2}}{4} + \langle \Delta^2 \hat{u}_\kappa \rangle \right) . \end{aligned} \tag{6.26}$$

[4] We use that coherent states have minimum variances $\langle \Delta^2 \hat{x}_{\alpha,\beta} \rangle = \langle \Delta^2 \hat{p}_{\alpha,\beta} \rangle = \frac{1}{4}$

Similarly, $\langle V_\kappa^2 \rangle = \frac{1}{2}\left(\frac{\kappa^2+\kappa^{-2}}{4} + \langle \Delta^2 \hat{v}_\kappa \rangle\right)$, and consequently we find that in the proposed scheme

$$\langle \mathrm{MDIEW}_\kappa \rangle = \frac{1}{2}\left(\frac{\kappa^2+\kappa^{-2}}{2} + \langle \mathrm{EW}_\kappa \rangle\right) . \tag{6.27}$$

The generalization to states that have non-zero averages of the quadratures is obtained by simply offsetting the outputs accordingly as follows: a_1 is the output of $\hat{A}_1 - \langle \hat{x}_A \rangle/\sqrt{2}$, a_2 of $\hat{A}_2 + \langle \hat{p}_A \rangle/\sqrt{2}$ and so on.
The violation of the inequality (6.21) is therefore found for any entangled Gaussian state: indeed, for any such state there exist quadratures and a value of κ such that $\langle \mathrm{EW}_\kappa \rangle < \frac{\kappa^2+\kappa^{-2}}{2}$, which implies from Eq. (6.27) that in the proposed scheme $\langle \mathrm{MDIEW}_\kappa \rangle < \frac{\kappa^2+\kappa^{-2}}{2}$. It is then sufficient to choose σ large enough to violate (6.21).

6.3.3 Two-mode squeezed state case, with noise tolerance

At last, to illustrate the feasibility of our scheme, we apply it to the case of ρ_{AB} being a TMSV state. By including noise tolerance, we pave the way for an experimental realization of our MDI entanglement witness.
A TMSV state can be described as the mixing of two single-mode squeezed states (one squeezed in $\hat{p}$ and one in $\hat{x}$) [63]. In the Heisenberg picture, this results in

$$\begin{aligned} \hat{x}_A &= \frac{e^r \hat{x}_1^{(0)} + e^{-r}\hat{x}_2^{(0)}}{\sqrt{2}}, \quad \hat{p}_A = \frac{e^{-r}\hat{p}_1^{(0)} + e^r \hat{p}_2^{(0)}}{\sqrt{2}}, \\ \hat{x}_B &= \frac{e^r \hat{x}_1^{(0)} - e^{-r}\hat{x}_2^{(0)}}{\sqrt{2}}, \quad \hat{p}_B = \frac{e^{-r}\hat{p}_1^{(0)} - e^r \hat{p}_2^{(0)}}{\sqrt{2}}, \end{aligned} \tag{6.28}$$

where the superscript $\{\hat{x}^{(0)}, \hat{p}^{(0)}\}$ represents operators acting on the vacuum. Consider now the two operators $\hat{u}_{\kappa=1} = \hat{x}_A - \hat{x}_B$ and $\hat{v}_{\kappa=1} = \hat{p}_A + \hat{p}_B$. From Eq. (6.18) we see, by choosing $\kappa = 1$, that these operators satisfy for any separable state $\langle \Delta^2 \hat{u}_{\kappa=1} \rangle + \langle \Delta^2 \hat{v}_{\kappa=1} \rangle \geq 1$. If we compute the above combination for the two-mode squeezed state, we obtain $\hat{u}_{\kappa=1} = \sqrt{2}e^{-r}\hat{x}_2^{(0)}$ and $\hat{v}_{\kappa=1} = \sqrt{2}e^{-r}\hat{p}_1^{(0)}$. Consequently, it holds

$$\langle \mathrm{EW}_{k=1} \rangle_{\mathrm{TMSV}} = e^{-2r} < 1. \tag{6.29}$$

This is not surprising: as soon as there is squeezing $r > 0$ the state is entangled. Substituting this value into expression (6.21), we obtain the entangled score for the MDI witness $\langle \mathrm{MDIEW}_{\kappa=1} \rangle = \frac{1}{2}\left(1 + e^{-2r}\right)$.

To check noise tolerance, we consider losses in the modes A and B modelled as a beam splitter

$$\begin{aligned}\hat{a}_A(\eta_A) &= \sqrt{1-\eta_A}\hat{a}_A(0) + \sqrt{\eta_A}\hat{a}^{(0)}_{NA}\,,\\ \hat{a}_B(\eta_B) &= \sqrt{1-\eta_B}\hat{a}_B(0) + \sqrt{\eta_B}\hat{a}^{(0)}_{NB}\,,\end{aligned} \tag{6.30}$$

where $\hat{a}^{(0)}_{NX}$ is a vacuum mode acting as a noise on mode X, while $\hat{a}_X(0)$ is the corresponding noiseless mode. We focus on a natural scenario in which the source producing the two-mode squeezed state is between Alice and Bob and losses affect the two modes, not necessarily in a symmetric way. At the same time, the same loss noise (6.30) applied to the input coherent states would lead only to a renormalization of α and β and can be compensated by increasing the variance σ. Using Eqs. (6.28) and (6.30) we can accordingly compute

$$\begin{aligned}&\langle \mathrm{EW}_\kappa\rangle_{\mathrm{TMSV},\eta_A,\eta_B} = \\ &\left\langle \Delta^2\left(\kappa\hat{x}_A(\eta_A) - \frac{\hat{x}_B(\eta_B)}{\kappa}\right) + \Delta^2\left(\kappa\hat{p}_A(\eta_A) + \frac{\hat{p}_B(\eta_B)}{\kappa}\right)\right\rangle = \\ &\frac{1}{2}\left(\kappa^2\eta_A + \frac{\eta_B}{\kappa^2}\right) + \frac{e^{2r}}{4}\left(\kappa\sqrt{1-\eta_A} - \frac{\sqrt{1-\eta_B}}{\kappa}\right)^2 + \frac{e^{-2r}}{4}\left(\kappa\sqrt{1-\eta_A} + \frac{\sqrt{1-\eta_B}}{\kappa}\right)^2.\end{aligned} \tag{6.31}$$

Notice that it is always possible to choose κ such that $\kappa\sqrt{1-\eta_A} - \kappa^{-1}\sqrt{1-\eta_B} = 0$ and r big enough to nullify the last term of (6.31), and obtain a score (6.27) $\langle \mathrm{MDIEW}_\kappa\rangle = \frac{1}{4}\left(\kappa^2(\eta_A^2+1) + (\eta_B^2+1)\kappa^{-2}\right)$, which is lower than the separable bound $\frac{1}{2}\left(\kappa^2+\kappa^{-2}\right)\left(\frac{\sigma^2}{1+\sigma^2}\right)$ for large enough σ. In this sense the entanglement witness we analysed here is loss-resistant. In Figure 6.3 we plot the trade-off between noise, entanglement, and variance of the prior distributions σ for the MDI detection of entanglement in the case of symmetric losses ($\eta_A = \eta_B$, $\kappa = 1$).

6.4 Other MDI certification tasks: quantum memory verification

The MDI framework can be used, in general, as a scenario for quantum certification with minimal assumptions. Besides the case of entanglement certification presented above, such framework has been used as well for quantum key distribution [207, 208], quantum secure communication [222], and quantum randomness certification [211].

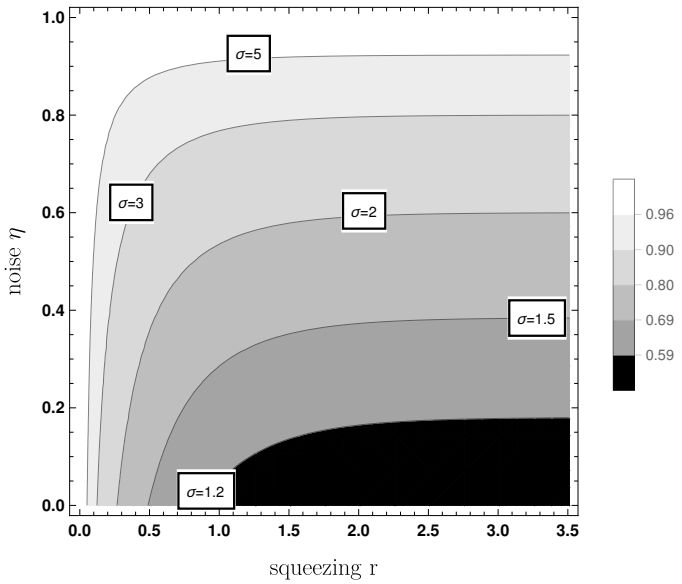

Figure 6.3: Contourplot of the obtainable value (6.27) of $\langle \mathrm{MDIEW}_{\kappa=1} \rangle$, for a two-mode squeezed vacuum state with squeezing parameter r, under the presence of losses with parameter $\eta_A = \eta_B \equiv \eta$ (cf. (6.31)). The contours are chosen to match the separable bound (6.21) for different values of σ, which corresponds to the width of the Gaussian prior used in the experiment. Therefore the area under each σ-contour defines the range of parameters for which MDI entanglement is certified.

Another task that is suited to the MDI scenario is that of *quantum memory verification*. The idea of quantum memory certification with minimal assumptions was introduced in Ref. [223]. Specifically, the problem is that of a honest party A wanting to test the quantum memory given by an untrusted provider E. Given the kind of task, it is natural to have *trusted quantum inputs*, i.e. the task is inherently suited for the class of measurement-device-independent protocols. Alice can send quantum inputs to Eve, who claims to have a "good" quantum memory, which can be seen as a channel $\mathcal{N}$, that maps the input state ϕ to $\mathcal{N}[\phi]$ after some time τ (i.e. the storage time). Ideally, the most desirable property of a memory is the fidelity between input and output. However, a unitary channel $\mathcal{N}[\phi] = U\phi U^\dagger$ is arguably as good as a memory as the identity channel $\mathcal{N} = \mathbb{1}$. In fact, unitary transformations preserve the information contained in the input state ϕ, and can be corrected by applying the inverse unitary. Instead, the fundamental property of a quantum memory is that of preserving "the quantumness" of what it stores, such as the entanglement of the states that are sent to it. More precisely, it is known that a channel is entanglement-breaking if and only if it can be simulated by *measure & prepare* transformations, that is, measuring the ϕ and preparing a new state according to the measurement result [223]. Such channels therefore only require a classical memory storing the measurement results.

Therefore, the main property that a quantum memory should satisfy is to be a non-entanglement-breaking channel, and the task of quantum memory verification can therefore be defined as the MDI certification of non-entanglement-breaking channels.

The authors in [223] showed constructively, in the discrete variable case, that any non-entanglement-breaking channel can be in fact certified in the MDI scenario. We sketch the protocol as follows: Alice sends to Bob a state φ at time $t = 0$, and another state ϕ at time $t = \tau$. Eve can perform any measurement on φ and ϕ, but given the time delay, she is forced to use $\mathcal{N}[\varphi]$ in case she wants to perform a joint measurement. If Eve performs projections on the maximally entangled state, then Alice can certify if $\mathcal{N}$ is non-entanglement breaking, just out of the statistics $P(b|\varphi, \phi)$. Let's say in fact that Eve's POVM is

$$M^{b=0} = |\psi^+\rangle \langle\psi^+| \ , \tag{6.32}$$

$$M^{b=1} = \mathbb{1} - |\psi^+\rangle \langle\psi^+| \ , \tag{6.33}$$

where ψ^+ is the maximally entangled state $|\psi^+\rangle = \frac{1}{\sqrt{d}} \sum_i^d |ii\rangle$. Then, one has

$$P(b = 0|\varphi, \phi) = \frac{1}{d} \mathrm{Tr}[\phi^T \mathcal{N}[\varphi]] \ . \tag{6.34}$$

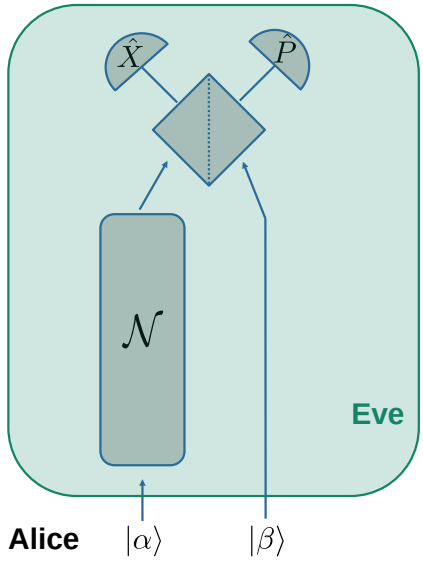

Figure 6.4: Alice sends random coherent states $|\alpha\rangle$, and, with a delay, $|\beta\rangle$. Eve, after conserving $|\alpha\rangle$ in the memory $\mathcal{N}$, measures jointly $\frac{\hat{x}_\alpha+\hat{x}_\beta}{\sqrt{2}}$ and $\frac{\hat{p}_\alpha-\hat{p}_\beta}{\sqrt{2}}$, via a beamsplitter and homodyne detection.

That is, by varying ϕ and φ it is possible to do tomography of $\mathcal{N}$ and verify if it is non-entanglement breaking. This is the underlying idea, which can be translated into a proper witness (see details in [223]).
The first experimental MDI verifications of quantum memories have been performed for the discrete variables case [224–226].

6.4.1 MDI certification of a continuous variables quantum memory

We propose a MDI protocol for the verification of continuous variable quantum memories, inspired both by [223] and [3], where we have seen that in the context of continous variables there is a strong connection of MDI certification to metrology, which we exploit again in the following proposal. The detailed study will be the object of a forthcoming work [227]. Here we briefly outline the proposal and the idea behind the security proof.
The setup is simple, as in Figure 6.4. Alice sends a random coherent state $|\alpha\rangle$, and after some time, $|\beta\rangle$. If the memory is perfect, Eve can measure the quadratures

$$\hat{X} = \frac{\hat{x}_\alpha + \hat{x}_\beta}{\sqrt{2}}\,, \quad \hat{P} = \frac{\hat{p}_\alpha - \hat{p}_\beta}{\sqrt{2}}\,, \tag{6.35}$$

via mixing the two inputs in a balanced beam splitter. This measurement can be seen as an estimate of the corresponding linear combination of the complex numbers defining the coherent states. That is, the average values of $\hat{X}$ and $\hat{P}$ are, respectively, $\frac{\alpha_x+\beta_x}{\sqrt{2}}$ and $\frac{\alpha_p-\beta_p}{\sqrt{2}}$, and the variance associated to both measurements (6.35) is the vacuum standard noise, equal $1/4$ in our normalization choice (cf. [63]), $\Delta^2\hat{X} = \Delta^2\hat{P} = \frac{1}{4}$. The measurement (6.35) is an entangled measurement on the two incoming modes. For this reason, we propose a witness based on such estimation. That is, we input random α and β, and we ask Eve two outputs ξ_x and ξ_p representing Eve's best estimate of $\frac{\alpha_x+\beta_x}{\sqrt{2}}$ and $\frac{\alpha_p+\beta_p}{\sqrt{2}}$. We thus compute as a score the average value of Eve's estimate error, i.e.

$$\text{score:=} \left\langle \left(\xi_x - \frac{\alpha_x+\beta_x}{\sqrt{2}}\right)^2 + \left(\xi_p - \frac{\alpha_p-\beta_p}{\sqrt{2}}\right)^2 \right\rangle . \tag{6.36}$$

As we chose units in which the vacuum has variance $= \frac{1}{4}$ on both quadratures, if the memory is perfect $\mathcal{N} = \mathbb{1}$ the resulting score will have a value of $1/2$,

$$\text{score}(\mathcal{N} = \mathbb{1}) = \Delta^2\left(\frac{\hat{x}_\alpha + \hat{x}_\beta}{\sqrt{2}}\right) + \Delta^2\left(\frac{\hat{p}_\alpha - \hat{p}_\beta}{\sqrt{2}}\right) = \frac{1}{2} . \tag{6.37}$$

Now, if instead the channel $\mathcal{N}$ is entanglement breaking, this means that it corresponds to a measure & prepare channel. Intuitively, the estimation of the above quantity will then need to go through all the four values of $\alpha_x, \beta_x, \alpha_p, \beta_p$, doubling the error. This is the main idea of the MDI proof in this case. That is, the optimal strategy if $\mathcal{N}$ is entanglement breaking, becomes measuring $\frac{\hat{x}_\alpha}{\sqrt{2}}$ and $\frac{\hat{p}_\alpha}{\sqrt{2}}$, and separately the same for $\frac{\hat{x}_\beta}{\sqrt{2}}$ and $\frac{\hat{p}_\beta}{\sqrt{2}}$. By doing so, the average error, i.e. the score (6.36) will be $4 * \frac{1}{4} = 1$.

$$\text{score}(\mathcal{N} \text{ ent.-breaking}) \gtrsim 1 . \tag{6.38}$$

What we presented until now is only the idea of the experimental proposal, without formal proofs, which will be given in a forthcoming work [227].

6.5 Discussion

We discussed different measurement-device-independent (MDI) protocols for the certification of quantum properties, such as entanglement, in experimental scenarios involving a minimal level of trust in the devices employed. This scenario consists and assuming only a precise description of the quantum states used as inputs of the experiment, while leaving the rest uncharacterised. In

particular, extended previous results and found new practical witnesses for the case of continuous-variable systems, showing how such framework is well suited to implementations in systems using e.g. Gaussian quantum optics. The implementation of our protocols for entanglement witnessing [3] and quantum memory verification [227] is feasible with current technology and is being investigated at the moment.

Our work also opens up a series of interesting directions. From a general perspective, our results opens the path to the use of CV quantum systems for MDI tasks beyond entanglement detection, such as randomness generation [211], secure communication [208, 222], or witnessing of indefinite causal order [228]. Another possible research direction would be to investigate the possible generality of the connection between entanglement detection and metrology exploited in [3]. In particular, can all MDI entanglement tests be translated into a parameter estimation problem?

OUTRO

Conclusions

In this thesis, we theoretically analysed the optimization of a series of *operational tasks* involving the use of *quantum resources*. We considered a wide range of protocols with different objectives, from measuring temperature with the best possible precision, to witnessing entanglement of optical states with minimal assumptions, to implementing efficient thermal cycles for refrigeration or energy extraction, et cetera.
As we argued in the "Motivation" Chapter, analysing these tasks to their limits can bring additional knowledge to the theory itself. This is in fact one of the classic tools of theoretical physics, i.e. that of using *gedankenexperiments* to gain understanding in our description of nature. This is arguably the case of this PhD thesis, too. In it, we could appreciate different theoretical insights, some of which we list here:

- While studying thermal engines, we were reminded once again of the central role of the heat capacity as a central quantifier that characterizes equilibrium properties and close-to-equilibrium properties of a working fluid.

- In the context of thermometry, we found out that it is possible to achieve an Heisenberg-like scaling of the estimation precision of temperature, which not only can be obtained with physically sound interacting systems, but also without any use of entanglement or other quantum properties.

- For what concerns the study of memory effects and non-Markovianity witnessing, our results seem to suggest that the natural geometry of the set of states should be the one given by the Fisher Information, thanks to which it is also possible to give a more intuitive notion of information backflow in an open dynamics.

- When studying network nonlocality of optical states, we contributed to answering a physical question that has always sparked debate in the quantum foundations community: can a single-particle entangled state show nonlocality with particle-detection measurements only?

- While investigating the detection of entanglement with minimal assumptions, we learnt how different protocols for the certification of quantum properties can be translated to problems of adversarial metrology, and viceversa.

Geometry played a central role in our work, facilitating the optimization of practical tasks, emerging both as a natural tool for mathematical analysis, and in the resulting optimal protocols.
Among the geometrical objects we studied, we find remarkable how the Fisher Information metric recurrently appears in the study of subjects that are, a priori, not related, such as the thermodynamic length metric (cf. Chap. 1 and App. A), the geometry of non-Markovianity witnessing (Chap. 4), as well as the metrological bounds that can be used to infer the best precision of a thermometer (Chap. 3) or the capacity of parameter estimation for entangling and non entangling measurements (Chap. 6).

The outlooks of our research have multiple directions, which reflect the variety of subjects studied. For topic-specific perspectives we refer the reader to the "Discussion" section at the end of each Chapter. There is however, a number of general, speculative questions that remain open. Among them:

- Together with Ref. [39], our result in [1] show that the main figure of merit of a thermal engine or refrigerator is given by the ratio between heat capacity and relaxation timescale, $\mathcal{C}/\tau_{\mathrm{rel}}$. When considering i.i.d. equilibrium thermometry, the main figure of merit becomes instead $\nu\mathcal{C}$ (see Chap. 3), where ν is the number of repetitions of the experiment. One might therefore be tempted to observe the following: the number of reads of a thermalized themometers is governed by the total time τ at disposal, divided by τ_{rel}. That is, $\nu \sim \tau/\tau_{\mathrm{rel}}$. This means that, when time becomes a resource, in the context of equilibrium thermometry, the main figure of merit becomes the same as that of thermal machines. Is this a coincidence? Are the tasks of energy extraction and temperature estimation related more intimately than what we know?

- Our results in Chapter 3 show that it is possible to engineer heat capacities $\mathcal{C}$ that scale quadratically in the size of the system. Is it possible to obtain a similar result when considering the whole ratio $\mathcal{C}/\tau_{\mathrm{rel}}$, in physically-sound systems? This would have consequences in all the above-mentioned problems. At the same time, it is an extremely challenging problem, as the parameter optimization involves also the interactions with the thermal reservoir, which dictate τ_{rel}.

- Our results show that, in the context of low dissipation thermal machines and equilibrium thermometry, no quantum advantage is observed. Rather, quantum effects seem to be detrimental. Are there reasonable tasks in thermodynamics where a quantum advantage exists and is effective? The most notable signatures of quantum advantage in thermal machines have been so far noted in limits of small action / small power (see e.g. [229] or [230]), which might therefore be seen as of minor importance.

- In the quest of quantum advantages in thermodynamics, can the advantage derive from nonlocal properties of quantum sources, perhaps in networks? Is there a thermodynamic protocol which is able to show a quantum advantage in a semi-device-independent way?

- What is the most general description of non-Markovianity outside the stochastic-divisibility framework? Can it be still described in terms of metric contraction/dilation?

- Is it possible to identify the most general tasks which can be certified between an honest user – which can only prepare and send trusted quantum inputs – and an untrusted quantum provider? Does this class coincide with a class of adversarial metrology tasks?

Finally, thus far we have stressed how our analysis of operational tasks can give new insights and understanding in theoretical Quantum Physics. It is evident, however, that the natural continuation of many of the results presented in this thesis will be their experimental demonstration, as well as their application to real-life useful protocols. In this sense, we can anticipate that a part of our results is currently under consideration for realization.

Acknowledgements

This thesis is the academic result of four eventful years that I spent (mostly) in Barcelona. During this time I have had the fortune to meet many people, friends and colleagues, arguably too many to thank properly in few lines here, and for that I apologize in advance.

I would like to start by thanking my academic mentors of these years:
Grazie Toni, per avermi dato tutta la libertá e tutte le opportunitá di cui potessi avere bisogno in questi anni, e per l'ambiente allegro e sereno che mantieni nel gruppo.
Moltes gràcies Martí, por ser mi "padrino académico" en estos años, por acogerme en Ginebra, por los consejos, y por todas las actividades juntos.

My time at ICFO has been truly happy, and I do believe that ICFO is an extraordinary place, excellent on the scientific point of view, and exceptionally enjoyable at the human level.
I was lucky enough too meet good friends on my very first day, Korbinian and Arturo, together with whom I started this journey and shared much more.
I am particularly grateful to the "old QIT group" which made me feel part of the family *literally* since the day I arrived, among them Joe, Alexia, Alex, Moha, Flavio, Ivan, Matteo1... I will never thank you enough. More in general, I really appreciate having spent time with all of the people who formed part of the QIT group during these years and now.

Some special thanks are due to very special co-authors: Matteo2, Dario, Jacopo, Tamás, PaoloE, Cristian. It has been a pleasure to share time with you, both in and out of the office.

Finally, I could not function without the constant support of my family and all of my friends in Italy and elsewhere.
Grazie a Mamma e Papà come sempre, da sempre.
Grazie a Pierluigi ed Anna per le innumerevoli ospitate tra un viaggio e l'altro.
Grazie a tutti i miei amici.
Grazie Erika, di tutto ♡.

Merci a totes i tots! Adéu i hasta pronto :)

Appendix A

Slow-driving derivation of the thermodynamic length

As we mentioned in Chap. 1, the thermodynamic length metric can be derived in different dynamical frameworks. In Ref. [2] we review some of them. In this Appendix, we show explicitly how to derive explicitly such metric in the case of open quantum systems [35], assuming standard Markovian dynamics and following approach of Ref. [29].
Consider a *slow* isothermal process where the Hamiltonian of the working fluid $\mathcal{S}$ is externally controlled by some experimental parameters $\vec{\lambda}_t$ – we will write for simplicity $H_t \equiv H_{\vec{\lambda}_t}$ – that is

$$H_t \ , \quad t \in [0, \tau] \ , \tag{A.1}$$

while in contact with a thermal bath at temperature $T = 1/\beta$. In order to characterise the process beyond the quasistatic limit, we need to assume some structure on the thermalization processes of $\mathcal{S}$ induced by the reservoir. In a rather generic scenario we consider that the relaxation of the quantum working fluid can be described by a master equation with the thermal state as a unique fixed point [51], i.e.

$$\dot{\rho}_t = \mathcal{L}_t[\rho] \ , \tag{A.2}$$

$$\mathcal{L}_t[\omega_t] = 0 \ . \tag{A.3}$$

Here and throughout the Appendix we use the shorthand subscript $_t$ to indicate any time dependence by $H_t \equiv H(\vec{\lambda}_t)$. In particular, notice that in general $\mathcal{L}_t$ depends on time through H_t, and

$$\omega_t \equiv \frac{e^{-\beta H_t}}{\mathrm{Tr}[e^{-\beta H_t}]} \ . \tag{A.4}$$

Following [29], the solution of this master equation can be found perturbatively around $1/\tau$. By writing the expansion

$$\rho_t = \rho_t^{(0)} + \rho_t^{(1)} + \rho^{(2)} + \dots \qquad \rho_t^{(i)} \sim \mathcal{O}(\tau^{-i}) \,, \tag{A.5}$$

it is possible to solve the dynamical equation by noticing that $\dot{\rho}_t^{(i)}$ is of order $\mathcal{O}(\tau^{-(i+1)})$ and starting the perturbative expansion around the quasistatic equilibrium solution

$$\rho_t^{(0)} = \lim_{\tau\to\infty} \rho_t = \omega_t \,. \tag{A.6}$$

That is, given

$$\dot{\rho}_t = \dot{\rho}_t^{(0)} + \rho_t^{(1)} + \dots = \mathcal{L}_t[\rho_t^{(0)} + \rho_t^{(1)} + \dots] \,, \tag{A.7}$$

And matching the same-order equations

$$\dot{\rho}_t^{(i)} = \mathcal{L}_t[\rho_t^{(i-1)}] \,, \tag{A.8}$$

it follows that the first order correction in ρ_t is

$$\rho_t^{(1)} = \mathcal{L}_t^{-1}[\dot{\rho}_t^{(0)}] = \mathcal{L}_t^{-1}[\dot{\omega}_t] \,. \tag{A.9}$$

Notice that in general $\mathcal{L}_t$ is not an invertible operator. It is however invertible when restricted to the subspace of density matrices having null trace (such as is $\dot{\omega}_t$), assuming that is has a unique steady state, as in the case of standard thermalizations (see [29] for the technical details). We will therefore use an abuse of notation indicating with $\mathcal{L}_t^{-1}$ the pseudoinverse of $\mathcal{L}_t$ on the traceless subspace of density matrices.
Plugging this expression into $W = \int_0^\tau \mathrm{d}t\ \mathrm{Tr}[\rho_t \dot{H}_t]$ it is possible to compute the corrections to the quasistatic limit i.e.

$$W^{(i)} = \int_0^\tau \mathrm{d}t\ \mathrm{Tr}[\rho_t^{(i)} \dot{H}_t] \,, \tag{A.10}$$

which immediately yields the quasistatic results

$$W^{(0)} = \int_0^\tau \mathrm{d}t\ \mathrm{Tr}[\omega_t \dot{H}_t] = \Delta F \tag{A.11}$$

and the first order correction

$$W^{(1)} = \int_0^\tau \mathrm{d}t \ \mathrm{Tr}[\mathcal{L}_t^{-1}[\dot{\omega}_t]\dot{H}_t] \ . \tag{A.12}$$

The limit of slow-driving [29] corresponds to neglecting all higher orders, and identifying $W^{(1)}$ as the irreversible work dissipated (cf. Chap. 1)

$$W - \Delta F = W^{(1)} = \frac{T\sigma}{\tau} \ , \tag{A.13}$$

where we chose the standard parametrization that factorizes the temperature T and the total duration τ of the transformation. It follows that the irreversible entropy production per unit time, expressed by the coefficient σ, can be expressed as

$$\sigma = \beta \int_0^1 \mathrm{d}s \ \mathrm{Tr}[\mathcal{L}_s^{-1}[\omega_s']H_s'] \ , \tag{A.14}$$

where we slightly Abiused the notation by using the adimensional time $s = t/\tau$, i.e. $A_s \equiv A_{s\tau}$ for any quantity A, as well as

$$A_s' \equiv \frac{\partial}{\partial s} A_s = \tau \dot{A}_{s\tau} \ \text{, for any quantity } A \ . \tag{A.15}$$

From now on, we will avoid writing the time s when clear from the context, and we will use the Hamiltonian in units of temperature, i.e.

$$G_s \equiv \beta H_s \ . \tag{A.16}$$

It follows that

$$\sigma = \int_0^1 \mathrm{d}s \ \mathrm{Tr}[\mathcal{L}^{-1}[\omega']G'] \quad \text{and} \quad \omega = \frac{e^{-G}}{\mathrm{Tr}[e^{-G}]} \ . \tag{A.17}$$

We can then use the formula for the derivative of an exponential,

$$\omega' = -\int_0^1 \mathrm{d}y \ \omega^{1-y}(G' - \mathrm{Tr}[\omega G'])\omega^y \equiv -\mathfrak{J}_\omega[G'], \tag{A.18}$$

where we defined the operator $\mathfrak{J}_\omega$.
σ can then be reexpressed in the convenient quadratic form:

$$\sigma = -\int_0^1 \mathrm{d}s \ \mathrm{Tr}[G'\mathcal{L}^{-1}[\mathfrak{J}_\omega[G']]] \ . \tag{A.19}$$

Considering the control parameters that drive G_s, e.g. in the form

$$G_s = \sum_j \lambda_j(s) X_j \,, \tag{A.20}$$

(where X_j is a vector of observables of the system) it follows that the above quadratic form (A.19) if projected to a metric m_{ij} onto the control vector $\vec{\lambda}$ as

$$\sigma = \sum_{ij} \int_0^1 \mathrm{d}s \; \lambda_i' m_{ij} \lambda_j' \tag{A.21}$$

with

$$m_{ij} = -\frac{1}{2} \left(\mathrm{Tr}[X_i \mathcal{L}_s^{-1} \mathfrak{J}_\omega[X_j] + X_j \mathcal{L}_s^{-1} \mathfrak{J}_\omega[X_i]] \right) . \tag{A.22}$$

The matrix m_{ij} is symmetric, positive-definite due to the second law $\mathrm{d}\sigma \geq 0$, and it depends smoothly on the base point ω; hence it defines a metric.

A.1 Thermodynamic metric for single or multiple time scales

As the general expression (A.19)-(A.22) of the thermodynamic length metric depends on the dynamical generator $\mathcal{L}_s$, in this section we consider the simple case of an exponential relaxation of ρ to equilibrium with a single time-scale, as described by the Lindbladian:

$$\mathcal{L}_s[\rho_s] = \tau_{\mathrm{eq}}^{-1}(\omega_s - \rho_s) \tag{A.23}$$

which has pseudo-inverse $\mathcal{L}_s^{-1}[.] = \tau_{\mathrm{eq}}(\omega_s \mathrm{Tr}(.) - \mathbb{I})$ that acts multiplicatively on the subspace of traceless operators. In this case, using that $\mathrm{Tr}[\mathfrak{J}_\omega[\rho]] = 0 \; \forall \rho$ one finds that σ in (A.19) is given by:

$$\sigma = \tau_{\mathrm{eq}} \int_0^1 \mathrm{d}s \; \mathrm{Tr}[G' [\mathfrak{J}_\omega[G']]] \,. \tag{A.24}$$

Notice that although we have assumed that the whole state ρ converges to equilibrium as in (A.23), strictly speaking it is only necessary that the driven observables (the X_j in (A.20)) converge to equilibrium with a single time-scale, i.e., $\langle \dot{X}_i \rangle_\rho = \tau_{\mathrm{eq}}^{-1}(\langle X_i \rangle_\omega - \langle X_i \rangle_\rho)$, with $\langle X \rangle_\rho = \mathrm{Tr}(X\rho)$. This is especially relevant in complex systems (e.g. many-body systems), where the full dissipative dynamics can be extremely complex but the equilibration of some macroscopic observables can be well described by an exponential relaxation with a suitable

time-scale. In this sense, it is also worth pointing out that if each generalised observable X_i decays with a different time-scale τ_i

$$\langle \dot{X}_i \rangle_\rho = \tau_i^{-1}(\langle X_i \rangle_\omega - \langle X_i \rangle_\rho) \,, \tag{A.25}$$

then the metric above can be easily generalised as [35]:

$$m_{ij} = \frac{\tau_i + \tau_j}{2} \frac{\partial^2}{\partial\lambda_i \partial\lambda_j} \ln \mathcal{Z}, \quad \mathcal{Z} \equiv \mathrm{Tr}[e^{-G}] \,, \tag{A.26}$$

$\mathcal{Z}$ being the partition function and where we have absorbed the dependence on τ_i in the metric and where τ_i can in principle depend on the point of the trajectory.

Connection to Fisher Information metric In the above analysis, we expressed the dissipation σ as a quadratic form in terms of the Hamiltonian change $G' = \beta H'$, as in (A.19). Notice however, how the same quantity can be expressed as a function of the thermal state variation ω', i.e. using the definition $\omega' = -\mathfrak{J}_\omega[G']$ (A.18), from (A.17) it follows

$$-\sigma = \int_0^1 \mathrm{d}s \; \mathrm{Tr}[\mathfrak{J}_\omega^{-1}[\omega']\mathcal{L}^{-1}\omega'] \tag{A.27}$$

Notice that both $\mathfrak{J}_\omega$ and $\mathcal{L}$ are invertible on the subspace of traceless operators, and that ω' is traceless. Without loss of generality we can assume also G to have null average value $\mathrm{Tr}[\omega G]$, as it does not influence the definition of ω (A.17). In such case we find

$$\omega' = -\int_0^1 \mathrm{d}x \; \omega^{1-x} G' \omega^x \equiv -\mathbb{J}_\omega[G'], \tag{A.28}$$

where $\mathbb{J}_\omega[.] = \mathfrak{J}_\omega[.] + \omega \mathrm{Tr}[\omega(.)]$. In the simple relaxation case in which $\mathcal{L}_s[\rho_s] = \tau_{\mathrm{eq}}^{-1}(\omega_s - \rho_s)$, we find

$$\sigma = \tau_{\mathrm{eq}} \int_0^1 \mathrm{d}s \; \mathrm{Tr}[\omega' \mathbb{J}_\omega^{-1}[\omega']] \,. \tag{A.29}$$

The operator $\mathcal{J}_\omega$ corresponds to the Kubo-Mori-Bogoliubov inner product, and the integrand $\mathrm{Tr}[\mathrm{d}\omega \mathbb{J}_\omega^{-1}[\mathrm{d}\omega]]$ corresponds to the infinitesimal square length associated to $\mathrm{d}\omega$ according to the associated Quantum Fisher metric (cf. Appendix C). In the limit of simple exponential relaxation, we thus find that the dissipation σ of the transformation is equivalent to the square distance that is travelled by the thermal state, starting at $\omega_{s=0}$ to $\omega_{s=1}$.

A.1.1 Thermodynamic metric on standard microscopical models

In this section, we apply our general considerations to derive thermodynamic metrics for systems described by quantum master equations that can be derived from standard physical models (particularly, for detailed-balanced dynamics). Before, let us discuss here some generic properties. Following [29,51], we have in standard scenarios where a quantum system is coupled to a bosonic bath:

$$\mathcal{L}[\rho] = \sum_{\nu>0} \gamma_0 \nu^\theta \left((N(\beta\nu)+1)\mathcal{D}_{A_\nu}[\rho] + N(\beta\nu)\mathcal{D}_{A_\nu^\dagger}[\rho] \right) \tag{A.30}$$

where $\gamma_0\nu^\theta$ is the spectral density of the bath (θ defines the Ohmicity),

$$\mathcal{D}_X[\rho] = X\rho X^\dagger - \frac{1}{2}(X^\dagger X\rho + \rho X^\dagger X) \tag{A.31}$$

and

$$N(\beta\nu) = \frac{1}{e^{\beta\nu}-1} \tag{A.32}$$

is the Bose-Einstein distribution (one can consider ferimonic baths by replacing $N(\beta\nu)$ by the Fermi distribution). The general form (A.30) of $\mathcal{L}$ depends on H_s, β and the spectral density $J(\nu) = \gamma_0\nu^\theta$. Consider now a Carnot cycle with two baths at temperature $\beta_{\rm c}$ and $\beta_{\rm h}$, and the action of each bath being described by a $\mathcal{L}(H_S, \beta, J(\nu))$. Assuming that the spectral density of both baths is the same, then both Lindbladians are related by the transformation $\beta \to \lambda^{-1}\beta$ and $H_S \to \lambda H_S$, with $\beta \equiv \beta_{\rm c}$ and $\lambda \equiv \beta_{\rm c}/\beta_{\rm h}$. In terms of the Lindbladian, note that [29]:

$$\mathcal{L}(\lambda H_S, \lambda^{-1}\beta) = \lambda^\theta \mathcal{L}(H_S, \beta), \tag{A.33}$$

which shows how the generators of the dynamics are related between the cold and hot isotherm, i.e.

$$\mathcal{L}((\beta_{\rm c}/\beta_{\rm h})H_S, \beta_{\rm h}) = \left(\frac{\beta_{\rm c}}{\beta_{\rm h}}\right)^\theta \mathcal{L}(H_S, \beta_{\rm c})\ . \tag{A.34}$$

After noticing that the dissipation is related to $\mathcal{L}$ through Eq. (A.19), we can write a simple proportionality relation for the metric describing the dissipation

on the hot and cold isotherm,

$$m_{ij}^{(\mathrm{h})} = \left(\frac{T_\mathrm{c}}{T_\mathrm{h}}\right)^{\theta} m_{ij}^{(\mathrm{c})} \ , \tag{A.35}$$

which implies, for time-reversal symmetric protocols described in the main text,

$$\sigma_\mathrm{h} = \left(\frac{T_\mathrm{c}}{T_\mathrm{h}}\right)^{\theta} \sigma_\mathrm{c} \ . \tag{A.36}$$

Appendix B

Proof of Lemma 1 (Chap. 3).

In this Appendix we prove the theoretical Lemma 1 presented in the main text in Chapter 3. The lemma considers two Hamiltonians, H_1 and H_2, such that H_1 has a single ground state and a k_1-degenerate excited state (k_1+1 levels in total), while H_2 has the same spectrum and additional k_2 excited states above (totaling $1+k_1+k_2$ levels) that is,

$$H_1 = 0\,|0\rangle\langle 0| + \sum_{i=1}^{k_1} E\,|i\rangle\langle i| \ , \tag{B.1}$$

$$H_2 = 0\,|0\rangle\langle 0| + \sum_{i=1}^{k_1} E\,|i\rangle\langle i| + \sum_{\alpha=k_1+1}^{k_1+k_2} E_\alpha\,|\alpha\rangle\langle\alpha| \ , \tag{B.2}$$

with $0 \leq E \leq E_\alpha \ \forall\alpha$. Consider now the realistic situation in which these Hamiltonians are controlled via internal coupling parameters $\vec{\lambda}$, such as is the case of our work. The lemma has two assumptions: *i)* it is possible to control the first excited gap $E(\vec{\lambda})$, contemporary to keeping the additional α-levels above *ii)* $E_\alpha(\vec{\lambda}) \geq E(\vec{\lambda})$. A simple scenario in which these assumptions are satisfied is the simple Hamiltonian

$$H_2(\lambda) = \lambda\left(\sum_{i=1}^{k_1}\epsilon\,|i\rangle\langle i| + \sum_{\alpha=k_1+1}^{k_1+k_2}\epsilon_\alpha\,|\alpha\rangle\langle\alpha|\right) \ . \tag{B.3}$$

Under assumptions *i)* and *ii)*, the lemma states that the maximum heat capacity obtainable with H_2 is always larger than the maximum heat capacity obtainable with H_1.

$$\max_E \mathcal{C}(H_1) \leq \max_{E\leq E_\alpha} \mathcal{C}(H_2) \ . \tag{B.4}$$

Proof of the Lemma. When computing the variance of the energy in a thermal state, global shifts in the energy do not matter. For this reason we rewrite the same Hamiltonians putting the k_1 $|i\rangle$ levels to zero, i.e.

$$H_1' = -E\,|0\rangle\langle 0|\ , \tag{B.5}$$

$$H_2' = -E\,|0\rangle\langle 0| + \sum_\alpha E_\alpha'\,|\alpha\rangle\langle\alpha|\ . \tag{B.6}$$

with $E \geq 0$ and $E_\alpha - E \equiv E_\alpha' \geq 0$.
We now use temperature units $\beta = 1$, to simplify the discussion. The thermal states are therefore

$$\rho^{(1)} = \frac{e^{-H_1'}}{\mathrm{Tr}[e^{-H_1'}]}\ , \quad \rho^{(2)} = \frac{e^{-H_2'}}{\mathrm{Tr}[e^{-H_2'}]}\ . \tag{B.7}$$

Let us call $p_0^{(1)}$ and $p_0^{(2)}$ the corresponding ground state populations,

$$p_0^{(1)} = \frac{1}{1+k_1 e^{-E}}\ , \quad p_0^{(2)} = \frac{1}{1+k_1 e^{-E}+\sum_\alpha e^{-(E_\alpha'+E)}}\ . \tag{B.8}$$

Notice that they both depend on the value of E, which we omit in the following for simplicity. It is easy to compute the heat capacity (equivalently, the energy variance) as

$$\Delta^2 H_1 = E^2\, p_0^{(1)}(1-p_0^{(1)})\ . \tag{B.9}$$

For what concerns H_2 instead (calling p_α the population of the level E_α)

$$\begin{aligned}\Delta^2 H_2 =& E^2 p_0^{(2)} + \sum_\alpha E_\alpha'^2 p_\alpha - \left(-E\,p_0^{(2)} + \sum_\alpha E_\alpha' p_\alpha\right)^2 \\ =& E^2 p_0^{(2)}(1-p_0^{(2)}) + A + B\ ,\end{aligned} \tag{B.10}$$

with

$$A := \sum_\alpha E_\alpha'^2 p_\alpha - \left(\sum_\alpha E_\alpha' p_\alpha\right)^2 \geq 0\ , \tag{B.11}$$

$$B := 2E p_0^{(2)} \left(\sum_\alpha E_\alpha' p_\alpha\right) \geq 0\ . \tag{B.12}$$

where the inequalities follow from B being trivially positive, while A corresponds to the variance of an Hamiltonian having levels E_α with population

p_α and all the rest of the population $1 - \sum_\alpha p_\alpha$ being at an energy $= 0$. It follows that

$$\Delta^2 H_2 \geq E^2 p_0^{(2)}(1 - p_0^{(2)}) \; . \tag{B.13}$$

Finally, let's compare the maximal value of the heat capacity in the two cases. Let's call $\bar{E}^{(1)}$ the optimal value for the Hamiltonian H_1, which induces a ground state population equal to $\bar{p}_0^{(1)} = p_0^{(1)}(\bar{E}^{(1)})$, i.e.

$$\max \Delta^2 H_1 = \max_E E^2 p_0^{(1)}(E)(1 - p_0^{(1)}(E)) = (\bar{E}^{(1)})^2 \bar{p}_0^{(1)}(1 - \bar{p}_0^{(1)}) \; . \tag{B.14}$$

It suffices to conclude now by noticing that $p_0^{(2)}(E)$ is an increasing function of E and is always smaller than $p_0^{(1)}(E)$, cf. Eq. (B.8). This means that

$$p_0^{(2)}(x) = p_0^{(1)}(y) \quad \text{implies} \quad x > y \; . \tag{B.15}$$

Therefore one can choose $p_0^{(2)}(E) = \bar{p}_0^{(1)}$ which will lead to $E > \bar{E}$ and therefore from (B.13)

$$\max \Delta^2 H_2 \geq (\bar{E}^{(1)})^2 \bar{p}_0^{(1)}(1 - \bar{p}_0^{(1)}) \; , \tag{B.16}$$

concluding the proof.

Appendix C

Fisher distance and Fisher metric

C.1 Classical Fisher metric and Fisher information

The Fisher distance is a fundamental object in statistical and mathematical physics. Here we first consider the case of classical systems, more precisely with a finite dimension D. This means that the state of the system can be described by a probability vector $p_i \quad i = 1, \dots, D$. The set of states is therefore the $D-1$-dimensional symplex in $\mathbb{R}^D$ defined as

$$\mathcal{S}(\mathbb{R}^D) := \{\boldsymbol{p} \in \mathbb{R}^D | p_i \geq 0 \wedge \sum_{i=1}^{D} p_i = 1\} \ . \tag{C.1}$$

The Fisher distance on $\mathcal{S}(\mathbb{R}^D)$ induces Riemannian manifold, as it derives from a *Riemannian metric*, that is a positive definite inner product that varies smoothly in the manifold. In particular, such inner product can be defined as

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle_{\boldsymbol{p}} := \sum \frac{a_i b_i}{2p_i} \ . \tag{C.2}$$

The corresponding infinitesimal length squared between $\boldsymbol{p}$ and $\boldsymbol{p}+\delta\boldsymbol{p}$ is given by

$$D^2_{\text{Fish}}(\boldsymbol{p}, \boldsymbol{p}+\delta\boldsymbol{p}) \simeq \langle \delta\boldsymbol{p}, \delta\boldsymbol{p} \rangle_{\boldsymbol{p}} := \frac{1}{2} \sum_i \frac{\delta p_i^2}{p_i} \ , \tag{C.3}$$

Notice that such metric coincides with the differential of the square root of $\boldsymbol{p}$. That is, at leading order,

$$D^2_{\text{Fish}}(\boldsymbol{p}, \boldsymbol{p}+\boldsymbol{d}) = 2|\sqrt{\boldsymbol{p}} - \sqrt{\boldsymbol{p}+\delta\boldsymbol{p}}|^2 \ . \tag{C.4}$$

This means that the Fisher metric maps the set of states from the symplex $\mathcal{S}(\mathbb{R}^D)$ to a sphere sector, via the change of variable $\mathbf{y} = \sqrt{\mathbf{p}}$ [56,130]. Accordingly, integrating Eq. (C.3) over $\mathcal{S}(\mathbb{R}^D)$ one can obtain the general expression of the distance [56, 153]

$$D_{\text{Fish}}(\boldsymbol{p}, \boldsymbol{q}) = \sqrt{2} \arccos\left(\sqrt{\boldsymbol{p}} \cdot \sqrt{\boldsymbol{q}}\right) . \tag{C.5}$$

However, notice that the same distance is well-defined in the cone of positive vectors of $\mathbb{R}^D$, that is

$$\mathfrak{C}(\mathbb{R}^D) := \{\vec{v} \in \mathbb{R}^D) | v_i \geq 0\} . \tag{C.6}$$

When integrating the distance (C.3) over two points $\mathfrak{C}(\mathbb{R}^D)$, the resulting distance is

$$D'_{\text{Fish}}(\vec{v}, \vec{w}) = \sqrt{2} \sqrt{\left(\text{Tr}[\vec{v} + \vec{w}] - 2\sqrt{\vec{v}} \cdot \sqrt{\vec{w}}\right)} . \tag{C.7}$$

It follows, for normalized vectors (i.e. physical states) that $D_{\text{Fish}}(\boldsymbol{p}, \boldsymbol{q})$ (C.5) corresponds to the angle (or arc with unit radius) between $\boldsymbol{p}$ and $\boldsymbol{q}$, while

$$D'_{\text{Fish}}(\boldsymbol{p}, \boldsymbol{q}) = \sqrt{2} \sqrt{\left(2 - 2\sqrt{\vec{\boldsymbol{p}}} \cdot \sqrt{\vec{\boldsymbol{q}}}\right)} . \tag{C.8}$$

is the euclidean distance between $\sqrt{\boldsymbol{p}}$ and $\sqrt{q}$, which clearly satisfies

$$D'_{\text{Fish}}(\boldsymbol{p}, \boldsymbol{q}) \leq D_{\text{Fish}}(\boldsymbol{p}, \boldsymbol{q}) . \tag{C.9}$$

The Fisher distance between two infinitesimally close points is sometimes called *Fisher Information*. More precisely, the Fisher Information is defined when considering parametric derivatives of the states. That is, consider the state $\boldsymbol{p}_\theta$ to be parametrized by some variable θ. Then it is possible to define

$$\lim_{\delta\theta \to 0} \frac{D^2_{\text{Fish}}(\boldsymbol{p}_\theta, \boldsymbol{p}_{\theta+\delta\theta})}{\delta\theta^2} = \sum_i \frac{(\partial_\theta p_i)^2}{p_i} . \tag{C.10}$$

This quantity is usually termed Fisher Information w.r.t. parameter θ.
This quantity has numerous operational interpretations: in metrology it is used to derive the Cramér-Rao bound [76,154], a fundamental limit on the precision with which θ can be estimated, which we clarify below in C.1.1. Moreover, it quantifies the asymptotic distinguishability between multiple copies of two states (Chernoff bound [155]); it also coincides with the infinitesimal

expansion of the relative entropy, or Kullback-Leibler divergence [156],

$$S(\boldsymbol{p}||\boldsymbol{p}+\delta\boldsymbol{p}) = \sum p_i \ln\left(\frac{p_i}{p_i+\delta p_i}\right) = -\sum_i p_i \ln\left(1+\frac{\delta p_i}{p_i}\right) = \\ -\sum_i p_i\left(\frac{\delta p_i}{p_i} - \frac{1}{2}\frac{\delta p_i^2}{p_i^2} + \dots\right) = \frac{1}{2}\sum_i \frac{\delta p_i^2}{p_i} + \mathcal{O}(\delta p^3)\ , \quad \text{(C.11)}$$

where we used that $\sum_i \delta p_i = 0$ due to the normalization constraint. More in general, it can be shown that any f-divergence locally behaves as the Fisher information [157]).
Finally, a theorem by Chentsov [56,57,159] identifies the Fisher metric (C.3) as the only Riemannian metric that is monotone under classical stochastic maps, i.e.

$$\langle \delta\boldsymbol{p}, \delta\boldsymbol{p}\rangle_{\boldsymbol{p}} \geq \langle T[\delta\boldsymbol{p}], T[\delta\boldsymbol{p}]\rangle_{T[\boldsymbol{p}]} \quad \forall T \text{ stochastic.} \quad \text{(C.12)}$$

Remarkably, the same property is not enough to single out a unique metric in the quantum case (see below Sec. C.2).

C.1.1 Single-parameter metrology and the Cramer-Rao bound

Consider the problem of estimating a parameter θ, which will be encoded in some state of the physical system at hand in the laboratory. In the classical case, the state is therefore given by some probability distribution

$$p_\theta(x)\ , \quad \text{(C.13)}$$

where we now take x to be either discrete or continuous, i.e. $\boldsymbol{p}_\theta$ can be either finite or infinite dimensional. An *estimator* is a function $\hat{\tau}(x)$ trying to reproduce the value of θ, based on the measurement output x (notice that in the classical case there is no ambiguity regarding the measurement observable. Typical estimators are not perfect: they might have, for example some systematic error $b(\theta)$, called *bias*: that is, taking the expectation value

$$E_\theta(\hat{\tau}) = \int \mathrm{d}x\ \hat{\tau}(x) p_\theta(x) = \theta + b(\theta)\ . \quad \text{(C.14)}$$

Consider now the derivative of the above equation in θ

$$1 + \partial_\theta b(\theta) = \int \mathrm{d}x\ \hat{\tau}(x)\partial_\theta p_\theta(x) = \int \mathrm{d}x\ (\hat{\tau}(x) - k(\theta))\frac{\partial_\theta p_\theta(x)}{p_\theta} p_\theta(x)\ , \quad \text{(C.15)}$$

where in the second equality we added for free any constant $k(\theta)$ which does not depend on x. Taking now $\int \mathrm{d}x\ A(x)B(x)p_\theta(x) := \langle A, B\rangle$ as a scalar product, and applying the Cauchy-Schwarz inequality, we obtain

$$(1+\partial_\theta b(\theta))^2 \leq \left(\int \mathrm{d}x\ (\hat{\tau}(x)-k(\theta))^2 p_\theta(x)\right)\left(\int \mathrm{d}x\ \frac{(\partial_\theta p_\theta(x))^2}{p_\theta(x)}\right) . \tag{C.16}$$

The above equation gives us a lower bound on the mean squared distance of the estimator from any constant, in particular the mean square error (choosing $k(\theta)=\theta$)

$$E_\theta[(\hat{\tau}-\theta)^2] \geq \frac{(1+\partial_\theta b(\theta))^2}{E_\theta[(\partial_\theta \ln p_\theta)^2]} . \tag{C.17}$$

This is the so called Cramer-Rao bound [76, 154]. Good estimators are unbiased, e.g. $\partial_\theta b(\theta)=0$. We see therefore that the average mistake is inversely proportional to the Fisher information

$$E_\theta[(\hat{\tau}-\theta)^2] \geq \left(E_\theta[(\partial_\theta \ln p_\theta)^2]\right)^{-1} \equiv \left(\int \mathrm{d}x \frac{(\partial_\theta p_\theta(x))^2}{p_\theta(x)}\right)^{-1} . \tag{C.18}$$

Parenthesis on the Bayesian case. A similar bound can be obtained in case some prior knowledge about the possible value of θ is given. That is, consider a prior distribution on the possible values of θ given by $K(\theta)$. We have, by definition, for fixed θ, $E_\theta[\hat{\tau}(x)] = \theta + b(\theta)$, which can be integrated over the distribution K to obtain

$$\int \mathrm{d}\theta \mathrm{d}x\ K(\theta)p_\theta(x)\,(\hat{\tau}(x)-\theta-b(\theta)) = 0 . \tag{C.19}$$

Taking the derivative in θ of the above equation (notice that for the equality to 0 it is sufficient to integrate in x) one obtains

$$\int \mathrm{d}\theta \mathrm{d}x\ \left(\partial_\theta K(\theta) + \partial_\theta p_\theta)\right)(\hat{\tau}-\theta-b(\theta)) = \int \mathrm{d}\theta \mathrm{d}x\ K(\theta)\,(1+\partial_\theta b(\theta)) . \tag{C.20}$$

Similarly as above, one can consider the integral weighted with $K(\theta)p_\theta(x)$ as a scalar product, and apply the Cauchy-Schwarz inequality to obtain a Bayesian

Cramer-Rao bound, which we express here for unbiased estimators

$$\int \mathrm{d}\theta \mathrm{d}x \; K(\theta) p_\theta(x) \left(\hat{\tau} - \theta - b(\theta)\right)^2 \geq \frac{\left(1 + \int \mathrm{d}\theta \; \partial_\theta b(\theta)\right)^2}{\int \mathrm{d}\theta \; \frac{(\partial_\theta K(\theta))^2}{K(\theta} + \int \mathrm{d}\theta \mathrm{d}x \; K(\theta) \frac{(\partial_\theta p_\theta(x))^2}{p_\theta(x)}} \tag{C.21}$$

This bound is on the variance of $\hat{\tau}$, but a similar bound can be obtained by neglecting the integral of $\partial_\theta(K(\theta)b(\theta))$, (which is justified under the very mild assumption of $K(\theta)b(\theta)$ going to zero for large values of θ), to obtain for the mean square error

$$\int \mathrm{d}\theta \mathrm{d}x \; K(\theta) p_\theta(x) \left(\hat{\tau} - \theta\right)^2 \geq \frac{1}{\int \mathrm{d}\theta \; \frac{(\partial_\theta K(\theta))^2}{K(\theta} + \int \mathrm{d}\theta \mathrm{d}x \; K(\theta) \frac{(\partial_\theta p_\theta(x))^2}{p_\theta(x)}} \; . \tag{C.22}$$

As we see, in the Bayesian case, the expected precision of the measurement is given by the average Fisher Information of the state $p_\theta(x)$, plus the Fisher Information of the prior $K(\theta)$ itself.

C.2 Quantum Fisher metric

The extension of the Fisher information metric to quantum systems is done by generalising Chentsov's theorem to completely positive, trace preserving maps (CPTP). That is, a metric $K^f_\rho(A, B)$ on quantum states is called monotone if it decreases under all CPTP maps $\mathcal{E}$, that is $K_\rho(A, A) \geq K_{\mathcal{E}[\rho]}(\mathcal{E}[A], \mathcal{E}[A])$. Then, it was shown by Petz in [160] that all such metrics are induced by scalar products of the form:

$$K^f_\rho(A, B) := \frac{1}{2}\mathrm{Tr}[A^\dagger \, \mathbb{J}^{-1}_{f,\rho}[B]], \tag{C.23}$$

where $\mathbb{J}_{f,\rho}$ is a self-adjoint superoperator given by:

$$\mathbb{J}_{f,\rho} := \mathbb{R}_\rho \, f(\mathbb{L}_\rho \mathbb{R}_\rho^{-1}), \tag{C.24}$$

and $\mathbb{L}_\rho/\mathbb{R}_\rho$ are the left/right multiplication operators acting as $\mathbb{L}_\rho[\pi] = \rho\pi$ (respectively $\mathbb{R}_\rho[\pi] = \pi\rho$) and $f : \mathbb{R}^+ \to \mathbb{R}^+$ is a standard operator monotone function, which satisfies

$$f(1) = 1 \; , \quad f(x) = x f(x^{-1}) \; . \tag{C.25}$$

Operator monotone functions satisfying the above requirements are all contained between two extrema

$$\frac{2x}{x+1} \leq f(x) \leq \frac{1+x}{2} \ . \tag{C.26}$$

Despite this complicated form, and the remaining arbitrariness of f, the interpretation of $K^f_\rho(A,B)$ as the natural extension of the Fisher information to the quantum regime is corroborated by the fact the $\mathbb{J}_{f,\rho}[A]$ acts as a symmetrized non-commutative version of the multiplication by ρ (consequently $\mathbb{J}^{-1}_{f,\rho}$ can be considered a non-commutative multiplication by ρ^{-1}, similarly to the classical case). See more in detail Sec. C.2.2.
For this reason, we define the family of quantum Fisher distances as

$$D^2_{\text{Fish},\,f}(\rho,\rho+\delta\rho) \simeq \ K^f_\rho(\delta\rho,\delta\rho) := \frac{1}{2}\text{Tr}[\delta\rho^\dagger\ \mathbb{J}^{-1}_{f,\rho}[\delta\rho]] \ . \tag{C.27}$$

The uniqueness of the classical Fisher metric is hence substituted with a whole family of different monotone metrics. In the following we provide few relevant examples. It should be kept in mind that all these different examples of Fisher metric collapse onto the same metric in the classical case of commuting operators (cf. C.2.2) .

C.2.1 Examples

KMB Fisher metric and the thermodynamic length
As mentioned in Appendix A.1, the expression of work dissipation/irreversible entropy production (A.29) corresponds to the integral of a particular Quantum Information metric, i.e. the one given by the Kubo-Mori-Bogoliubov inner product, which corresponds to the choice of f

$$f(x) = \int_0^1 \mathrm{d}y\ x^y \ , \tag{C.28}$$

which gives

$$\mathbb{J}_{f(x)=\int_0^1 \mathrm{d}y\ x^y,\rho}[A] = \int_0^1 \mathrm{d}y\ \rho^y A \rho^{1-y} \ . \tag{C.29}$$

This can be inverted as

$$\mathbb{J}^{-1}_{f(x)=\int_0^1 \mathrm{d}y\ x^y,\rho}[A] = \int_0^\infty \mathrm{d}y (\rho+y)^{-1} A (\rho+y)^{-1} \ . \tag{C.30}$$

The corresponding infinitesimal length squared $\frac{1}{2}\mathrm{Tr}[\delta\rho\mathbb{J}_\rho^{-1}[\delta\rho]]$ is the one appearing in (A.29).

Bures Fisher metric and the Quantum Cramer-Rao bound.
By choosing f being equal to the maximum (C.26)

$$f(x) = \frac{x+1}{2} \,, \tag{C.31}$$

$\mathbb{J}_{f,\rho}$ (C.24) becomes proportional to the anticommutator

$$\mathbb{J}_{f=\frac{x+1}{2},\rho}[A] = \frac{1}{2}\{A,\rho\} \,, \tag{C.32}$$

which has inverse equal to

$$\mathbb{J}^{-1}_{f=\frac{x+1}{2},\rho}[A] = 2\int_0^\infty \mathrm{d}y\; e^{-y\rho}Ae^{-y\rho} \,. \tag{C.33}$$

The resulting metric (C.23) is the so called *Bures metric*, which corresponds to the smallest of all the Quantum Information metrics. Moreover, the Bures metric, when considered as a parametric derivative, is the metrological generalization of the Fisher Information (C.18) entering the Cramer Rao bound. That is, in the context of Quantum parameter estimation, the Quantum Cramer Rao bound reads [128], for an unbiased estimators $\hat{\tau}$ of θ,

$$E_\theta[(\hat{\tau}-\theta)^2] \geq \left(\mathrm{Tr}[\partial_\theta\rho_\theta \mathbb{J}_{f=\frac{x+1}{2}^{-1},\rho_\theta}[\partial_\theta\rho_\theta]]\right)^{-1} \,. \tag{C.34}$$

Geometric Fisher metric, complete positivity and the Petz recovery map
Another important example is given by choosing

$$f(x) = \sqrt{x} \,, \tag{C.35}$$

from which it follows immediately

$$\mathbb{J}_{f=\sqrt{x},\rho}[A] = \sqrt{\rho}A\sqrt{\rho} \,, \quad \mathbb{J}^{-1}_{f=\sqrt{x},\rho}[A] = \sqrt{\rho^{-1}}A\sqrt{\rho^{-1}} \,. \tag{C.36}$$

This is corresponds to the so called *geometric* $\mathbb{J}_{f,\rho}$ operator, and is the only Quantum Fisher metric for which it holds that both $\mathbb{J}_{f,\rho}$ and $\mathbb{J}^{-1}_{f,\rho}$ are Completely-Positive operators. Related to this, this is the Fisher metric that is used to generalized the Bayesian retrodiction map (4.26), to the Quantum Petz recovery map [165, 231].

C.2.2 Quantum reduction Classical

It is interesting to point out, though, that when $\mathbb{J}_{f,\rho}$ acts on diagonal states it behaves (independently of f) as the multiplication by ρ. That is, if $[A,\rho]=0$

$$\mathbb{J}_{f,\rho}[A]=\rho A=A\rho\,,\quad \mathbb{J}^{-1}_{f,\rho}[A]=\rho^{-1}A=A\rho^{-1}\,. \tag{C.37}$$

And therefore, if $A,\rho]=[B,\rho]=0$, with a small abuse of notation we have

$$K^f_\rho(A,B):=\frac{1}{2}\mathrm{Tr}\left[\frac{A^\dagger B}{\rho}\right]=\langle A,B\rangle_\rho\,, \tag{C.38}$$

where A,B and ρ can be seen as the vectors corresponding to the diagonal elements in the common diagonal basis. That is, all the quantum Fisher metrics collapse to the classical one for diagonal states. This is the main feature allows us, in Ref. [7] to lift to the quantum scenario the results presented in Chapter 4 for the classical case, without further complications.
Specifically, we present in the following few particularly relevant Lemmas.

Lemma 2. *Given a state $\rho=\sum_i \rho_i\,|i\rangle\langle i|$, any perturbation $\delta\rho$ can be decomposed in diagonal and coherent part as:*

$$\delta\rho=\delta\rho_\Delta+\delta\rho_C\quad \textit{with}\quad [\rho,\delta\rho_\Delta]=0\quad \textit{and}\quad \langle i|\,\delta\rho_C\,|i\rangle=0\,, \tag{C.39}$$

that is, $\delta\rho_\Delta$ is the diagonal part of the matrix $\delta\rho$ and $\delta\rho_C$ the off-diagonal. Then for all f

$$\begin{aligned} Tr[\delta\rho\,\mathbb{J}^{-1}_{f,\rho}[\delta\rho]] &= Tr[\delta\rho_\Delta\,\mathbb{J}^{-1}_{f,\rho}[\delta\rho_\Delta]]+Tr[\delta\rho_C\,\mathbb{J}^{-1}_{f,\rho}[\delta\rho_C]]\;= && \text{(C.40)}\\ &= 2\,\langle\boldsymbol{\delta},\boldsymbol{\delta}\rangle_{\boldsymbol{\rho}}+Tr[\delta\rho_C\,\mathbb{J}^{-1}_{f,\rho}[\delta\rho_C]]\,, && \text{(C.41)}\end{aligned}$$

where the components of the vectors $\boldsymbol{\delta}$, $\boldsymbol{\rho}$ are specified as

$$\boldsymbol{\rho}_i=\langle i|\,\rho\,|i\rangle\,,\;\boldsymbol{\delta}_i=\langle i|\,\delta\rho\,|i\rangle=\langle i|\,\delta\rho_\Delta\,|i\rangle\,. \tag{C.42}$$

This result directly follows from the fact that

$$\mathrm{Tr}\delta\rho_\Delta\,\mathbb{J}^{-1}_{f,\rho}[\delta\rho_C]=\mathrm{Tr}\delta\rho_C\,\mathbb{J}^{-1}_{f,\rho}[\delta\rho_\Delta]=0\,, \tag{C.43}$$

since $\mathbb{J}^{-1}_{f,\rho}[\delta\rho_\Delta]$ is itself diagonal in the basis $|i\rangle$ of eigenvectors of ρ. Then, we can use the Lemma above to prove the following corollary:

Corollary 2. *Consider a perturbation of the form $\delta\rho = \delta\rho_\Delta + \mathrm{d}t\,\delta\rho_C$, where $\mathrm{d}t$ is an infinitesimal quantity. Then, from Lemma 2 it follows that*

$$Tr[\delta\rho\, \mathbb{J}^{-1}_{f,\rho}[\delta\rho]] = Tr[\delta\rho_\Delta\, \mathbb{J}^{-1}_{f,\rho}[\delta\rho_\Delta]] + \mathcal{O}\left(\mathrm{d}t^{\,2}\right) = 2\,\langle \boldsymbol{\delta}, \boldsymbol{\delta}\rangle_{\boldsymbol{\rho}} + \mathcal{O}\left(\mathrm{d}t^{\,2}\right). \tag{C.44}$$

In particular, the time derivative of the Fisher Information between ρ and $\rho+\delta\rho$ for $[\rho,\delta\rho]=0$ coincides with the derivative of the classical Fisher Information. That is:

$$\frac{1}{2}\,Tr[\delta\rho\, \mathbb{J}^{-1}_{f,\rho}[\delta\rho]] = \langle \boldsymbol{\delta}, \boldsymbol{\delta}\rangle_{\boldsymbol{\rho}} \quad \textit{and} \quad \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\,Tr[\delta\rho\, \mathbb{J}^{-1}_{f,\rho}[\delta\rho]] = \frac{\mathrm{d}}{\mathrm{d}t}\,\langle \boldsymbol{\delta}, \boldsymbol{\delta}\rangle_{\boldsymbol{\rho}}\ . \tag{C.45}$$

In fact, consider the scenario in which initially the perturbation is of the form $\delta\rho \equiv \delta\rho_\Delta$. In order to compute the derivative, one considers the evolution of the state $\rho+\delta\rho$ for a time $\mathrm{d}t$, which we denote by $\tilde{\rho}+\delta\tilde{\rho}$. Then, the perturbation has the form $\delta\tilde{\rho} = \delta\tilde{\rho}_\Delta + \mathrm{d}t\,\delta\tilde{\rho}_C$, so that we are in the situation of Eq. (C.44) (notice that it doesn't matter whether we take $\delta\tilde{\rho}_\Delta$ to be diagonal with respect to ρ or to $\tilde{\rho}$, as this difference only contributes to order $\mathcal{O}(\mathrm{d}t)$). Since quadratic terms in $\mathrm{d}t$ do not contribute to the derivative, these considerations prove Eq. (C.45).

Finally, in the next Lemma we show that there are special points for which the trace distance and the Fisher information locally coincide. This result allows us to lift the many constructions present in the literature for the trace distance to the study of the Fisher information metric.

Lemma 3. *Choose an arbitrary perturbation $\delta\rho$. Then, consider the state $\rho_{\delta\rho} = \frac{|\delta\rho|}{Tr|\delta\rho|}$. It holds that*

$$D^2_{\mathrm{Fish}}(\rho_{\delta\rho}, \rho_{\delta\rho}+\delta\rho) = \frac{1}{2}\,D^2_{\mathrm{Tr}}(\rho_{\delta\rho}, \rho_{\delta\rho}+\delta\rho) = \frac{1}{2}\,Tr|\delta\rho|^2\,. \tag{C.46}$$

Moreover, since $[\rho_{\delta\rho}, \delta\rho]=0$, one can use Corollary 2 to show that:

$$\frac{\mathrm{d}}{\mathrm{d}t}D^2_{\mathrm{Fish}}(\rho_{\delta\rho}, \rho_{\delta\rho}+\delta\rho) = \frac{1}{2}\,\frac{\mathrm{d}}{\mathrm{d}t}D^2_{\mathrm{Tr}}(\rho_{\delta\rho}, \rho_{\delta\rho}+\delta\rho) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\,Tr|\delta\rho|^2 \tag{C.47}$$

In fact, since $[\rho_{\delta\rho}, \delta\rho] = 0$, the quantum Fisher information and its derivative can be studied just by looking at the quantity $\langle \boldsymbol{\delta}, \boldsymbol{\delta} \rangle_{\boldsymbol{\rho}_{\boldsymbol{\delta}\boldsymbol{\rho}}}$. But this is given by:

$$2 \, \langle \boldsymbol{\delta}, \boldsymbol{\delta} \rangle_{\boldsymbol{\rho}_{\boldsymbol{\delta}\boldsymbol{\rho}}} = \sum_i \frac{\delta_i^2}{\left(\frac{|\delta_i|}{\mathrm{Tr}|\delta\rho|} \right)} = \sum_i \frac{\delta_i^2}{|\delta_i|} \sum_j |\delta_j| =$$

$$= \left(\sum_i |\delta_i| \right)^2 = D^2_{\mathrm{Tr}}(\rho_{\delta\rho}, \rho_{\delta\rho} + \delta\rho) \, . \quad \text{(C.48)}$$

The fact that not only one can locally identify the Fisher distance and the trace distance, but also their first derivatives is of key importance in many of our derivations in [7].

Bibliography

[1] P. Abiuso and M. Perarnau-Llobet, *Optimal cycles for low-dissipation heat engines*, Phys. Rev. Lett. **124**, 110606 (2020), doi:10.1103/PhysRevLett.124.110606.

[2] P. Abiuso, H. J. D. Miller, M. Perarnau-Llobet and M. Scandi, *Geometric optimisation of quantum thermodynamic processes*, Entropy **22**(10) (2020), doi:10.3390/e22101076.

[3] P. Abiuso, S. Bäuml, D. Cavalcanti and A. Acín, *Measurement-device-independent entanglement detection for continuous-variable systems*, Phys. Rev. Lett. **126**, 190502 (2021), doi:10.1103/PhysRevLett.126.190502.

[4] V. Cavina, P. A. Erdman, P. Abiuso, L. Tolomeo and V. Giovannetti, *Maximum-power heat engines and refrigerators in the fast-driving regime*, Phys. Rev. A **104**, 032226 (2021), doi:10.1103/PhysRevA.104.032226.

[5] P. Abiuso, T. Kriváchy, E.-C. Boghiu, M.-O. Renou, A. Pozas-Kerstjens and A. Acín, *Single-photon nonlocality in quantum networks*, Phys. Rev. Research **4**, L012041 (2022), doi:10.1103/PhysRevResearch.4.L012041.

[6] P. Terrén Alonso, P. Abiuso, M. Perarnau-Llobet and L. Arrachea, *Geometric optimization of nonequilibrium adiabatic thermal machines and implementation in a qubit system*, PRX Quantum **3**, 010326 (2022), doi:10.1103/PRXQuantum.3.010326.

[7] P. Abiuso, M. Scandi, D. De Santis and J. Surace, *Characterizing (non-)markovianity through fisher information* (2022), doi:10.48550/ARXIV.2204.04072.

[8] P. Abiuso, P. A. Erdman, M. Ronen, G. Haack, F. Noé and M. Perarnau-Llobet, *Discovery of optimal thermometers with spin networks aided by machine-learning* (2022), doi:10.48550/ARXIV.2211.01934.

[9] P. Abiuso and V. Giovannetti, *Non-markov enhancement of maximum power for quantum thermal machines*, Phys. Rev. A **99**, 052106 (2019), doi:10.1103/PhysRevA.99.052106.

[10] P. Abiuso, V. Holubec, J. Anders, Z. Ye, F. Cerisola and M. Perarnau-Llobet, *Thermodynamics and optimal protocols of multidimensional quadratic brownian systems*, Journal of Physics Communications **6**(6), 063001 (2022), doi:10.1088/2399-6528/ac72f8.

[11] Z. Ye, F. Cerisola, P. Abiuso, J. Anders, M. Perarnau-Llobet and V. Holubec, *Optimal finite-time heat engines under constrained control* (2022), doi:10.48550/ARXIV.2202.12953.

[12] P. A. Erdman, A. Rolandi, P. Abiuso, M. Perarnau-Llobet and F. Noé, *Pareto-optimal cycles for power, efficiency and fluctuations of quantum heat engines using reinforcement learning* (2022), doi:10.48550/ARXIV.2207.13104.

[13] S. Vinjanampathy and J. Anders, *Quantum thermodynamics*, Contemporary Physics **57**(4), 545 (2016), doi:10.1080/00107514.2016.1201896.

[14] S. Deffner and S. Campbell, *Quantum Thermodynamics*, 2053-2571. Morgan & Claypool Publishers, ISBN 978-1-64327-658-8, doi:10.1088/2053-2571/ab21c6 (2019).

[15] K. E. Dorfman, D. V. Voronine, S. Mukamel and M. O. Scully, *Photosynthetic reaction center as a quantum heat engine*, Proceedings of the National Academy of Sciences **110**(8), 2746 (2013), doi:10.1073/pnas.1212666110.

[16] J. V. Koski, V. F. Maisi, J. P. Pekola and D. V. Averin, *Experimental realization of a szilard engine with a single electron*, Proceedings of the National Academy of Sciences **111**(38), 13786 (2014), doi:10.1073/pnas.1406966111.

[17] J. P. Pekola, *Towards quantum thermodynamics in electronic circuits*, Nature Physics **11**(2), 118 (2015), doi:10.1038/nphys3169.

[18] J. Roßnagel, S. T. Dawkins, K. N. Tolazzi, O. Abah, E. Lutz, F. Schmidt-Kaler and K. Singer, *A single-atom heat engine*, Science **352**(6283), 325 (2016), doi:10.1126/science.aad6320.

[19] A. Ronzani, B. Karimi, J. Senior, Y.-C. Chang, J. T. Peltonen, C. Chen and J. P. Pekola, *Tunable photonic heat transport in a quantum heat valve*, Nature Physics **14**(10), 991 (2018), doi:10.1038/s41567-018-0199-4.

[20] N. M. Myers, O. Abah and S. Deffner, *Quantum thermodynamic devices: From theoretical proposals to experimental reality*, AVS Quantum Science **4**(2), 027101 (2022), doi:10.1116/5.0083192.

[21] V. Cavina, A. Mari, A. Carlini and V. Giovannetti, *Optimal thermodynamic control in open quantum systems*, Phys. Rev. A **98**, 012139 (2018), doi:10.1103/PhysRevA.98.012139.

[22] L. S. Pontryagin, *Mathematical theory of optimal processes*, CRC press, ISBN 9782881240775 (1987).

[23] P. A. Erdman and F. Noé, *Identifying optimal cycles in quantum thermal machines with reinforcement-learning*, npj Quantum Information **8**(1), 1 (2022), doi:10.1038/s41534-021-00512-0.

[24] P. A. Erdman, V. Cavina, R. Fazio, F. Taddei and V. Giovannetti, *Maximum power and corresponding efficiency for two-level heat engines and refrigerators: optimality of fast cycles*, New Journal of Physics **21**(10), 103049 (2019), doi:10.1088/1367-2630/ab4dca.

[25] T. Schmiedl and U. Seifert, *Optimal finite-time processes in stochastic thermodynamics*, Phys. Rev. Lett. **98**, 108301 (2007), doi:10.1103/PhysRevLett.98.108301.

[26] T. Schmiedl and U. Seifert, *Efficiency at maximum power: An analytically solvable model for stochastic heat engines*, EPL (Europhysics Letters) **81**(2), 20003 (2007), doi:10.1209/0295-5075/81/20003.

[27] A. Gomez-Marin, T. Schmiedl and U. Seifert, *Optimal protocols for minimal work processes in underdamped stochastic thermodynamics*, The Journal of Chemical Physics **129**(2), 024114 (2008), doi:10.1063/1.2948948.

[28] A. Dechant, N. Kiesel and E. Lutz, *Underdamped stochastic heat engine at maximum efficiency*, EPL (Europhysics Letters) **119**(5), 50003 (2017), doi:10.1209/0295-5075/119/50003.

[29] V. Cavina, A. Mari and V. Giovannetti, *Slow dynamics and thermodynamics of open quantum systems*, Phys. Rev. Lett. **119**, 050601 (2017), doi:10.1103/PhysRevLett.119.050601.

[30] F. Weinhold, *Metric geometry of equilibrium thermodynamics*, The Journal of Chemical Physics **63**(6), 2479 (1975), doi:10.1063/1.431689.

[31] P. Salamon and R. S. Berry, *Thermodynamic length and dissipated availability*, Phys. Rev. Lett. **51**, 1127 (1983), doi:10.1103/PhysRevLett.51.1127.

[32] G. Ruppeiner, *Riemannian geometry in thermodynamic fluctuation theory*, Rev. Mod. Phys. **67**, 605 (1995), doi:10.1103/RevModPhys.67.605.

[33] L. Diósi, K. Kulacsy, B. Lukács and A. Rácz, *Thermodynamic length, time, speed, and optimum path to minimize entropy production*, The Journal of Chemical Physics **105**(24), 11220 (1996), doi:10.1063/1.472897.

[34] P. R. Zulkowski, D. A. Sivak, G. E. Crooks and M. R. DeWeese, *Geometry of thermodynamic control*, Phys. Rev. E **86**, 041148 (2012), doi:10.1103/PhysRevE.86.041148.

[35] M. Scandi and M. Perarnau-Llobet, *Thermodynamic length in open quantum systems*, Quantum **3**, 197 (2019), doi:10.22331/q-2019-10-24-197.

[36] B. Bhandari, P. T. Alonso, F. Taddei, F. von Oppen, R. Fazio and L. Arrachea, *Geometric properties of adiabatic quantum thermal machines*, Phys. Rev. B **102**, 155407 (2020), doi:10.1103/PhysRevB.102.155407.

[37] J. P. Pekola, B. Karimi, G. Thomas and D. V. Averin, *Supremacy of incoherent sudden cycles*, Phys. Rev. B **100**, 085405 (2019), doi:10.1103/PhysRevB.100.085405.

[38] A. Das and V. Mukherjee, *Quantum-enhanced finite-time otto cycle*, Phys. Rev. Research **2**, 033083 (2020), doi:10.1103/PhysRevResearch.2.033083.

[39] M. Campisi and R. Fazio, *The power of a critical heat engine*, Nature Communications **7**(1), 11895 (2016), doi:10.1038/ncomms11895.

[40] A. De Pasquale and T. M. Stace, *Quantum Thermometry*, pp. 503–527, Springer International Publishing, Cham, ISBN 978-3-319-99046-0, doi:10.1007/978-3-319-99046-0_21 (2018).

[41] M. Mehboudi, A. Sanpera and L. A. Correa, *Thermometry in the quantum regime: recent theoretical progress*, Journal of Physics A: Mathematical and Theoretical **52**(30), 303001 (2019), doi:10.1088/1751-8121/ab2828.

[42] L. A. Correa, M. Mehboudi, G. Adesso and A. Sanpera, *Individual quantum probes for optimal thermometry*, Phys. Rev. Lett. **114**, 220405 (2015), doi:10.1103/PhysRevLett.114.220405.

[43] L. Buffoni and M. Campisi, *Thermodynamics of a quantum annealer*, Quantum Science and Technology **5**(3), 035013 (2020), doi:10.1088/2058-9565/ab9755.

[44] Z. G. Izquierdo, I. Hen and T. Albash, *Testing a quantum annealer as a quantum thermal sampler*, ACM Transactions on Quantum Computing **2**(2) (2021), doi:10.1145/3464456.

[45] T. Albash and J. Marshall, *Comparing relaxation mechanisms in quantum and classical transverse-field annealing*, Phys. Rev. Applied **15**, 014029 (2021), doi:10.1103/PhysRevApplied.15.014029.

[46] R. Zwanzig, A. Szabo and B. Bagchi, *Levinthal's paradox.*, Proceedings of the National Academy of Sciences **89**(1), 20 (1992), doi:10.1073/pnas.89.1.20.

[47] R. Zwanzig, *Simple model of protein folding kinetics.*, Proceedings of the National Academy of Sciences **92**(21), 9801 (1995), doi:10.1073/pnas.92.21.9801.

[48] E. Farhi and S. Gutmann, *Analog analogue of a digital quantum computation*, Phys. Rev. A **57**, 2403 (1998), doi:10.1103/PhysRevA.57.2403.

[49] J. Roland and N. J. Cerf, *Quantum search by local adiabatic evolution*, Phys. Rev. A **65**, 042308 (2002), doi:10.1103/PhysRevA.65.042308.

[50] N. Chancellor, S. Zohren, P. A. Warburton, S. C. Benjamin and S. Roberts, *A Direct Mapping of Max k-SAT and High Order Parity Checks to a Chimera Graph*, Scientific Reports **6**(1), 37107 (2016), doi:10.1038/srep37107.

[51] H.-P. Breuer, F. Petruccione *et al.*, *The theory of open quantum systems*, Oxford University Press on Demand, doi:10.1093/acprof:oso/9780199213900.001.0001 (2002).

[52] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, Cambridge University Press, doi:10.1017/CBO9780511976667 (2010).

[53] Á. Rivas, S. F. Huelga and M. B. Plenio, *Quantum non-markovianity: characterization, quantification and detection*, Reports on Progress in Physics **77**(9), 094001 (2014), doi:10.1088/0034-4885/77/9/094001.

[54] H.-P. Breuer, E.-M. Laine, J. Piilo and B. Vacchini, *Colloquium: Non-markovian dynamics in open quantum systems*, Rev. Mod. Phys. **88**, 021002 (2016), doi:10.1103/RevModPhys.88.021002.

[55] B. Bylicka, M. Johansson and A. Acín, *Constructive method for detecting the information backflow of non-markovian dynamics*, Phys. Rev. Lett. **118**, 120501 (2017), doi:10.1103/PhysRevLett.118.120501.

[56] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, Cambridge University Press, doi:10.1017/CBO9780511535048 (2006).

[57] N. N. Cencov, *Statistical decision rules and optimal inference*, 53. American Mathematical Soc., doi:https://doi.org/10.1090/mmono/053 (2000).

[58] J. S. Bell, *On the einstein podolsky rosen paradox*, Physics Physique Fizika **1**, 195 (1964), doi:10.1103/PhysicsPhysiqueFizika.1.195.

[59] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani and S. Wehner, *Bell nonlocality*, Rev. Mod. Phys. **86**, 419 (2014), doi:10.1103/RevModPhys.86.419.

[60] A. Tavakoli, A. Pozas-Kerstjens, M.-X. Luo and M.-O. Renou, *Bell nonlocality in networks*, Reports on Progress in Physics **85**(5), 056001 (2022), doi:10.1088/1361-6633/ac41bb.

[61] T. Fritz, *Beyond bell's theorem: correlation scenarios*, New Journal of Physics **14**(10), 103001 (2012), doi:10.1088/1367-2630/14/10/103001.

[62] M.-O. Renou, E. Bäumer, S. Boreiri, N. Brunner, N. Gisin and S. Beigi, *Genuine quantum nonlocality in the triangle network*, Phys. Rev. Lett. **123**, 140401 (2019), doi:10.1103/PhysRevLett.123.140401.

[63] S. L. Braunstein and P. van Loock, *Quantum information with continuous variables*, Rev. Mod. Phys. **77**, 513 (2005), doi:10.1103/RevModPhys.77.513.

[64] C. Branciard, D. Rosset, Y.-C. Liang and N. Gisin, *Measurement-device-independent entanglement witnesses for all entangled quantum states*, Phys. Rev. Lett. **110**, 060405 (2013), doi:10.1103/PhysRevLett.110.060405.

[65] F. Buscemi, *All entangled quantum states are nonlocal*, Phys. Rev. Lett. **108**, 200401 (2012), doi:10.1103/PhysRevLett.108.200401.

[66] M. Scandi, P. Abiuso, J. Surace and D. De Santis, Quantum Fisher Information: divisible dynamics, reverse processes and witnesses for Non-Markovianity *(In preparation)* .

[67] N. Linden, S. Popescu and P. Skrzypczyk, *How small can thermal machines be? the smallest possible refrigerator*, Phys. Rev. Lett. **105**, 130401 (2010), doi:10.1103/PhysRevLett.105.130401.

[68] C. Gogolin and J. Eisert, *Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems*, Reports on Progress in Physics **79**(5), 056001 (2016), doi:10.1088/0034-4885/79/5/056001.

[69] J. P. S. Peterson, T. B. Batalhão, M. Herrera, A. M. Souza, R. S. Sarthour, I. S. Oliveira and R. M. Serra, *Experimental characterization of a spin quantum heat engine*, Phys. Rev. Lett. **123**, 240601 (2019), doi:10.1103/PhysRevLett.123.240601.

[70] J. Klatzow, J. N. Becker, P. M. Ledingham, C. Weinzetl, K. T. Kaczmarek, D. J. Saunders, J. Nunn, I. A. Walmsley, R. Uzdin and E. Poem, *Experimental demonstration of quantum effects in the operation of microscopic heat engines*, Phys. Rev. Lett. **122**, 110601 (2019), doi:10.1103/PhysRevLett.122.110601.

[71] D. von Lindenfels, O. Gräb, C. T. Schmiegelow, V. Kaushal, J. Schulz, M. T. Mitchison, J. Goold, F. Schmidt-Kaler and U. G. Poschinger, *Spin heat engine coupled to a harmonic-oscillator flywheel*, Phys. Rev. Lett. **123**, 080602 (2019), doi:10.1103/PhysRevLett.123.080602.

[72] J. M. Parrondo, J. M. Horowitz and T. Sagawa, *Thermodynamics of information*, Nature physics **11**(2), 131 (2015), doi:10.1038/nphys3230.

[73] J. Goold, M. Huber, A. Riera, L. del Rio and P. Skrzypczyk, *The role of quantum information in thermodynamics—a topical review*, Journal of Physics A: Mathematical and Theoretical **49**(14), 143001 (2016), doi:10.1088/1751-8113/49/14/143001.

[74] M. Lostaglio, *An introductory review of the resource theory approach to thermodynamics*, Reports on Progress in Physics **82**(11), 114001 (2019), doi:10.1088/1361-6633/ab46e5.

[75] M. Esposito, R. Kawai, K. Lindenberg and C. Van den Broeck, *Efficiency at maximum power of low-dissipation carnot engines*, Phys. Rev. Lett. **105**, 150603 (2010), doi:10.1103/PhysRevLett.105.150603.

[76] H. Cramér, *Mathematical methods of statistics*, Princeton University Press, ISBN 0691005478, doi:https://doi.org/10.1515/9781400883868 (1946).

[77] G. Benenti, G. Casati, K. Saito and R. S. Whitney, *Fundamental aspects of steady-state conversion of heat to work at the nanoscale*, Physics Reports **694**, 1 (2017), doi:https://doi.org/10.1016/j.physrep.2017.05.008.

[78] F. Tonner and G. Mahler, *Autonomous quantum thermodynamic machines*, Phys. Rev. E **72**, 066118 (2005), doi:10.1103/PhysRevE.72.066118.

[79] A. Roulet, S. Nimmrichter, J. M. Arrazola, S. Seah and V. Scarani, *Autonomous rotor heat engine*, Phys. Rev. E **95**, 062131 (2017), doi:10.1103/PhysRevE.95.062131.

[80] R. Alicki, *The quantum open system as a model of the heat engine*, Journal of Physics A: Mathematical and General **12**(5), L103 (1979), doi:10.1088/0305-4470/12/5/007.

[81] T. D. Kieu, *The second law, maxwell's demon, and work derivable from quantum heat engines*, Phys. Rev. Lett. **93**, 140403 (2004), doi:10.1103/PhysRevLett.93.140403.

[82] A. C. Barato and U. Seifert, *Thermodynamic uncertainty relation for biomolecular processes*, Phys. Rev. Lett. **114**, 158101 (2015), doi:10.1103/PhysRevLett.114.158101.

[83] T. R. Gingrich, J. M. Horowitz, N. Perunov and J. L. England, *Dissipation bounds all steady-state current fluctuations*, Phys. Rev. Lett. **116**, 120601 (2016), doi:10.1103/PhysRevLett.116.120601.

[84] P. Pietzonka and U. Seifert, *Universal trade-off between power, efficiency, and constancy in steady-state heat engines*, Phys. Rev. Lett. **120**, 190602 (2018), doi:10.1103/PhysRevLett.120.190602.

[85] G. Guarnieri, G. T. Landi, S. R. Clark and J. Goold, *Thermodynamics of precision in quantum nonequilibrium steady states*, Phys. Rev. Research **1**, 033021 (2019), doi:10.1103/PhysRevResearch.1.033021.

[86] G. Falasco, M. Esposito and J.-C. Delvenne, *Unifying thermodynamic uncertainty relations*, New Journal of Physics **22**(5), 053046 (2020), doi:10.1088/1367-2630/ab8679.

[87] T. Koyuk, U. Seifert and P. Pietzonka, *A generalization of the thermodynamic uncertainty relation to periodically driven systems*, Journal of Physics A: Mathematical and Theoretical **52**(2), 02LT02 (2018), doi:10.1088/1751-8121/aaeec4.

[88] T. Koyuk and U. Seifert, *Operationally accessible bounds on fluctuations and entropy production in periodically driven systems*, Phys. Rev. Lett. **122**, 230601 (2019), doi:10.1103/PhysRevLett.122.230601.

[89] T. Van Vu and Y. Hasegawa, *Thermodynamic uncertainty relations under arbitrary control protocols*, Phys. Rev. Research **2**, 013060 (2020), doi:10.1103/PhysRevResearch.2.013060.

[90] H. J. D. Miller, M. H. Mohammady, M. Perarnau-Llobet and G. Guarnieri, *Thermodynamic uncertainty relation in slowly driven quantum heat engines*, Phys. Rev. Lett. **126**, 210603 (2021), doi:10.1103/PhysRevLett.126.210603.

[91] L. D. Landau and E. M. Lifshitz, *Statistical Physics: Volume 5*, vol. 5, Elsevier, ISBN 9780750633727 (2013).

[92] G. E. Crooks, *Measuring thermodynamic length*, Phys. Rev. Lett. **99**, 100602 (2007), doi:10.1103/PhysRevLett.99.100602.

[93] Y.-H. Ma, D. Xu, H. Dong and C.-P. Sun, *Universal constraint for efficiency and power of a low-dissipation heat engine*, Phys. Rev. E **98**, 042112 (2018), doi:10.1103/PhysRevE.98.042112.

[94] P. Salamon, K. H. Hoffmann, S. Schubert, R. S. Berry and B. Andresen, *What conditions make minimum entropy production equivalent to maximum power production?* **26**(1), 73 (2001), doi:doi:10.1515/JNETDY.2001.006.

[95] B. Andresen and J. M. Gordon, *Constant thermodynamic speed for minimizing entropy production in thermodynamic processes and simulated annealing*, Phys. Rev. E **50**, 4346 (1994), doi:10.1103/PhysRevE.50.4346.

[96] P. Salamon, A. Nitzan, B. Andresen and R. S. Berry, *Minimum entropy production and the optimization of heat engines*, Phys. Rev. A **21**, 2115 (1980), doi:10.1103/PhysRevA.21.2115.

[97] P. Salamon, J. Nulton, G. Siragusa, T. Andersen and A. Limon, *Principles of control thermodynamics*, Energy **26**(3), 307 (2001), doi:https://doi.org/10.1016/S0360-5442(00)00059-1.

[98] B. Andresen, *Current trends in finite-time thermodynamics*, Angewandte Chemie International Edition **50**(12), 2690 (2011), doi:https://doi.org/10.1002/anie.201001411, https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.201001411.

[99] K. Brandner, M. Bauer and U. Seifert, *Universal coherence-induced power losses of quantum heat engines in linear response*, Phys. Rev. Lett. **119**, 170602 (2017), doi:10.1103/PhysRevLett.119.170602.

[100] V. Holubec and A. Ryabov, *Maximum efficiency of low-dissipation heat engines at arbitrary power*, Journal of Statistical Mechanics: Theory and Experiment **2016**(7), 073204 (2016), doi:10.1088/1742-5468/2016/07/073204.

[101] A. C. Hernández, A. Medina and J. M. M. Roco, *Time, entropy generation, and optimization in low-dissipation heat devices*, New Journal of Physics **17**(7), 075011 (2015), doi:10.1088/1367-2630/17/7/075011.

[102] V. Holubec and Z. Ye, *Maximum efficiency of low-dissipation refrigerators at arbitrary cooling power*, Phys. Rev. E **101**, 052124 (2020), doi:10.1103/PhysRevE.101.052124.

[103] J. P. Palao, R. Kosloff and J. M. Gordon, *Quantum thermodynamic cooling cycle*, Phys. Rev. E **64**, 056130 (2001), doi:10.1103/PhysRevE.64.056130.

[104] N. Brunner, N. Linden, S. Popescu and P. Skrzypczyk, *Virtual qubits, virtual temperatures, and the foundations of thermodynamics*, Phys. Rev. E **85**, 051117 (2012), doi:10.1103/PhysRevE.85.051117.

[105] A. Ros, *The isoperimetric problem*, Global theory of minimal surfaces **2**, 175 (2001).

[106] E. Parini, *An introduction to the cheeger problem*, Surv. Math. Appl. **6**, 9 (2011).

[107] G. P. Leonardi, *An Overview on the Cheeger Problem*, pp. 117–139, Springer International Publishing, Cham, ISBN 978-3-319-17563-8, doi:10.1007/978-3-319-17563-8_6 (2015).

[108] V. Blåsjö, *The isoperimetric problem*, The American Mathematical Monthly **112**(6), 526 (2005), doi:10.1080/00029890.2005.11920227.

[109] H. J. D. Miller, M. Scandi, J. Anders and M. Perarnau-Llobet, *Work fluctuations in slow processes: Quantum signatures and optimal control*, Phys. Rev. Lett. **123**, 230603 (2019), doi:10.1103/PhysRevLett.123.230603.

[110] T. Denzler and E. Lutz, *Power fluctuations in a finite-time quantum carnot engine*, Phys. Rev. Research **3**, L032041 (2021), doi:10.1103/PhysRevResearch.3.L032041.

[111] N. Pancotti, M. Scandi, M. T. Mitchison and M. Perarnau-Llobet, *Speed-ups to isothermality: Enhanced quantum thermal machines through control of the system-bath coupling*, Phys. Rev. X **10**, 031015 (2020), doi:10.1103/PhysRevX.10.031015.

[112] M. Scandi, D. Barker, S. Lehmann, K. A. Dick, V. F. Maisi and M. Perarnau-Llobet, *Constant dissipation rate is optimal for thermodynamic protocols: experimental implementation of landauer erasure through thermodynamic length*, doi:10.48550/ARXIV.2209.01852 (2022).

[113] R.-X. Zhai, F.-M. Cui, Y.-H. Ma, C. P. Sun and H. Dong, *Experimental implementation of finite-time carnot cycle*, doi:10.48550/ARXIV.2206.10153 (2022).

[114] M. Esposito, R. Kawai, K. Lindenberg and C. Van den Broeck, *Quantum-dot carnot engine at maximum power*, Phys. Rev. E **81**, 041106 (2010), doi:10.1103/PhysRevE.81.041106.

[115] O. Abah, J. Roßnagel, G. Jacob, S. Deffner, F. Schmidt-Kaler, K. Singer and E. Lutz, *Single-ion heat engine at maximum power*, Phys. Rev. Lett. **109**, 203006 (2012), doi:10.1103/PhysRevLett.109.203006.

[116] K. Zhang, F. Bariani and P. Meystre, *Quantum optomechanical heat engine*, Phys. Rev. Lett. **112**, 150602 (2014), doi:10.1103/PhysRevLett.112.150602.

[117] J. Deng, Q.-h. Wang, Z. Liu, P. Hänggi and J. Gong, *Boosting work characteristics and overall heat-engine performance via shortcuts to adiabaticity: Quantum and classical systems*, Phys. Rev. E **88**, 062122 (2013), doi:10.1103/PhysRevE.88.062122.

[118] E. Torrontegui, S. Ibáñez, S. Martínez-Garaot, M. Modugno, A. del Campo, D. Guéry-Odelin, A. Ruschhaupt, X. Chen and J. G. Muga, *Chapter 2 - shortcuts to adiabaticity*, In E. Arimondo, P. R. Berman and C. C. Lin, eds., *Advances in Atomic, Molecular, and Optical Physics*, vol. 62 of *Advances In Atomic, Molecular, and Optical Physics*, pp. 117–169. Academic Press, doi:https://doi.org/10.1016/B978-0-12-408090-4.00002-5 (2013).

[119] A. d. Campo, J. Goold and M. Paternostro, *More bang for your buck: Super-adiabatic quantum engines*, Scientific Reports **4**(1), 6208 (2014), doi:10.1038/srep06208.

[120] P. W. Claeys, M. Pandey, D. Sels and A. Polkovnikov, *Floquet-engineering counterdiabatic protocols in quantum many-body systems*, Phys. Rev. Lett. **123**, 090602 (2019), doi:10.1103/PhysRevLett.123.090602.

[121] T. Villazon, A. Polkovnikov and A. Chandran, *Swift heat transfer by fast-forward driving in open quantum systems*, Phys. Rev. A **100**, 012126 (2019), doi:10.1103/PhysRevA.100.012126.

[122] A. E. Allahverdyan, K. V. Hovhannisyan, A. V. Melkikh and S. G. Gevorkian, *Carnot cycle at finite power: Attainability of maximal efficiency*, Phys. Rev. Lett. **111**, 050601 (2013), doi:10.1103/PhysRevLett.111.050601.

[123] R. Kosloff and Y. Rezek, *The quantum harmonic otto cycle*, Entropy **19**(4) (2017), doi:10.3390/e19040136.

[124] E. Geva and R. Kosloff, *A quantum-mechanical heat engine operating in finite time. a model consisting of spin-1/2 systems as the working fluid*, The Journal of Chemical Physics **96**(4), 3054 (1992), doi:10.1063/1.461951.

[125] V. Cavina, A. Mari, A. Carlini and V. Giovannetti, *Variational approach to the optimal control of coherently driven, open quantum system dynamics*, Phys. Rev. A **98**, 052125 (2018), doi:10.1103/PhysRevA.98.052125.

[126] U. Seifert, *Stochastic thermodynamics, fluctuation theorems and molecular machines*, Reports on Progress in Physics **75**(12), 126001 (2012), doi:10.1088/0034-4885/75/12/126001.

[127] P. Menczel and K. Brandner, *Limit cycles in periodically driven open quantum systems*, Journal of Physics A: Mathematical and Theoretical **52**(43), 43LT01 (2019), doi:10.1088/1751-8121/ab435a.

[128] M. G. A. Paris, *Quantum estimation for quantum technology*, International Journal of Quantum Information **07**(supp01), 125 (2009), doi:10.1142/S0219749909004839.

[129] G. Tóth and I. Apellaniz, *Quantum metrology from a quantum information science perspective*, Journal of Physics A: Mathematical and Theoretical **47**(42), 424006 (2014), doi:10.1088/1751-8113/47/42/424006.

[130] J. S. Sidhu and P. Kok, *Geometric perspective on quantum parameter estimation*, AVS Quantum Science **2**(1), 014701 (2020), doi:10.1116/1.5119961.

[131] W.-K. Mok, K. Bharti, L.-C. Kwek and A. Bayat, *Optimal probes for global quantum thermometry*, Communications Physics **4**(1), 62 (2021), doi:10.1038/s42005-021-00572-w.

[132] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT Press, `http://www.deeplearningbook.org` (2016).

[133] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, doi:10.48550/ARXIV.1412.6980 (2014).

[134] *D-wave systems*.

[135] G. Mussardo, *Statistical field theory: an introduction to exactly solved models in statistical physics*, Oxford University Press, ISBN 9780198788102 (2010).

[136] M. E. Fisher, *The theory of equilibrium critical phenomena*, Reports on Progress in Physics **30**(2), 615 (1967), doi:10.1088/0034-4885/30/2/306.

[137] N. S. Izmailian and C.-K. Hu, *Exact amplitude ratio and finite-size corrections for the $m \times n$ square lattice ising model*, Phys. Rev. E **65**, 036103 (2002), doi:10.1103/PhysRevE.65.036103.

[138] A. B. Dodds, V. Kendon, C. S. Adams and N. Chancellor, *Practical designs for permutation-symmetric problem hamiltonians on hypercubes*, Phys. Rev. A **100**, 032320 (2019), doi:10.1103/PhysRevA.100.032320.

[139] P. Sekatski and M. Perarnau-Llobet, *Optimal nonequilibrium thermometry in finite time*, doi:10.48550/ARXIV.2107.04425 (2021).

[140] S. Milz and K. Modi, *Quantum stochastic processes and quantum non-markovian phenomena*, PRX Quantum **2**, 030201 (2021), doi:10.1103/PRXQuantum.2.030201.

[141] A. Rivas and S. F. Huelga, *Open quantum systems*, vol. 10, Springer, doi:10.1007/978-3-642-23354-8 (2012).

[142] R. Vasile, S. Olivares, M. A. Paris and S. Maniscalco, *Continuous-variable quantum key distribution in non-markovian channels*, Phys. Rev. A **83**, 042321 (2011), doi:10.1103/PhysRevA.83.042321.

[143] A. W. Chin, S. F. Huelga and M. B. Plenio, *Quantum metrology in non-markovian environments*, Phys. Rev. Lett. **109**, 233601 (2012), doi:10.1103/PhysRevLett.109.233601.

[144] S. F. Huelga, A. Rivas and M. B. Plenio, *Non-markovianity-assisted steady state entanglement*, Phys. Rev. Lett. **108**, 160402 (2012), doi:10.1103/PhysRevLett.108.160402.

[145] B. Bylicka, D. Chruściński and S. Maniscalco, *Non-Markovianity and reservoir memory of quantum channels: a quantum information theory perspective*, Scientific Reports **4**(1), 5720 (2014), doi:10.1038/srep05720.

[146] E.-M. Laine, H.-P. Breuer and J. Piilo, *Nonlocal memory effects allow perfect teleportation with mixed states*, Scientific Reports **4**(1), 4620 (2014), doi:10.1038/srep04620.

[147] W. F. Stinespring, *Positive functions on c*-algebras*, Proceedings of the American Mathematical Society **6**(2), 211 (1955), doi:https://doi.org/10.1090/S0002-9939-1955-0069403-4.

[148] F. Buscemi and N. Datta, *Equivalence between divisibility and monotonic decrease of information in classical and quantum stochastic processes*, Phys. Rev. A **93**, 012101 (2016), doi:10.1103/PhysRevA.93.012101.

[149] B. Bylicka, D. Chruściński and S. Maniscalco, *Non-markovianity and reservoir memory of quantum channels: a quantum information theory perspective*, Scientific Reports **4** (2014), doi:10.1038/srep05720.

[150] S. Lorenzo, F. Plastina and M. Paternostro, *Geometrical characterization of non-markovianity*, Phys. Rev. A **88**, 020102 (2013), doi:10.1103/PhysRevA.88.020102.

[151] D. D. Santis and M. Johansson, *Equivalence between non-markovian dynamics and correlation backflows*, New Journal of Physics **22**(9), 093034 (2020), doi:10.1088/1367-2630/abaf6a.

[152] J. Kołodyński, S. Rana and A. Streltsov, *Entanglement negativity as a universal non-markovianity witness*, Phys. Rev. A **101**, 020303 (2020), doi:10.1103/PhysRevA.101.020303.

[153] W. K. Wootters, *Statistical distance and hilbert space*, Phys. Rev. D **23**, 357 (1981), doi:10.1103/PhysRevD.23.357.

[154] C. R. Rao, *Information and the accuracy attainable in the estimation of statistical parameters*, Kotz S., Johnson N.L. (eds) Breakthroughs in Statistics. Springer Series in Statistics (Perspectives in Statistics) (1992), doi:10.1007/978-1-4612-0919-5.

[155] H. Chernoff, *A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the sum of Observations*, The Annals of Mathematical Statistics **23**(4), 493 (1952), doi:10.1214/aoms/1177729330.

[156] S. Kullback, *Information theory and statistics*, Dover. Publications, ISBN 978-0486696843 (1997).

[157] I. Csiszár and P. C. Shields, *Information Theory and Statistics: A Tutorial*, Foundations and Trends® in Communications and Information Theory **1**(4), 417 (2004), doi:10.1561/0100000004.

[158] X.-M. Lu, X. Wang and C. P. Sun, *Quantum fisher information flow and non-markovian processes of open systems*, Phys. Rev. A **82**, 042103 (2010), doi:10.1103/PhysRevA.82.042103.

[159] L. L. Campbell, *An extended čencov characterization of the information metric*, Proceedings of the American Mathematical Society **98**(1), 135 (1986), doi:https://doi.org/10.2307/2045782.

[160] D. Petz, *Monotone metrics on matrix spaces*, Linear Algebra and its Applications **244**, 81 (1996), doi:https://doi.org/10.1016/0024-3795(94)00211-8.

[161] G. Lindblad, *On the generators of quantum dynamical semigroups*, Communications in Mathematical Physics **48**(2), 119 (1976), doi:10.1007/BF01608499.

[162] V. Gorini, A. Kossakowski and E. C. G. Sudarshan, *Completely positive dynamical semigroups of n-level systems*, Journal of Mathematical Physics **17**(5), 821 (1976), doi:10.1063/1.522979.

[163] M. J. W. Hall, J. D. Cresser, L. Li and E. Andersson, *Canonical form of master equations and characterization of non-markovianity*, Phys. Rev. A **89**, 042120 (2014), doi:10.1103/PhysRevA.89.042120.

[164] D. Chruściński and S. Maniscalco, *Degree of non-markovianity of quantum evolution*, Phys. Rev. Lett. **112**, 120404 (2014), doi:10.1103/PhysRevLett.112.120404.

[165] F. Buscemi and V. Scarani, *Fluctuation theorems from bayesian retrodiction*, Phys. Rev. E **103**, 052111 (2021), doi:10.1103/PhysRevE.103.052111.

[166] J. Surace and M. Scandi, *State retrieval beyond bayes' retrodiction and reverse processes* (2022), doi:10.48550/ARXIV.2201.09899.

[167] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, *Proposed experiment to test local hidden-variable theories*, Phys. Rev. Lett. **23**, 880 (1969), doi:10.1103/PhysRevLett.23.880.

[168] B. S. Cirel'son, *Quantum generalizations of Bell's inequality*, Letters in Mathematical Physics **4**(2), 93 (1980), doi:10.1007/BF00417500.

[169] R. Colbeck, *Quantum and Relativistic Protocols For Secure Multi-Party Computation*, Ph.D. thesis, University of Cambridge, doi:10.48550/ARXIV.0911.3814 (2007).

[170] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning and C. Monroe, *Random numbers certified by Bell's theorem*, Nature **464**(7291), 1021 (2010), doi:10.1038/nature09008.

[171] A. K. Ekert, *Quantum cryptography based on bell's theorem*, Phys. Rev. Lett. **67**, 661 (1991), doi:10.1103/PhysRevLett.67.661.

[172] A. Acín, N. Gisin and L. Masanes, *From bell's theorem to secure quantum key distribution*, Phys. Rev. Lett. **97**, 120405 (2006), doi:10.1103/PhysRevLett.97.120405.

[173] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio and V. Scarani, *Device-independent security of quantum cryptography against collective attacks*, Phys. Rev. Lett. **98**, 230501 (2007), doi:10.1103/PhysRevLett.98.230501.

[174] I. Šupić and J. Bowles, *Self-testing of quantum systems: a review*, Quantum **4**, 337 (2020), doi:10.22331/q-2020-09-30-337.

[175] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri *et al.*, *Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres*, Nature **526**(7575), 682 (2015), doi:10.1038/nature15759.

[176] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-A. Larsson, C. Abellán, W. Amaya, V. Pruneri *et al.*, *Significant-loophole-free test of Bell's theorem with entangled photons*, Phys. Rev. Lett. **115**, 250401 (2015), doi:10.1103/PhysRevLett.115.250401.

[177] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer *et al.*, *Strong loophole-free test of local realism*, Phys. Rev. Lett. **115**, 250402 (2015), doi:10.1103/PhysRevLett.115.250402.

[178] S. M. Tan, D. F. Walls and M. J. Collett, *Nonlocality of a single photon*, Phys. Rev. Lett. **66**, 252 (1991), doi:10.1103/PhysRevLett.66.252.

[179] L. Hardy, *Nonlocality of a single photon revisited*, Phys. Rev. Lett. **73**, 2279 (1994), doi:10.1103/PhysRevLett.73.2279.

[180] C. C. Gerry, *Nonlocality of a single photon in cavity qed*, Phys. Rev. A **53**, 4583 (1996), doi:10.1103/PhysRevA.53.4583.

[181] L. Vaidman, *Nonlocality of a single photon revisited again*, Phys. Rev. Lett. **75**, 2063 (1995), doi:10.1103/PhysRevLett.75.2063.

[182] Y. Aharonov and L. Vaidman, *Nonlocal aspects of a quantum wave*, Phys. Rev. A **61**, 052108 (2000), doi:10.1103/PhysRevA.61.052108.

[183] B. Hessmo, P. Usachev, H. Heydari and G. Björk, *Experimental demonstration of single photon nonlocality*, Phys. Rev. Lett. **92**, 180401 (2004), doi:10.1103/PhysRevLett.92.180401.

[184] J. B. Brask, R. Chaves and N. Brunner, *Testing nonlocality of a single photon without a shared reference frame*, Phys. Rev. A **88**, 012111 (2013), doi:10.1103/PhysRevA.88.012111.

[185] O. Morin, J.-D. Bancal, M. Ho, P. Sekatski, V. D'Auria, N. Gisin, J. Laurat and N. Sangouard, *Witnessing trustworthy single-photon entanglement with local homodyne measurements*, Phys. Rev. Lett. **110**, 130401 (2013), doi:10.1103/PhysRevLett.110.130401.

[186] T. Das, M. Karczewski, A. Mandarino, M. Markiewicz, B. Woloncewicz and M. Żukowski, *Can single photon excitation of two spatially separated modes lead to a violation of bell inequality via weak-field homodyne measurements?*, New Journal of Physics **23**(7), 073042 (2021), doi:10.1088/1367-2630/ac0ffe.

[187] B. Yurke and D. Stoler, *Bell's-inequality experiments using independent-particle sources*, Phys. Rev. A **46**, 2229 (1992), doi:10.1103/PhysRevA.46.2229.

[188] B. Yurke and D. Stoler, *Einstein-podolsky-rosen effects from independent particle sources*, Phys. Rev. Lett. **68**, 1251 (1992), doi:10.1103/PhysRevLett.68.1251.

[189] C. Branciard, N. Gisin and S. Pironio, *Characterizing the nonlocal correlations created via entanglement swapping*, Phys. Rev. Lett. **104**, 170401 (2010), doi:10.1103/PhysRevLett.104.170401.

[190] C. Branciard, D. Rosset, N. Gisin and S. Pironio, *Bilocal versus nonbilocal correlations in entanglement-swapping experiments*, Phys. Rev. A **85**, 032119 (2012), doi:10.1103/PhysRevA.85.032119.

[191] D. Cavalcanti, M. L. Almeida, V. Scarani and A. Acín, *Quantum networks reveal quantum nonlocality*, Nature Communications **2**(1), 184 (2011), doi:10.1038/ncomms1193.

[192] T. Fritz, *Beyond Bell's Theorem II: Scenarios with Arbitrary Causal Structure*, Communications in Mathematical Physics **341**(2), 391 (2016), doi:10.1007/s00220-015-2495-5.

[193] D. Cavalcanti, R. Rabelo and V. Scarani, *Nonlocality tests enhanced by a third observer*, Phys. Rev. Lett. **108**, 040402 (2012), doi:10.1103/PhysRevLett.108.040402.

[194] T. Van Himbeeck, J. Bohr Brask, S. Pironio, R. Ramanathan, A. B. Sainz and E. Wolfe, *Quantum violations in the instrumental scenario and their relations to the Bell scenario*, Quantum **3**, 186 (2019), doi:10.22331/q-2019-09-16-186.

[195] M.-O. Renou, Y. Wang, S. Boreiri, S. Beigi, N. Gisin and N. Brunner, *Limits on correlations in networks for quantum and no-signaling resources*, Phys. Rev. Lett. **123**, 070403 (2019), doi:10.1103/PhysRevLett.123.070403.

[196] I. Šupić, J.-D. Bancal and N. Brunner, *Quantum nonlocality in networks can be demonstrated with an arbitrarily small level of independence between the sources*, Phys. Rev. Lett. **125**, 240403 (2020), doi:10.1103/PhysRevLett.125.240403.

[197] T. Kraft, S. Designolle, C. Ritz, N. Brunner, O. Gühne and M. Huber, *Quantum entanglement in the triangle network*, Phys. Rev. A **103**, L060401 (2021), doi:10.1103/PhysRevA.103.L060401.

[198] A. Pozas-Kerstjens, N. Gisin and M.-O. Renou, *Proofs of network quantum nonlocality aided by machine learning* (2022), doi:10.48550/ARXIV.2203.16543.

[199] S. Boreiri, A. Girardin, B. Ulu, P. Lypka-Bartosik, N. Brunner and P. Sekatski, *Towards a minimal example of quantum nonlocality without inputs* (2022), doi:10.48550/ARXIV.2207.08532.

[200] M.-O. Renou and S. Beigi, *Network nonlocality via rigidity of token counting and color matching*, Phys. Rev. A **105**, 022408 (2022), doi:10.1103/PhysRevA.105.022408.

[201] T. Kriváchy, Y. Cai, D. Cavalcanti, A. Tavakoli, N. Gisin and N. Brunner, *A neural network oracle for quantum nonlocality problems in networks*, npj Quantum Information **6**(1), 70 (2020), doi:10.1038/s41534-020-00305-x.

[202] R. Jozsa and N. Linden, *On the role of entanglement in quantum-computational speed-up*, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **459**(2036), 2011 (2003), doi:10.1098/rspa.2002.1097.

[203] G. Tóth and I. Apellaniz, *Quantum metrology from a quantum information science perspective*, Journal of Physics A: Mathematical and Theoretical **47**(42), 424006 (2014), doi:10.1088/1751-8113/47/42/424006.

[204] O. Gühne and G. Tóth, *Entanglement detection*, Physics Reports **474**(1), 1 (2009), doi:https://doi.org/10.1016/j.physrep.2009.02.004.

[205] D. Rosset, R. Ferretti-Schöbitz, J.-D. Bancal, N. Gisin and Y.-C. Liang, *Imperfect measurement settings: Implications for quantum state tomography and entanglement witnesses*, Phys. Rev. A **86**, 062325 (2012), doi:10.1103/PhysRevA.86.062325.

[206] T. Moroder, O. Gühne, N. Beaudry, M. Piani and N. Lütkenhaus, *Entanglement verification with realistic measurement devices via squashing operations*, Phys. Rev. A **81**, 052342 (2010), doi:10.1103/PhysRevA.81.052342.

[207] H.-K. Lo, M. Curty and B. Qi, *Measurement-device-independent quantum key distribution*, Phys. Rev. Lett. **108**, 130503 (2012), doi:10.1103/PhysRevLett.108.130503.

[208] S. Pirandola, C. Ottaviani, G. Spedalieri, C. Weedbrook, S. L. Braunstein, S. Lloyd, T. Gehring, C. S. Jacobsen and U. L. Andersen, *High-rate measurement-device-independent quantum cryptography*, Nature Photonics **9**(6), 397 (2015), doi:10.1038/nphoton.2015.83.

[209] Z. Li, Y.-C. Zhang, F. Xu, X. Peng and H. Guo, *Continuous-variable measurement-device-independent quantum key distribution*, Phys. Rev. A **89**, 052301 (2014), doi:10.1103/PhysRevA.89.052301.

[210] X.-C. Ma, S.-H. Sun, M.-S. Jiang, M. Gui and L.-M. Liang, *Gaussian-modulated coherent-state measurement-device-independent quantum key distribution*, Phys. Rev. A **89**, 042335 (2014), doi:10.1103/PhysRevA.89.042335.

[211] I. Šupić, P. Skrzypczyk and D. Cavalcanti, *Measurement-device-independent entanglement and randomness estimation in quantum networks*, Phys. Rev. A **95**, 042340 (2017), doi:10.1103/PhysRevA.95.042340.

[212] B. Regula, L. Lami, G. Ferrari and R. Takagi, *Operational quantification of continuous-variable quantum resources*, Phys. Rev. Lett. **126**, 110403 (2021), doi:10.1103/PhysRevLett.126.110403.

[213] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory*, vol. 1, Edizioni della Normale Pisa, doi:10.1007/978-88-7642-378-9 (2011).

[214] G. M. D'Ariano, L. Maccone and M. G. A. Paris, *Quorum of observables for universal quantum estimation*, Journal of Physics A: Mathematical and General **34**(1), 93 (2000), doi:10.1088/0305-4470/34/1/307.

[215] E. C. G. Sudarshan, *Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams*, Phys. Rev. Lett. **10**, 277 (1963), doi:10.1103/PhysRevLett.10.277.

[216] R. J. Glauber, *Coherent and incoherent states of the radiation field*, Phys. Rev. **131**, 2766 (1963), doi:10.1103/PhysRev.131.2766.

[217] L.-M. Duan, G. Giedke, J. I. Cirac and P. Zoller, *Inseparability criterion for continuous variable systems*, Phys. Rev. Lett. **84**, 2722 (2000), doi:10.1103/PhysRevLett.84.2722.

[218] R. Simon, *Peres-horodecki separability criterion for continuous variable systems*, Phys. Rev. Lett. **84**, 2726 (2000), doi:10.1103/PhysRevLett.84.2726.

[219] H. Yuen and M. Lax, *Multiple-parameter quantum estimation and measurement of nonselfadjoint observables*, IEEE Transactions on Information Theory **19**(6), 740 (1973), doi:10.1109/TIT.1973.1055103.

[220] M. G. Genoni, M. G. A. Paris, G. Adesso, H. Nha, P. L. Knight and M. S. Kim, *Optimal estimation of joint parameters in phase space*, Phys. Rev. A **87**, 012107 (2013), doi:10.1103/PhysRevA.87.012107.

[221] S. Morelli, A. Usui, E. Agudelo and N. Friis, *Bayesian parameter estimation using gaussian states and measurements*, Quantum Science and Technology **6**(2), 025018 (2021), doi:10.1088/2058-9565/abd83d.

[222] Z. Zhou, Y. Sheng, P. Niu, L. Yin, G. Long and L. Hanzo, *Measurement-device-independent quantum secure direct communication*, Science China Physics, Mechanics & Astronomy **63**(3), 230362 (2019), doi:10.1007/s11433-019-1450-8.

[223] D. Rosset, F. Buscemi and Y.-C. Liang, *Resource theory of quantum memories and their faithful verification with minimal assumptions*, Phys. Rev. X **8**, 021033 (2018), doi:10.1103/PhysRevX.8.021033.

[224] Y. Mao, Y.-Z. Zhen, H. Liu, M. Zou, Q.-J. Tang, S.-J. Zhang, J. Wang, H. Liang, W. Zhang, H. Li, L. You, Z. Wang *et al.*, *Experimentally verified approach to nonentanglement-breaking channel certification*, Phys. Rev. Lett. **124**, 010502 (2020), doi:10.1103/PhysRevLett.124.010502.

[225] F. Graffitti, A. Pickston, P. Barrow, M. Proietti, D. Kundys, D. Rosset, M. Ringbauer and A. Fedrizzi, *Measurement-device-independent verification of quantum channels*, Phys. Rev. Lett. **124**, 010503 (2020), doi:10.1103/PhysRevLett.124.010503.

[226] Y. Yu, P.-F. Sun, Y.-Z. Zhang, B. Bai, Y.-Q. Fang, X.-Y. Luo, Z.-Y. An, J. Li, J. Zhang, F. Xu, X.-H. Bao and J.-W. Pan, *Measurement-device-independent verification of a quantum memory*, Phys. Rev. Lett. **127**, 160502 (2021), doi:10.1103/PhysRevLett.127.160502.

[227] P. Abiuso and A. Acín, Quantum memory verification for continuous-variable systems *(In preparation)* .

[228] H. Dourdent, A. A. Abbott, N. Brunner, I. Šupić and C. Branciard, *Semi-device-independent certification of causal nonseparability with trusted quantum inputs* (2021), doi:10.48550/ARXIV.2107.10877.

[229] R. Uzdin, A. Levy and R. Kosloff, *Equivalence of quantum heat machines, and quantum-thermodynamic signatures*, Phys. Rev. X **5**, 031044 (2015), doi:10.1103/PhysRevX.5.031044.

[230] M. Lostaglio, *Certifying quantum signatures in thermodynamics and metrology via contextuality of quantum linear response*, Phys. Rev. Lett. **125**, 230603 (2020), doi:10.1103/PhysRevLett.125.230603.

[231] M. Ohya and D. Petz, *Quantum entropy and its use*, Springer Science & Business Media (2004).

This page is intentionally left blank.