

ULTIMATE PRECISION BOUNDS ON QUANTUM THERMOMETRY

A dissertation submitted in partial fulfilment of the
requirements for the degree of philosophiae doctor

Mathias Rønnow Jørgensen
September 2021

QUANTUM PHYSICS AND INFORMATION TECHNOLOGY
DEPARTMENT OF PHYSICS
TECHNICAL UNIVERSITY OF DENMARK

Project period:
September 2018 – September 2021

Supervisors:
Associate Prof. Jonatan B. Brask
Associate Prof. Alexander Huck

External research stay:
University of Geneva
Prof. Nicolas Brunner

Contact:
mathias93@me.com

Copyright:
© 2021 Mathias Rønnow Jørgensen

Summary

This theoretical PhD project investigates the fundamental limits to the precision with which temperature can be estimated in quantum systems. The first part of the thesis presents a Bayesian formulation of temperature estimation theory, which is based on the thermodynamic length between thermal sample states. We show how the resulting temperature estimation theory can be mapped onto a Euclidean parameter estimation problem, and this insight makes it possible to apply precision bounds developed in this context. When mapping these bound from the Euclidean theory, back to the space of temperatures, it is found that the measure of precision resulting from a framework based on the thermodynamic length, has a natural interpretation as a generalized relative error.

The Bayesian approach to temperature estimation, in which a priori incomplete information is assumed, points to the need of effective adaptive strategies, with which the measurement strategy can be updated as temperature information is extracted. We propose two such adaptive strategies. In the case of equilibrium probe thermometry, we show that no non-adaptive measurement strategy can exhibit Heisenberg-like scaling with the probe dimension. Furthermore, it is illustrated that an adaptive strategy can restore Heisenberg-like scaling with the probe dimension. These results highlight the essential role of adaptation for thermometry.

The last part of this thesis investigates optimal thermometry under realistic constraints on the available measurements. In particular, we focus on the challenges associated with thermometry at very low temperatures. It is found that finite energy-resolution fundamentally constrains the attainable sensitivity of a measurement to temperature in the ultracold regime. The relevance of the derived bounds is illustrated by considering a numerically exact simulation of a probe-based thermometry protocol. Furthermore, we investigate fundamental limitations under constraints of measurements with only a finite number of distinguishable outcomes, and derive the associated precision bounds. It is shown that for the majority of many-body quantum systems, there exist coarse-grained measurements achieving a temperature sensitivity comparable to that of the many-body system itself.

Resume

Dette teoretiske PhD projekt undersøger de fundamentale begrænsninger af præcisionen hvormed temperatur kan måles i kvantesystemer. I den første del af afhandlingen præsenterer vi en Bayesisk formulering af temperatur estimerings problemet, som er baseret på den termodynamiske længde mellem termiske tilstande. Vi viser at temperatur estimerings problemet har en ækvivalent formulering som et Euklidisk estimations problem, og bruger denne indsigt til at anvende fundamentale præcisions begrænsninger udviklet i den Euklidiske kontekst. Når disse udtrykkes direkte som begrænsninger på temperatur målinger, finder vi et mål for præcision som naturligt kan fortolkes som en generaliseret relativ fejl.

Den Bayesiske tilgang til temperatur estimering, hvori kun begrænset information om temperature er tilgængelig a priori, motiverer formuleringen af effektive adaptive strategier, hvormed en målings-strategi kan opdateres i takt med at temperatur information bliver tilgængelig. Vi præsenterer to adaptive strategier. For ligevægts-termometri med kvante prober viser vi at ingen ikke-adaptive strategi kan realisere Heisenberg-lignende skalering med dimensionen af proben. Derudover viser vi at hvis målingen kan adapteres, så kan Heisenberg-lignende skalering med probens dimension genetableres. Disse resultater illustrerer vigtigheden af adaptation i termometri.

I den sidste del af afhandlingen undersøger vi optimal termometri under realistiske begrænsninger på tilgængelige målinger. Især fokuserer vi på de udfordringer der er forbundne med termometri ved meget lave temperature. Vi finder at målinger med endelig energiopløsning er fundamentalt begrænsede med hensyn til deres sensitivitet overfor meget lave temperature. Vi illustrerer relevansen af disse begrænsninger ved at studere en numerisk eksakt løsning af en probe-baseret termometri protokol. Derudover undersøger vi fundamentale begrænsninger når det antages at kun målinger med endeligt mange forskellige udfald er tilgængelige. Vi viser at for størstedelen af kvante mangelegeme systemer, eksisterer der sådanne målinger som realiserer en sensitivitet overfor temperatur som er sammenlignelig med temperatur sensitiviteten af mangelegeme systemet selv.

Preface

This thesis was completed at the section for Quantum Physics and Information Technology (QPIT) at the Department of Physics, as part of the fulfilment of the requirements for the PhD degree. The project was carried out under the supervision of Associate Prof. Jonatan B. Brask, with Associate Prof. Alexander Huck acting as co-supervisor, and was funded by the Independent Research Fund Denmark. As part of my PhD project I spent three months of 2020 in the group of Prof. Nicolas Brunner at the Department of Applied Physics at the University of Geneva. In Geneva I worked directly with scientific collaborator Martí Perarnau-Llobet.

First and foremost I would like to thank my direct supervisor Jonatan. I deeply appreciate the trust he has shown me in providing me with the freedom to pursue my research interests in my own way. His competent guidance, and constantly good mood, has been a unwavering source of support, especially when considering that he has had to balance the requirements of a young family, a growing research group and a global pandemic.

I would also like to thank the group in Geneva, and Prof. Nicolas Brunner for hosting my visit. In particular, I would like to thank Martí Perarnau-Llobet and Mohammad Mehboudi for welcoming me into the group both scientifically and socially. Martí has been an invaluable resource in driving research projects forward, and putting me in contact with collaborators around the world.

My gratitude also goes out to the QPIT group at the Technical University of Denmark. In particular, the other PhDs and Post. Docs. working with Jonatan: Bradley, Kieran, Carles, Anders and Michael. After having been the only non-senior member of the group for a good two years, and after being kept at home by a global pandemic, it is invaluable to have the support of other group members who are always up for a discussion of either physics or other things of interest.

In addition I would like to thank my family home in Jutland. Although the work for my PhD often kept me away more than it should have, you have always been a source of love and support. Lastly, I would like to thank Mie. Both for being a superb, albeit unintentional, office mate for a year. But also for being a daily reminder that there is more to life than physics.

Dissemination of research

Publications included in this thesis:

Mohammad Mehboudi, **Mathias R. Jørgensen**, Stella Seah, Jonatan B. Brask, Jan Kołodyński and Martí Perarnau-Llobet, "*Fundamental limits in Bayesian thermometry and attainability via adaptive strategies*", arXiv ID:2108.05932 (Submitted to Physical Review Letters).

Mathias R. Jørgensen, Mohammad Mehboudi, Jan Kołodyński, Martí Perarnau-Llobet and Jonatan B. Brask, "*Bayesian quantum thermometry based on thermodynamic length*", arXiv ID:2108.05901 (Submitted to Physical Review A).

Karen V. Hovhannisyanyan, **Mathias R. Jørgensen**, Gabriel T. Landi, Álvaro M. Alhambra, Jonatan B. Brask and Martí Perarnau-Llobet, "*Optimal Quantum Thermometry with Coarse-grained Measurements*", PRX Quantum **2**, 020322 (2021).

Mathias R. Jørgensen, Patrick P. Potts, Matteo G. A. Paris and Jonatan B. Brask, "*Tight bound on finite-resolution quantum thermometry at low temperatures*", Phys. Rev. Research **2**, 033394 (2020).

Publications beyond the scope of the thesis:

Mathias R. Jørgensen and Felix A. Pollock, "*A discrete memory-kernel for multi-time correlations in non-Markovian quantum processes*", Phys. Rev. A **102**, 052206 (2020)¹.

Mathias R. Jørgensen and Felix A. Pollock, "*Exploiting the causal tensor network structure of quantum processes to efficiently simulate non-Markovian path integrals*", Phys. Rev. Lett. **123**, 240602 (2019)².

¹This work was initiated during my MSc thesis, but significantly expanded and eventually completed during my PhD

²This work was initiated during my MSc thesis but significantly expanded and eventually completed during my PhD. This paper forms the basis of the numerical simulations performed in chapter 6, and is included in appendix B for reference.

Conference contributions:

"Optimal Quantum Thermometry with Coarse-grained Measurements", Oral presentation, Lake Como School on quantum thermodynamic systems and processes (2021).

"Tight bound on finite-resolution quantum thermometry at low temperatures", Oral presentation, WE-Heraeus-Seminar on Quantum Thermodynamics for Young Scientists (2020).

"Exploiting the causal tensor network structure of quantum processes to efficiently simulate non-Markovian path integrals", Oral presentation, Polaron day, Technical University of Vienna (2019).

"A discrete memory-kernel for multi-time correlations in non-Markovian quantum processes", Oral presentation, Poster presentation, YQIS Conference and CoQuS Summer School (2018).

Contents

Summary	v
Resume	vii
Preface	ix
Dissemination of research	xi
1 Introductory material	1
1.1 Introduction	1
1.2 Basic theoretical framework	3
1.2.1 Quantum mechanical modelling	3
1.2.2 Quantum thermodynamic modelling	6
1.3 Outline of the thesis	7
2 Estimation Theory: Quantum Model and Invariance Conditions	11
2.1 Introduction	11
2.2 Quantum statistical model and likelihood	11
2.2.1 Manifold of quantum states	11
2.2.2 Measurements and the likelihood function	12
2.2.3 Likelihood for adaptive measurements	13
2.3 Frequentist estimation	14
2.3.1 The error-propagation formula	14
2.3.2 The frequentist Cramér-Rao bound	16
2.4 Bayesian inference	17
2.4.1 Probability measure on the manifold	17
2.4.2 Bayesian updating	17
2.4.3 Criteria for parameterization invariance	18
2.5 Invariant estimation theory	19
2.5.1 Criteria for invariant point estimation	19
2.5.2 Singling out a parameter estimate	20
2.5.3 Distance function and the mean-squared error	21
2.5.4 Metric structure and the geodesic distance	22
2.5.5 Invariant MAP estimates	24
2.5.6 Selecting an initial prior probability	26
2.6 Properties of the Fisher information	26
2.6.1 Derivation of the quantum Fisher information	27
2.6.2 Additivity of independent measurements	27

2.7	Concluding remarks	28
3	Estimation Theory: Precision Bounds and Adaptive Strategies	29
3.1	Introduction	29
3.2	Posterior mean-squared error	29
3.2.1	Minimal mean-squared error estimator	30
3.3	Posterior Cramér-Rao bound	31
3.3.1	Saturability of the CRB	32
3.3.2	Asymptotic statistics	33
3.4	Bayesian mean-squared error	35
3.4.1	Total variance decomposition	36
3.4.2	Bayesian Cramér-Rao bound	37
3.4.3	Van Trees bound	38
3.5	Adaptive measurement design	40
3.5.1	Asymptotic expression for the added information	40
3.5.2	Alternative optimization strategy	41
3.6	Concluding remarks	41
4	Temperature Estimation Based on the Thermodynamic Length	43
4.1	Introduction	43
4.1.1	Motivation	43
4.1.2	Outline of chapter	44
4.2	Canonical thermal states	45
4.2.1	Quantum Fisher information metric	45
4.2.2	Statistical interpretation of thermodynamic length	46
4.3	The thermalizing channel	48
4.4	Non-interacting spin one-half particles	49
4.5	Difficulty of thermometry at low temperatures	52
4.5.1	Relation to conventional thermometry	52
4.5.2	Thermometry of a Bose-Einstein condensate	54
4.5.3	Thermometry of a tight-binding chain	57
4.6	Concluding remarks	60
5	Fundamental Limits in Bayesian Thermometry and Attainability via Adaptive Strategies	61
5.1	Introduction	61
5.2	Preliminaries and setup.	63
5.3	Main results	64
5.4	Case study	65
5.5	Conclusions and future directions	68
5.6	Appendix	68
5.6.1	Van Trees inequality in equilibrium thermometry	68
5.6.2	Model-independent super-extensive upper bound on Γ	69
5.6.3	Tight upper bound on the thermal energy density	69
5.6.4	Extensive bounds for the non-adaptive scenario	70
5.6.5	Model-independent bound for a Gaussian prior	71

6	Tight bound on finite-resolution quantum thermometry at low temperatures	73
6.1	Introduction	73
6.2	Temperature estimation	75
6.2.1	Quantifying the estimation precision	75
6.2.2	Accounting for measurement limitations	76
6.2.3	Low-temperature scaling behaviour	78
6.3	Scaling bound for large systems	79
6.3.1	Finite-resolution criterion	79
6.3.2	Finite-resolution bound	81
6.3.3	Proving tightness of bound	82
6.4	Generalization to noisy measurements	83
6.4.1	Noisy temperature measurements	83
6.4.2	Illustration of noisy measurement	84
6.5	Single-qubit probe	87
6.5.1	Measurement protocol	87
6.5.2	Excitation-preserving interaction	88
6.5.3	Excitation-non-preserving interaction	88
6.6	Conclusion	90
6.7	Appendix	91
6.7.1	Density of states for a bosonic bath	91
6.7.2	Scaling behaviour for the noisy model	91
6.7.3	The non-excitation-preserving interaction as a noisy POVM	93
6.7.4	Tensor network simulation	93
7	Optimal Quantum Thermometry with Coarse-grained Measurements	97
7.1	Introduction	97
7.2	Framework	98
7.3	Optimal coarse graining	100
7.3.1	Optimal POVM	100
7.3.2	Optimal binning	102
7.3.3	Illustrative examples	104
7.3.4	General remarks and extension to imperfect measurements	107
7.4	Many-body lattice models	109
7.4.1	Gaussian density of states	109
7.4.2	Noncritical, interacting systems on lattices	112
7.4.3	Critical systems	113
7.5	Probe-based measurements	116
7.5.1	Jaynes–Cummings model	117
7.6	Conclusions	119
A	Probability theory	121
B	Exploiting the Causal Tensor Network Structure of Quantum Processes to Efficiently Simulate Non-Markovian Path Integrals	125
B.1	Introduction	125
B.2	Process tensor framework	127
B.3	Gaussian influence functional	127

B.4	Tensor network simulation	129
B.5	Network complexity for a two level system	131
B.6	Conclusion	133
B.7	Appendix	134
	Bibliography	141

Chapter 1

Introductory material

1.1 Introduction

Quantum mechanics provide the best theory we have for describing natural phenomena occurring at all length and energy scales [1]. Recent decades have seen an explosion of interest in the so-called second quantum revolution, in which highly controlled quantum systems form the basis for quantum-enhanced technologies [2]. Potential quantum technologies include secure communication systems, powerful computing machines, highly efficient material simulations, and quantum-enhanced sensing protocols. In particular, quantum sensing holds the promise for highly accurate determination of physical quantities of interest [3, 4]. Examples include the measurement of electric and magnetic fields, but also extends to the accurate determination of the minuscule space-time distortions associated with gravitational waves [5].

A topic intimately connected to sensing is what we might call characterization. When constructing controlled quantum devices, e.g. a simulation system or a sensor, the state of the system, and by extension how it operates, typically depends on a set of classical, or latent, parameters [6]. In the case of a sensor, the set of latent parameters might be attributed a significance extrinsic to the sensor itself. I.e. we want to know these parameters and the sensor is the tool we apply to that end. The aim is then to measure the sensing system, and from the observed data, infer the most probable values of the latent parameters. A different, but deeply connected, scenario might be that the latent parameters are intrinsic to our model of the controlled quantum device. In this case we face a characterization problem, i.e. we want to accurately determine the latent parameters for the sake of modelling the operation of the quantum device. In both cases, the fundamental limits to estimation would be determined by the sensitivity of the system to small changes in the latent parameters. However, the physical implications of a given sensitivity depends on the scenario. If the system manifest a vanishing sensitivity to the latent parameter, it implies that it constitutes a poor sensor. However, at the same time it signifies that no detailed knowledge of the latent parameter is required to accurately model the operation of the system.

Principle among such latent parameters is the operating temperature of the device, typically taken to describe energy dissipation between the device and its surroundings [7]. Determining the temperature of a physical system is an essential task throughout science and technology. Current attempts at pushing the limits of accurate temperature determination typically involves the use of controlled quantum systems

acting as temperature sensors, or probes. Examples include, but are not limited to, the following:

1. Explorations of real-time biological function of living organisms via in-vivo nanometer-scale temperature sensors, e.g. optically monitored colour centers in nanodiamonds [8, 9, 10]. Biological systems can typically be modelled locally as heat reservoirs with a well-defined temperature, and by monitoring the physics of continuously thermalizing quantum probes, it is possible to extract information on the temperature.
2. Precise characterization of ultracold atomic gases confined in optical lattices, e.g. verification of a state preparation procedure, exploring thermodynamic phase diagrams or investigating transport phenomena [11, 12, 13, 14, 15]. Traditionally, temperature estimation in ultracold atomic gases involved time-of-flight or in-situ absorption imaging, both of which are inherently destructive [16]. Recently there has been proposals in which impurity atoms act as nonequilibrium quantum probes of the gas temperature [17, 18, 19, 20], and such strategies could provide non-destructive alternatives.
3. Within the field of microelectronic physics, precise thermometry techniques play a crucial role in mapping out dissipative phenomena [21, 22]. State-of-the-art thermometry methods typically employ nanoscale superconducting quantum interference devices, and has enabled thermal-imaging measurements of energy dissipation at the fundamental Landauer bound [23, 24]. More broadly, the availability of high-precision thermometry techniques makes it possible to study quantum thermodynamic phenomena, involving the interplay of heat, work and information, in microelectronic systems [25, 26].
4. Lastly, we mention that precise thermometry techniques are also crucial in experiments with micro-mechanical resonators. In particular, thermometry plays a key role in the verification of cooling protocols attempting to prepare the resonator near its ground state of motion [27, 28, 29]. The accurate characterization of such resonators would pave the way for the observation of quantum superposition states of macroscopic quantum objects.

With the exception of the exploration of biological function in living organisms, all the examples mentioned involves phenomena occurring at very low temperatures. In fact, a sizeable part of modern physics focus on phenomena occurring at very low temperatures [30]. The principal reason is that as a system is cooled down its distinctly quantum mechanical features become increasingly manifest. As an example we mention the study of quantum phase transitions occurring at zero temperature in some quantum many-body systems [31]. In the case of a second-order phase transition, an ordered phase in which quantum correlations are manifest at all length scales, emerge at the critical point. It is well-known that quantum many-body systems tuned to criticality, constitutes a key resource for quantum-enhanced metrology [32]. However, finite temperature effects generally result in the fading away of these phenomena, and thus constitute an unwanted nuisance. In particular, a zero-temperature critical point typically broadens into a critical region, and the system exhibit a smooth crossover between different phases [31, 33]. Thus, certifying the presence or absence of a quantum phase

transition, requires accurate determination of the system temperature deep into the ultracold regime.

The study of quantum phase transitions is but one example of the importance of accurate temperature determination, with respect to the detailed characterization, control and ultimately utilization, of examined quantum systems and processes. The subject of this thesis is the fundamental question of how accurately it is possible to determine the temperature of a quantum system. In Sec. 1.3 we describe this problem in greater detail, and outline the specific contribution of this thesis. First, however, in Sec. 1.2 we will outline the basic theoretical framework with which we will approach the problem. In particular, we focus on the role played by temperature within a quantum mechanical model of a physical experiment.

1.2 Basic theoretical framework

1.2.1 Quantum mechanical modelling

The following section describes basic quantum mechanics [1, 34]. In particular, the treatment is inspired by Gammelmark [35]. We make use of some basic notions of measure-theoretic probability theory. For an in-depth treatment see Ref. [36]. A brief discussion is provided in Appendix A.

Quantum mechanics asserts that a physical system is associated with a Hilbert space \mathcal{H} , and that a dynamical variable of the system is represented by a Hermitian operator on \mathcal{H} . The eigenvalues of the operator correspond to the possible values of the dynamical variable. An experiment involving the system is divided into a state preparation procedure followed by a measurement process. Quantum mechanics aims to assign probabilities to the various possible outcomes of the measurement process. The state preparation procedure is represented by a state operator $\hat{\rho}$ on \mathcal{H} , which is Hermitian, non-negative and of unit trace. We will discuss the state operator in the following section. The measurement process is represented by a measurement operator, $\hat{\Omega}$, mapping an element of the space of measurement outcomes, \mathcal{X} , to a bounded operator on \mathcal{H} . We write this as

$$\hat{\Omega} : \mathcal{X} \rightarrow \hat{\Omega}(x) \in \mathcal{B}(\mathcal{H}), \quad (1.1)$$

where $x \in \mathcal{X}$ and $\mathcal{B}(\mathcal{H})$ denotes the space of bounded operators on \mathcal{H} . Given a reference probability measure $\mathbb{P}_0 : \mathcal{X} \rightarrow \mathbb{R}$ on the space of possible measurement outcomes, here \mathbb{R} denotes the real numbers, the measurement operators must satisfy the normalization condition

$$\int_{\mathcal{X}} d\mathbb{P}_0(x) \hat{\Omega}(x)^\dagger \hat{\Omega}(x) = \mathbb{1}. \quad (1.2)$$

The reference measure is to be understood as a degree of freedom in the theory, i.e. the measurement operators can be specified relative to a convenient reference process [35]. We will only make rudimentary use of the notion of a measure. For our purposes we can simply understand it as a function which takes a subset of outcomes $\mathcal{A} \subseteq \mathcal{X}$ and assigns to this subset a probability. Note, that in the case of a discrete sample space, \mathcal{X} , the reference measure reduce to an arbitrary set of weights, $\int d\mathbb{P}_0(x) = \sum \mathbb{P}_0(x)$, and can be eliminated via a rescaling, $\hat{\Omega}(x) \rightarrow \sqrt{\mathbb{P}_0(x)} \hat{\Omega}(x)$, of the measurement operators [35]. Hence, in this case we recover the more standard formulation of quantum measurement theory [1, 34].

The operator $\hat{\Pi}(x) := \hat{\Omega}(x)^\dagger \hat{\Omega}(x)$ is called an effect operator, and is both Hermitian and non-negative. Together with the reference measure, the effect operator forms a positive operator-valued measure (POVM) on \mathcal{X} given by $d\mathbb{P}_0(x) \hat{\Pi}(x)$. The statistical postulate of quantum mechanics asserts that the space of possible measurement outcomes is equipped with a probability measure given by the Hilbert-Schmidt inner product of the POVM and the state operator, that is

$$d\mathbb{P}(x) = d\mathbb{P}_0(x) \operatorname{Tr}(\hat{\Pi}(x)\hat{\rho}), \quad (1.3)$$

where Tr denotes the trace. Furthermore, the post-measurement state operator, conditional on observing a measurement outcome $x \in \mathcal{X}$, is given by $\hat{\rho}|x := \hat{\Omega}(x)\hat{\rho}\hat{\Omega}(x)^\dagger$. Note that the conditional state operator is generally sub-normalized, that is, the trace is less than one. Furthermore, we see that predicting the post-measurement state requires knowledge of the measurement operators. If we only have available the set of effect operators, then it is not possible to predict the post-measurement state. The reason for this is that a given effect operator does not have a single unique decomposition in terms of measurement operators [34].

We now ask the question whether it is possible to measure the value of a dynamical variable of the system. The spectral theorem states that to any dynamical variable \hat{Q} there corresponds a unique family of non-negative projection operators $\hat{E}(q)$, for $q \in \mathbb{R}$, onto the subspace of \hat{Q} with eigenvalues less than q , such that \hat{Q} can be written on the form [1]

$$\hat{Q} = \int_{\mathbb{R}} d\hat{E}(q) q, \quad (1.4)$$

where $d\hat{E}(q)$ denotes the POVM associated with the family of projection operators. A key concept in probability theory is that of a measurable function $Q : \mathcal{X} \rightarrow \mathbb{R}$. For our purposes, we can think of a measurable function as one with which we can meaningfully associate an image measure defined as [35, 36]

$$d\hat{\Pi}_Q(q) := d\mathbb{P}_0(Q^{-1}(q)) \hat{\Pi}(Q^{-1}(q)), \quad q \in \mathbb{R}, \quad (1.5)$$

where Q^{-1} is the pre-image of the measurable function. The image measure takes the POVM on \mathcal{X} and maps it to a POVM on the space of function values. Given the image measure, we say that a given dynamical variable can be measured if it is possible to design an experiment, and construct a measurable function Q , such that the corresponding image measure equals the POVM of the dynamical variable, that is

$$d\hat{\Pi}_Q(q) = d\hat{E}(q). \quad (1.6)$$

Furthermore, from the statistical postulate it follows that the space of possible values of the dynamical variable is equipped with a probability measure $d\mathbb{P}_Q(q) = \operatorname{Tr}(d\hat{E}(q)\hat{\rho})$. Lastly, we note that although all dynamical variables are associated with a POVM, it is not true that all POVMs correspond to a dynamical variable. In other words, most measurements do not correspond directly to measuring a dynamical variable.

Given a physical experiment, quantum mechanics allows us to assign probabilities to possible measurement outcomes. Typically we are interested in statistical properties of the experiment. If we consider again an arbitrary measurable function $Q : \mathcal{X} \rightarrow Q(x) \in \mathbb{R}$ on the space of measurement outcomes, then the expectation value is defined as

$$\langle Q \rangle := \int_{\mathcal{X}} d\mathbb{P}(x) Q(x) = \int_{\mathbb{R}} d\mathbb{P}_Q(q) q. \quad (1.7)$$

If an infinite number of identical repetitions of the experiment is performed, then the average value of the function Q computed over the acquired data converges to the expectation. If the image measure $d\mathbb{P}_Q(q)$ equals the POVM associated with a dynamical variable \hat{Q} , then it is possible to write the expectation value of the dynamical variable on the simple form

$$\langle \hat{Q} \rangle = \text{Tr}(\hat{Q}\hat{\rho}). \quad (1.8)$$

Note that an expectation value is not something which is measured in an experiment. Rather, an expectation must be estimated based on obtained measurement data. This is also true for asymptotic statistical quantities beyond the expectation, e.g. the variance

$$\text{Var}(Q) = \langle (Q - \langle Q \rangle)^2 \rangle. \quad (1.9)$$

Although the dynamical variables of a system are not what we exclusively, or even typically, measure directly in an experiment, they play an essential role in the construction of the measurement operators. As a simple example we consider a dynamical variable of particular importance – the system energy. The system energy is represented by the energy operator $\hat{H}(\boldsymbol{\xi})$, also called the system Hamiltonian, that typically depends on a vector of extrinsic control parameters $\boldsymbol{\xi}$. These controls could for instance be applied magnetic or electric fields. The particular importance of the Hamiltonian operator, stems from the fact that it serves as the generator of the intrinsic evolution of the system in time [1]. That is, if the system is closed, and initially in a state $\hat{\rho}_i$, then over a time duration t the state will evolve to

$$\hat{\rho}(t, \boldsymbol{\xi}) = e^{-it\hat{H}(\boldsymbol{\xi})} \hat{\rho}_i e^{it\hat{H}(\boldsymbol{\xi})}. \quad (1.10)$$

Intrinsic evolution is typically an idealization as most physical systems are not closed [37, 38], however, it will be sufficient as an illustration. If we imagine a measurement in which the system evolves intrinsically over a time duration t , and is subsequently subjected to a measurement process represented by measurement operators $\hat{\Omega}$, then from Eq. (1.3) it follows that the probability measure takes the form

$$\begin{aligned} d\mathbb{P}(x; t, \boldsymbol{\xi}) &= d\mathbb{P}_0(x) \text{Tr}(\hat{\Pi}(x)\hat{\rho}(t, \boldsymbol{\xi})), \\ &= d\mathbb{P}_0(x) \text{Tr}(\hat{\Pi}(x; t, \boldsymbol{\xi}) \hat{\rho}_i), \end{aligned} \quad (1.11)$$

where the second equality follows from the cyclic property of the trace and from defining the effect and measurement operators

$$\hat{\Pi}(x; t, \boldsymbol{\xi}) := \hat{\Omega}(x; t, \boldsymbol{\xi})^\dagger \hat{\Omega}(x; t, \boldsymbol{\xi}), \quad (1.12)$$

$$\hat{\Omega}(x; t, \boldsymbol{\xi}) := \hat{\Omega}(x) e^{-it\hat{H}(\boldsymbol{\xi})}. \quad (1.13)$$

The second equation shows that the full measurement operator is given by a time-ordered product of an operator describing the intrinsic evolution, generated by the Hamiltonian, and an operator describing the final measurement. Thus, we see that the simple scenario of an unmeasured system is captured by the system Hamiltonian. This illustrates the importance of the dynamical variables of the system, namely, they are the building blocks from which we construct measurement operators.

At this point we have encountered two kinds of objects: 1) The POVMs, e.g. the one associated with a dynamical variable like the energy, which are in principle directly

measurable, and 2) the latent variables, such as the time duration t or the extrinsic control parameters $\boldsymbol{\xi}$, which are intrinsic to the mechanical description [6]. Latent variables are by definition not directly measurable. To determine their values we must either assume that sufficient information is available a priori to fix the value, or we must resort to indirectly estimating the value via the statistics of the obtained measurement data [6]. This is also true for asymptotic quantities such as an expectation value. In this sense expectations themselves can be considered as latent variables of the model.

1.2.2 Quantum thermodynamic modelling

In the preceding section we saw how systems and processes are modelled quantum mechanically. In this section, we outline what is meant by a quantum thermodynamic model, in the specific context considered in this thesis. It should be noted, however, that the subject of quantum thermodynamic modelling extends far beyond the rudimentary picture discussed here [39, 40].

A key ingredient of the quantum mechanical description of a physical experiment is the initial state operator, $\hat{\rho}_i$, of the system. Once the state operator is specified, the statistical properties of any conceivable measurement of the system is fixed. In principle, given complete microscopic information, it is possible to mechanically specify the state preparation procedure. In practice, however, such complete microscopic information is rarely available for systems of even modest complexity.

Following E.T. Jaynes [41, 42], we say that a quantum thermodynamic description comes about, when prior information can be expressed as constraints on the expectation values of a set of dynamical variables. Furthermore, we adopt an assumption of sufficiency, namely, we assume that if the set of expectation values are known precisely, then that knowledge is sufficient to fix the statistics of any other measurement of the system. In particular, we define the internal energy of the system as the expectation of the Hamiltonian operator

$$U(\boldsymbol{\xi}) := \langle \hat{H}(\boldsymbol{\xi}) \rangle = \text{Tr}(\hat{H}(\boldsymbol{\xi}) \hat{\rho}_i). \quad (1.14)$$

We then imagine a thermodynamic scenario in which sufficient prior information exist to fix the values of the internal energy $U(\boldsymbol{\xi})$ and the latent parameters $\boldsymbol{\xi}$. Given these constraints, we can ask the inverse problem of finding an initial state operator $\hat{\rho}_i$ consistent with the fixed variables. Based on information-theoretic arguments, the so-called maximum-entropy principle, it can be shown that the state operator must take the form of a thermodynamic equilibrium state [41, 42]

$$\hat{\rho}(T; \hat{H}(\boldsymbol{\xi})) := e^{-\hat{H}(\boldsymbol{\xi})/k_B T} / \mathcal{Z}(T; \hat{H}(\boldsymbol{\xi})), \quad (1.15)$$

where k_B is the Boltzmann constant, $T \in \mathbb{R}_+$ is a parameter known as the system temperature, and we have defined the partition function

$$\mathcal{Z}(T; \hat{H}(\boldsymbol{\xi})) := \text{Tr}(e^{-\hat{H}(\boldsymbol{\xi})/k_B T}). \quad (1.16)$$

Given the thermodynamic equilibrium state, it follows that the temperature of the system is itself a latent variable of the quantum mechanical model. It encodes the constraint that the internal energy is fixed. To be specific, the temperature must be such that the following relation is satisfied [41, 42]

$$U(\boldsymbol{\xi}) = k_B T^2 \partial_T \log \mathcal{Z}(T; \hat{H}(\boldsymbol{\xi})), \quad (1.17)$$

where \log refers to the natural logarithm. Information-theoretic reasoning is not the only way of arriving at the thermodynamic equilibrium state as the suitable description of a quantum system. Alternatively, we could consider dynamic thermalization in quantum systems [7, 39, 43, 44]. Rather than pursuing this topic further, however, throughout this thesis we take the suitability of the thermodynamic equilibrium state as a given phenomenological fact.

1.3 Outline of the thesis

Above we saw how our quantum thermodynamic model of a physical system depends on the set of latent variables $(T, \boldsymbol{\xi})$. Knowledge of these variables determines the initial state of the system, and thus fix the statistics of any conceivable measurement which could be performed. The setting considered in this thesis is one in which the latent parameters $\boldsymbol{\xi}$ are taken to be known precisely, that is the system Hamiltonian is perfectly specified. On the other hand, the system temperature is taken to be only partially known. In other words, we consider an experimental scenario in which the suitability of the family of thermodynamic equilibrium states is taken for granted. However, it is not known precisely which state in the family is the correct one.

The subject of this thesis is the exploration of the fundamental limits to how well temperature can be determined in quantum systems. In particular, we focus on the challenges emerging when attempting to estimate the temperature of quantum systems in the ultracold regime. This subject matter is referred to as quantum thermometry, and aims at guiding the design of optimal thermometric measurement processes, and the processing of measurement data into an accurate temperature. The subject matter of quantum thermometry is rapidly expanding. For recent reviews see Refs. [45, 46]. The humble contributions of this thesis is outlined below:

1. In chapter 2 we develop the Bayesian estimation theory forming the backbone of the remaining thesis. In contrast to previous approaches to quantum thermometry [45, 46], we construct the estimation theory around the estimation of the thermal state itself, rather than the temperature parameter. We find that this approach forces us to adopt measures of precision respecting the invariance properties of the quantum states. In particular, we find that the only consistent way of constructing a distance function between different states, is an operational approach in which the distance is gauged by the ability of a reference measurement to distinguish the states. The mathematical formalism required to complete this construction is well developed [47, 48], and our contribution here is merely in the realization that these tools are relevant for quantum thermometry.
2. In chapter 3 we build on the foundations put down in chapter 2, and start from the key insight, which constitutes our humble contribution, that a thermometry problem can be formally mapped onto a Euclidean parameter estimation problem. Given this correspondence, it follows straightforwardly that we can apply precision bounds known from Euclidean parameter estimation theory [6]. The chapter goes through the development of known results on precision bounds from estimation theory within the Euclidean picture. In addition we discuss single-shot adaptive strategies, and propose two distinct optimization strategies.

3. In chapter 4 we apply the developed estimation theory to the problem of thermometry. The operational distance measure proposed in chapter 2 is shown to be equivalent to the concept of thermodynamic length [49]. The chapter goes through a number of applications of the developed framework: 1) Thermometry of a heat reservoir, 2) thermometry involving spin-1/2 particles, 3) adaptive probe thermometry of a non-interacting Bose gas, and 4) local thermometry of a fermionic tight-binding chain. Common to these examples is the realization that an analysis based on the thermodynamic length force us to consider a generalized relative error. In particular, the thermal sensitivity of a given measurement strategy must be determined relative to the thermal sensitivity of the thermal system itself. This point seems intuitive, perhaps even trivial, but it has not previously been properly appreciated within the field of quantum thermometry [45, 46].
4. In chapter 5 we consider the specific case of equilibrium probe thermometry of a heat reservoir. Previous work has shown that the achievable precision, when using a quantum probe of a certain dimension, is fundamentally limited to scale with the square of the probe dimension [50]. This conclusion is generally reached within the local estimation regime, in which good knowledge of the temperature is available a priori. In this chapter, we show that if we start out with only limited information on the temperature, then the precision scaling is limited to scale linearly in the probe dimension, rather than quadratically. The key constraint, which brings about this loss of super-linear scaling, is that we are not allowed to adaptively change the employed measurement strategy as new information is acquired. If we are able to update the measurement, then we show, and numerically illustrate, that the super-linear scaling can be restored. Thus this chapter highlight the essential role played by adaptation within quantum thermometry.
5. In chapter 6 we discuss the fundamental bounds on the scaling of the precision with temperature, as the temperature approach absolute zero. In particular, we seek to take into account realistic constraints on available measurement strategies. We propose a finite-resolution criterion, and show that this constraint on the available measurements, gives rise to a tighter bound on temperature scaling of the precision, as compared to a previous bound which was derived based on the unattainability principle [51]. Furthermore, we show that the tightened bound can be saturated. We both provide a formal proof of this statement, and we also demonstrate tightness in the specific scenario of monitoring the nonequilibrium dynamics of a spin-1/2 probe interacting with a thermal phononic environment. Interestingly, the enhanced sensitivity of the spin-1/2 probe is only realized when full account is taken of non-Markovian effects in the dynamical evolution. We model these effects using the tensor network algorithm outlined in appendix B.
6. In chapter 7 we take a slightly different approach to the question of fundamental precision bounds. Instead of considering finite-resolution measurements, we consider the fundamental limitations which arise in cases where the available measurement has only a finite number of distinct observable outcomes. This scenario is referred to as coarse-grained thermometry. In the chapter we develop a prescription for how optimal coarse-grained measurement are to be designed, and we study the performance of the optimal solutions for a number of different many-body thermal sample systems. From an experimental point of view, perhaps the

most relevant result of this investigation, is the construction of ultimate bounds on the precision achievable, when employing finite-dimensional quantum probes as thermometers.

Chapter 2

Estimation Theory: Quantum Model and Invariance Conditions

Parts of this chapter contains text and figures from Ref. [52]. It will be explicitly indicated when a given section or figure has appeared in Ref. [52].

2.1 Introduction

The aim of this chapter is to formulate the problem of temperature estimation as a problem of statistical inference — see Refs. [6, 53, 54, 55, 56] for standard treatments of this topic. One task of statistical inference theory is to determine how best to extract information on a quantity of interest from an experimental data set, where the data set is generated according to a known stochastic process. In this chapter we take a general approach, i.e. one not specific to temperature estimation, and formulate the problem as that of identifying a quantum state which is known to belong to a known family of states. Such a family is typically specified by a number of parameters, and the identification task can thus be formulated as a parameter estimation problem.

The structure of the chapter is as follows: In Sec. 2.2 we describe how the preparation and measurement of a system is described within the quantum mechanical formalism. In Sec. 2.3 we briefly review the frequentist approach to parameter estimation. In Sec. 2.4 we begin our treatment of the Bayesian approach to estimation, in particular we describe how probabilities on a parameter space are updated. In Sec. 2.5 we develop an invariant estimation theory, where invariance refers to the fact that it is typically possible to work with multiple equivalent parameterizations of the family of quantum states. We end the chapter in Sec. 2.6 where we outline some technical results.

2.2 Quantum statistical model and likelihood

2.2.1 Manifold of quantum states

Let us first define the quantity of interest, i.e., what we are trying to estimate. Consider a physical system, referred to as the sample system. Within quantum theory, a system is associated with a Hilbert space \mathcal{H} , and the system state is represented by a linear operator $\hat{\rho}$ on \mathcal{H} , which is non-negative, Hermitian and of unit trace — see for instance Ref. [1]. We are interested in a prepare-and-measure scenario, in which the

experimental background conditions, e.g. the nature of the experimental setup and the specific preparation procedure, are such that the true quantum state of the sample being prepared, $\hat{\rho}_{\text{true}}$, is an element of a smooth manifold of quantum states, \mathcal{S} . A manifold is a topological space that locally resembles a Euclidean space, and a smooth manifold is a manifold on which the notion of differentiability exists. We will only make rudimentary use of these concepts, and further details can be found in the mathematics literature — see for instance Refs. [47, 57, 58]. The general question of how experimental background conditions are mapped to a specific manifold is beyond the scope of the present discussion, and here we simply take the existence of the manifold for granted. Any smooth manifold of states can be equipped with a system of coordinates, or in other words it can be given a continuous parameterization. If we consider a parameter space $\Theta \subseteq \mathbb{R}$ then

$$\mathcal{S} = \{\hat{\rho}_{\Theta}(\theta) \text{ for } \theta \in \Theta\} := \mathcal{S}_{\Theta}, \quad (2.1)$$

where $\hat{\rho}_{\Theta}(\theta)$ denotes a Θ -parameterized quantum state-operator on the manifold. Throughout this thesis, we focus on the case where the manifold is one-dimensional, and simply note that a great deal of the presented formalism generalises to the multi-dimensional case — see for instance Ref. [47].

We consider a large number of repetitions of the prepare-and-measure scenario. The state preparation procedure is assumed perfectly reproducible, i.e. the sample state $\hat{\rho}_{\text{true}}$ is prepared with certainty in each repetition of the experiment. The true state of the sample is thus well-defined, however, we assume that the true state is imprecisely known to the experimenter, and our task is to provide an estimate of this state. Following the approach of Braunstein and Caves [48], we note that since the sample state is an element of a parameterized manifold of states, the state estimation problem can equivalently be formulated as a parameter estimation problem. We should keep in mind, however, that we are fundamentally interested in the state itself. This point is relevant as there might exist a number of different ways in which to parameterize the manifold of states. E.g., thermal states can be parameterized by the temperature or the inverse temperature, and a parameter estimate in one parameterization must be consistent with an estimate in another parameterization.

2.2.2 Measurements and the likelihood function

The second question we address is how to acquire information about the true sample state. In a single repetition of the prepare-and-measure scenario, an experimenter probes the sample state by performing a measurement of the sample system, and thus acquires measurement data. The basic setup is illustrated in Fig. 2.1. The acquired data is generally represented as a specific realization of a stochastic variable X , taking values $x \in \mathcal{X}$ where \mathcal{X} denotes the data space of the measurement. This data space could be a discrete set, or, in the case of continuous measurements, the data space could refer to a set of continuous functions on a given interval of time.

Within quantum mechanics, a measurement is represented by a set of measurement operators $\hat{\Omega}(x)$ for $x \in \mathcal{X}$, which are bounded operators on the system Hilbert space. The measurement operators are generally defined with respect to a reference probability measure $\mathbb{P}_0 : \mathcal{X} \rightarrow \mathbb{R}$, which constitutes a degree of freedom in the theory [35], and which allows us to treat discrete and continuous sample spaces in a unified way. The probability of observing an outcome within the subset $\mathcal{A} \subseteq \mathcal{X}$, given the sample-system

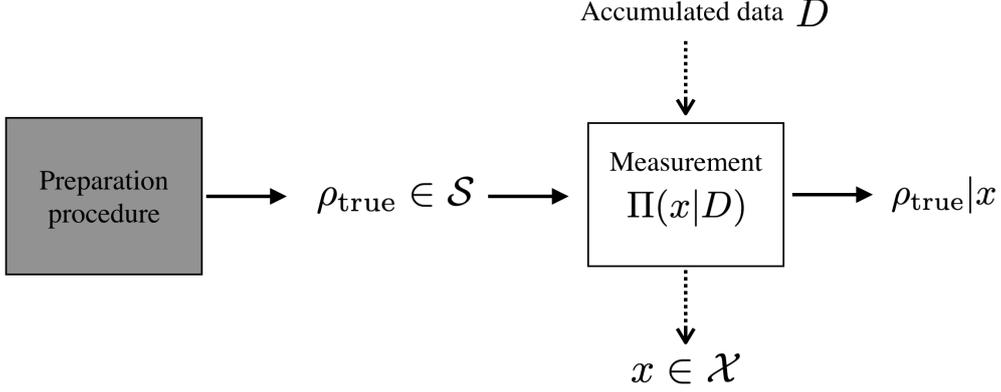


Figure 2.1: Illustration of estimation scenario. A true state of the sample system is generated according to an unknown preparation procedure (black box). The sample is subjected to a measurement, producing an outcome $x \in \mathcal{X}$. The measurement is designed based on the accumulated data, denoted D . After the measurement the sample is in a state $\rho_{\text{true}}|x$, which is conditional on the observed outcome. Throughout this thesis, we only consider scenarios in which the sample state is discarded after the measurement and a fresh state is prepared.

state labelled by $\theta \in \Theta$, is given by Born's rule [35]

$$\mathbb{P}(x \in \mathcal{A}|\theta) = \int_{\mathcal{A}} d\mathbb{P}(x|\theta), = \int_{\mathcal{A}} d\mathbb{P}_0(x) \text{Tr} \left[\hat{\Pi}(x) \hat{\rho}_{\Theta}(\theta) \right], \quad (2.2)$$

where $\hat{\Pi}(x) := \hat{\Omega}(x)^\dagger \hat{\Omega}(x)$ is called an effect operator. From here on will refer to the operator $\hat{\Pi}(x)$ as an element of the positive operator-valued measure (POVM) representing the implemented measurement. Strictly speaking this is not correct, as it is the combination $d\mathbb{P}_0(x)\Pi(x)$ which constitutes the POVM element. However, this slight abuse of nomenclature will not lead to any serious confusion. Furthermore, we will drop the hats on operators whenever it does not lead to any confusion.

The quantity $\text{Tr} [\Pi(x)\rho_{\Theta}(\theta)]$ serves as a generalized probability density function (PDF) on the data space \mathcal{X} , with respect to the reference probability measure \mathbb{P}_0 . Throughout we will refer to this PDF as the likelihood function associated with the measurement

$$p_{\mathcal{X}}(x|\theta) := \text{Tr} [\Pi(x)\rho(\theta)] = \frac{d\mathbb{P}(x|\theta)}{d\mathbb{P}_0(x)}, \quad (2.3)$$

where the second equality serves to remind us that the likelihood function is formally defined as the Radon-Nikodym derivative with respect to a reference measure [35]. Any measurement, for which the associated likelihood function can be written in the form of Eq. (2.3), is considered a single-shot measurement. Notice that for a single-shot measurement, the POVM element representing the measurement is required to be independent of the sample state.

2.2.3 Likelihood for adaptive measurements

The most general estimation protocol we will consider consists of multiple single-shot measurements performed in a time-ordered sequence. For a sequence consisting of n

measurements, the accumulated measurement data is represented by a trajectory of measurement outcomes

$$\mathbf{x}_n := \{x_1, \dots, x_n\}, \quad (2.4)$$

where $x_k \in \mathcal{X}_k$ and \mathcal{X}_k denotes the data space associated with the k th measurement. For later convenience we denote the full data space by $\mathcal{X}_n := \mathcal{X}_1 \times \dots \times \mathcal{X}_n$. In general, we do not require the implemented measurements to be identical, nor do we require the measurements to be independent. By dependent measurements, we mean that the POVM elements describing the j th measurement, can depend on the past measurement trajectory \mathbf{x}_{j-1} . We will express the likelihood function associated with a correlated measurement, at the j th measurement step, as

$$p_{\mathcal{X}_j}(x_j|\theta, \mathbf{x}_{j-1}) = \text{Tr}(\Pi(x_j|\mathbf{x}_{j-1})\rho_{\Theta}(\theta)), \quad (2.5)$$

where \mathbf{x}_0 should be understood as an empty vector, i.e. no measurement data, and $\Pi(x_j|\mathbf{x}_{j-1})$ denotes the POVM element associated with the j th measurement conditional on the past measurement trajectory. Notice that the quantum state of the sample system is not conditional on the measurement trajectory. We only consider the restricted case where a fresh sample state is prepared between single-shot measurements — this is described in Fig. 2.1 as discarding the conditional sample-state post measurement. Alternatively, this lack of conditioning could model a *non-invasive* measurement which is effectively not disturbing the sample.

Physically, the conditioning of the present measurement on the past measurement trajectory, could mean one of two things. The first possibility is that the state of the measurement apparatus, e.g. a probing system which interacts with the sample, becomes conditioned on the observed measurement outcome, and is reused for a subsequent measurement. The second possibility, which we focus on exclusively, is that the experimenter is allowed to adapt the measurement based on the accumulated data. For instance, we might imagine that the measurement is described by a set of control variables, e.g. interaction strength between the sample and a quantum probe, the measurement time, etc., and the adaptation could consist of tuning these variables. The ability to design the measurement based on accumulated data, makes it possible to optimize the expected performance of the measurement protocol with respect to the parameter estimation task.

2.3 Frequentist estimation

2.3.1 The error-propagation formula

In this section we provide a rudimentary description of the frequentist approach to parameter estimation [6]. We perform a measurement represented by the POVM elements $\Pi(x)$ for $x \in \mathcal{X}$. Based on the acquired measurement data we build an observable $\mathcal{O}(x)$ as measurable function of the observation. In the special case where the implemented measurement corresponds to a projection onto the eigenstates of a quantum observable $\hat{\mathcal{O}}$, the observable function $\mathcal{O}(x)$ would be the eigenvalue associated with an eigenvector labelled by x [1]. The expectation value of the observable function is defined as

$$\langle \mathcal{O} \rangle := \int_{\mathcal{X}} d\mathbb{P}(x|\theta) \mathcal{O}(x), \quad (2.6)$$

and is a function of the parameter to be estimated. For generality, we suppose that we want to estimate a bijective function g of the parameter, i.e. a function with a well-defined inverse. If we assume that the observable expectation can be theoretically computed to high precision, and that the observable expectation is a monotonically increasing function of g , then in principle there exist a function f such that

$$f(\langle \mathcal{O} \rangle) = g(\theta). \quad (2.7)$$

Hence, knowing precisely the expectation of the observable function makes it possible to infer the value of the function g and by extension the parameter value.

Naturally, the observable expectation is not a directly measurable quantity, and it must itself be estimated based on the acquired measurement data. If we perform n identical measurements, i.e. no adaptation, then we can compute the sample mean of the observable function. Denoting the sample mean by $\tilde{\mathcal{O}}(\mathbf{x}_n)$, we can define a frequentist estimate of the parameter

$$f(\tilde{\mathcal{O}}(\mathbf{x}_n)) := g(\tilde{\theta}(\mathbf{x}_n)) := \tilde{g}(\mathbf{x}_n), \quad (2.8)$$

The law of large numbers states that the sample mean converge to the expectation value as n increases [59]. Thus in the asymptotic limit, i.e. for sufficiently large n , we can locally Taylor-expand the estimate around the expectation

$$f(\tilde{\mathcal{O}}(\mathbf{x}_n)) = g(\theta) + \dot{g}(\theta) \left| \frac{d\langle \mathcal{O} \rangle}{d\theta} \right|^{-1} (\tilde{\mathcal{O}}(\mathbf{x}_n) - \langle \mathcal{O} \rangle) + O(\delta_n^2), \quad (2.9)$$

where the dot denotes derivative, and the O notation indicates that we neglect higher-order terms in the deviation $\delta_n := \tilde{\mathcal{O}}(\mathbf{x}_n) - \langle \mathcal{O} \rangle$. In writing down the above Taylor-expansion we take the true parameter value as known, i.e. given a true parameter value, we quantify the deviation between the true g value and the frequentist estimate.

Armed with the above Taylor-expansion, we can provide an asymptotic quantification of the fluctuations of the g estimate around the true function value, via the frequentist mean-squared error defined as

$$\begin{aligned} \Delta^2 \tilde{g}_n(\theta) &:= \int_{\mathbf{x}_n} d\mathbb{P}(\mathbf{x}_n|\theta) (\tilde{g}(\mathbf{x}_n) - g(\theta))^2 \\ &= (\dot{g}(\theta))^2 \left| \frac{d\langle \mathcal{O} \rangle}{d\theta} \right|^{-2} \text{Var}[\tilde{\mathcal{O}}_n] + O(\langle \delta_n^3 \rangle), \end{aligned} \quad (2.10)$$

where $d\mathbb{P}(\mathbf{x}_n|\theta) = \prod_{k=1}^n d\mathbb{P}(x_k|\theta)$, and the second equality follows by substituting the Taylor-expanded estimate, and we have defined the variance of the observable function

$$\text{Var}[\tilde{\mathcal{O}}_n] := \int_{\mathbf{x}_n} d\mathbb{P}(\mathbf{x}_n|\theta) (\tilde{\mathcal{O}}(\mathbf{x}_n) - \langle \mathcal{O} \rangle)^2. \quad (2.11)$$

The second equality in Eq. (2.10) is called the error-propagation formula [60], and quantifies the fluctuations of the computed frequentist estimate around a given true value. These fluctuations arise due to the stochastic nature of the employed measurement.

Naturally, the dependence on a true parameter value in the error-propagation formula is problematic from both a conceptual and an experimental point of view, i.e. before we can quantify the error of an obtained estimate we need to know the true value,

which is exactly what we are attempting to estimate. In practice, this issue can typically be resolved by specifying that the frequentist approach is valid within the local paradigm [6]. If we obtain a frequentist estimate $\tilde{\theta}(\mathbf{x}_n)$, and we consider a parameter interval $\theta \in [\tilde{\theta}(\mathbf{x}_n) - \delta\theta/2, \tilde{\theta}(\mathbf{x}_n) + \delta\theta/2]$ around the estimate. Then if we find that $\Delta^2 \tilde{g}_n(\theta)$ is approximately constant with respect to the parameter across this interval, it follows that $\Delta^2 \tilde{g}_n(\theta)$ provides a meaningful notion of error once $\Delta^2 \tilde{g}_n(\theta) \leq \delta\theta^2$. Thus, within this local paradigm, the error-propagation formula provides us with a tool to quantify the performance of an estimation protocol utilizing a given observable function.

2.3.2 The frequentist Cramér-Rao bound

The error-propagation formula quantifies the error of the frequentist estimate resulting from a specific choice of observable function. We now ask whether a lower bound on this quantity exist, which is given directly in terms of the acquired measurement data, i.e. without any dependence on the choice of observable function. One way of deriving such a bound is to first point out the following equality

$$\frac{d\langle \mathcal{O} \rangle}{d\theta} = \int_{\mathbf{x}_n} d\mathbb{P}(\mathbf{x}_n|\theta) \left(\tilde{\mathcal{O}}(\mathbf{x}_n) - \langle \mathcal{O} \rangle \right) \frac{\partial \log \mathbb{P}(\mathbf{x}_n|\theta)}{\partial \theta}, \quad (2.12)$$

which can be shown straightforwardly by performing the integral on the right-hand side. From the Cauchy-Schwarz inequality [59], we then obtain directly the lower bound

$$\left| \frac{d\langle \mathcal{O} \rangle}{d\theta} \right|^{-2} \text{Var}[\tilde{\mathcal{O}}_n] \geq (n\mathcal{F}_\Pi(\theta))^{-1}, \quad (2.13)$$

where we have defined the Fisher information of the implemented measurement [48]

$$\mathcal{F}_\Pi(\theta) := \int_{\mathcal{X}} d\mathbb{P}(x|\theta) \left(\frac{\partial \log \mathbb{P}(x|\theta)}{\partial \theta} \right)^2, \quad (2.14)$$

and made use of the fact that the Fisher information of independent measurements is additive, we will prove this in Sec. 2.6.2. The Fisher information is a central quantity in estimation theory as it quantifies the amount of information on the parameter which can in principle be extracted from the measurement data [6, 53, 54, 55, 56]. Taking Eq. (2.13) and substituting into the error-propagation formula Eq. (2.10), we obtain the frequentist Cramér-Rao bound [6]

$$\Delta^2 \tilde{g}_n(\theta) \geq \frac{(\dot{g}(\theta))^2}{n\mathcal{F}_\Pi(\theta)} + O(\langle \delta_n^3 \rangle). \quad (2.15)$$

The frequentist Cramér-Rao bound place a lower bound on the fluctuations of our frequentist estimate of the function g at a known true parameter value. Within the frequentist approach the next step is to solve two problems: 1) The so-called classical problem is to find an observable function for which the bound is tight, and 2) the so-called quantum problem is to design the measurement itself to maximize the associated Fisher information [48]. We will not pursue the topic further, as we will primarily concern ourselves with the Bayesian approach to estimation.

2.4 Bayesian inference

2.4.1 Probability measure on the manifold

At this point we have specified the quantity of interest, i.e. the sample state, and described a method of data acquisition. We now discuss how to translate the acquired data into information about the true sample state, by approaching the problem from the perspective of Bayesian inference [6, 53, 54, 55, 61]. From this point of view, the information available about a quantity in a smooth manifold \mathcal{S} , is represented by a probability measure $\mathbb{P}_{\mathcal{S}}$ assigning a probability to a region of the manifold. This probability measure is conditional on the obtained measurement data, i.e. we assume that there exist a rule for mapping available information to a probability measure. If we consider a parameterization $\Theta \subseteq \mathbb{R}$ of the smooth manifold, and indicate the available data as D , then we can express the probability assigned to an infinitesimal interval $[\theta, \theta + d\theta] \subset \Theta$, where $d\theta$ denotes an infinitesimal increment, in terms of a PDF on Θ as

$$\mathbb{P}_{\mathcal{S}}(\theta|D) = d\theta p_{\Theta}(\theta|D). \quad (2.16)$$

The mathematically inclined reader will notice that in decomposing the probability measure as above, we are implicitly assuming that the parameter space Θ can be equipped with a non-negative Riemannian metric, $\dot{\lambda}_{\Theta}$, such that given this metric we can construct an integration measure on the manifold as $d\lambda_{\Theta}(\theta) := d\theta \dot{\lambda}_{\Theta}(\theta)$, with the prime denoting derivative [57, 58, 47]. In the above we have included the factor $\dot{\lambda}_{\Theta}(\theta)$ as part of the definition of the PDF on Θ , and in what follows we will work with this PDF directly.

2.4.2 Bayesian updating

We have argued that the available information should be represented by a PDF conditioned on the acquired data. However, we have said nothing about how to actually map the acquired data into a PDF. The theoretical foundation of Bayesian inference theory, is a prescription for how to update the PDF on Θ as additional measurement data is acquired. This prescription is provided by the Bayes rule. For our purpose we express Bayes rule as [53]

$$p_{\Theta}(\theta|\mathbf{x}_j) = \frac{p_{\mathcal{X}_j}(x_j|\theta, \mathbf{x}_{j-1})p_{\Theta}(\theta|\mathbf{x}_{j-1})}{p_{\mathcal{X}_j}(x_j|\mathbf{x}_{j-1})}, \quad (2.17)$$

which is a single-shot relation between the prior PDF $p_{\Theta}(\theta|\mathbf{x}_{j-1})$ and a posterior PDF $p_{\Theta}(\theta|\mathbf{x}_j)$ reflecting the additional knowledge acquired by observing the measurement outcome $x_j \in \mathcal{X}_j$. The posterior PDF is obtained from the prior PDF via multiplication by the single-shot likelihood function and division by the prior-averaged likelihood function defined by

$$p_{\mathcal{X}_j}(x_j|\mathbf{x}_{j-1}) := \int_{\Theta} d\theta p_{\mathcal{X}_j}(x_j|\theta, \mathbf{x}_{j-1})p_{\Theta}(\theta|\mathbf{x}_{j-1}). \quad (2.18)$$

The prior-averaged likelihood function is also called the evidence function. Note that Bayes' rule as formulated above assumes that the evidence function is positive, or at the very least that the ratio of the likelihood function to the evidence function is finite. For later reference we also define the joint PDF:

$$p_{\mathcal{X}_j \times \Theta}(x_j, \theta|\mathbf{x}_{j-1}) := p_{\mathcal{X}_j}(x_j|\theta, \mathbf{x}_{j-1})p_{\Theta}(\theta|\mathbf{x}_{j-1}). \quad (2.19)$$

Before proceeding we mention two subtleties. The first is fundamental, namely, in outlining the Bayesian updating prescription we are merely postponing the question of how the acquired data is mapped to a posterior PDF. Eventually we must specify an initial prior PDF $p_{\Theta}(\theta|I_0)$ based solely on the available background information, denoted I_0 . The background information represents knowledge of the experiment prior to the acquisition of any measurement data. We will return to the problem of specifying an initial prior towards the end of this chapter. The second subtlety is notational, in that all the PDFs introduced above should be conditional on I_0 . To simplify the notation, we will not indicate this dependence explicitly, except in special cases.

The single-shot Bayes rule as formulated above can be understood as an iterative updating prescription. If we consider a sequence of n measurements, then we can apply the Bayes rule repeatedly, and write the posterior PDF associated with the measurement trajectory \mathbf{x}_n at measurement step n as:

$$p_{\Theta}(\theta|\mathbf{x}_n) = \frac{\prod_{j=1}^n p_{\mathcal{X}_j}(x_j|\theta, \mathbf{x}_{j-1})}{\prod_{j=1}^n p_{\mathcal{X}_j}(x_j|\mathbf{x}_{j-1})} p_{\Theta}(\theta). \quad (2.20)$$

This provides a direct relation between the posterior PDF and the initial prior PDF. If we recall that the full data space is denoted by $\mathcal{X}_n := \mathcal{X}_1 \times \dots \times \mathcal{X}_n$, then it is convenient to define the sequential likelihood and the sequential evidence functions:

$$p_{\mathcal{X}_n}(\mathbf{x}_n|\theta) := \prod_{j=1}^n p_{\mathcal{X}_j}(x_j|\theta, \mathbf{x}_{j-1}), \quad (2.21)$$

$$p_{\mathcal{X}_n}(\mathbf{x}_n) := \prod_{j=1}^n p_{\mathcal{X}_j}(x_j|\mathbf{x}_{j-1}), \quad (2.22)$$

where \mathbf{x}_0 is to be understood as an empty vector, that can be omitted whenever it appears as a conditioning in a PDF. Furthermore, we can define the full joint PDF as

$$p_{\mathcal{X}_n \times \Theta}(\mathbf{x}_n, \theta) := p_{\mathcal{X}_n}(\mathbf{x}_n|\theta) p_{\Theta}(\theta). \quad (2.23)$$

In working with both the joint and the sequential PDFs above, we should keep in mind that they are generally classically correlated, and that care should be taken when manipulating these distributions according to the standard rules of probability theory [53].

2.4.3 Criteria for parameterization invariance

The text and figures of this section is taken from Ref. [52].

Having introduced the manifold of quantum states, and discussed PDFs on the parameter space, we must pause and face the important subtlety of parameterization invariance. In the above we work with a parameterization Θ , however the manifold of quantum states itself is invariant with respect to the specific choice of parameterization. For example, the manifold of thermal states is the same whether we parameterize it using the temperature or the inverse temperature — this fact is illustrated in Fig. 2.2. In general we can express this invariance as follows: if we consider a one-to-one mapping $\phi: \Theta \rightarrow \Phi$, where $\Phi \subseteq \mathbb{R}$ is the image of the map, then the function ϕ provides an equally valid parameterization of the manifold of states, i.e. $\mathcal{S}_{\Phi} = \mathcal{S}_{\Theta}$ with

$$\mathcal{S}_{\Phi} := \{\rho_{\Phi}(\phi) \text{ for } \phi \in \Phi\}, \quad (2.24)$$

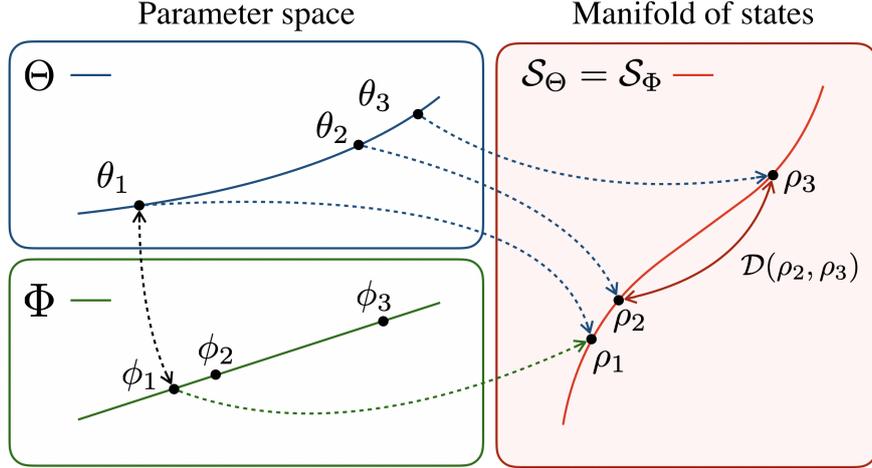


Figure 2.2: *This figure and caption is adopted from Ref. [52].* Illustration of a parameter space Θ , or an alternative parameterization Φ , mapped into a one-parameter curve of quantum states $\mathcal{S}_\Theta = \mathcal{S}_\Phi$. The distinguishability of parameter values is only meaningful, when expressed as a distance \mathcal{D} between the corresponding quantum states, i.e., the state space induces a geometry on the parameter space. The induced geometry may be simpler for certain choices of parameterization. In particular, there may exist a Euclidean parameterization with a flat geometry. By considering the distance directly between quantum states, it follows that the distance is invariant with respect to the specific choice of parameterization of the states.

where we explicitly indicate that each state is given with respect to the parameterization Φ , while the *invariance condition* reads

$$\forall \phi : \phi = \phi(\theta), \quad \rho_\Theta(\theta) = \rho_\Phi(\phi), \quad (2.25)$$

and expresses the parameterization invariance of every equivalent quantum state. Furthermore, the probability assigned to a given region of state space must be independent of the specific parameterization employed. Thus, to be consistent, the PDF on the parameter space must satisfy the invariance condition [47, 61]

$$\forall \phi : \phi = \phi(\theta), \quad d\theta p_\Theta(\theta|D) = d\phi p_\Phi(\phi|D). \quad (2.26)$$

Since the likelihood function only depends on the sample state itself it is inherently parameterization invariant, and it thus follows that if the above invariance condition holds for the initial prior PDF $p_\Phi(\phi|I_0)$, then it will also hold for the sequence of posterior PDFs. To simplify our notation we will in the following chapters not indicate the parameterization explicitly in all cases. The choice of parameterization is then implicitly indicated by the parameter value itself.

2.5 Invariant estimation theory

2.5.1 Criteria for invariant point estimation

At this point we have described how, within the Bayesian approach to parameter estimation, the available information is translated into a posterior PDF $p_\Theta(\theta|D)$ on a parameter

space Θ . This posterior PDF represents our degree of belief in different parameter values being the true parameter value, given the acquired measurement data. Furthermore, we have pointed out the required parameterization-invariance of the estimation task, i.e. the available information could equivalently be translated into a posterior PDF $p_{\Phi}(\phi|D)$ on a parameter space Φ , with a one-to-one mapping $\phi : \Theta \rightarrow \phi(\theta) \in \Phi$ connecting these two parameter spaces. Given these tools, we are ready to discuss how to go about developing a well-defined, i.e. an invariant, estimation theory.

A classical estimation theory, at least when considering point estimation, is (i) a prescription for singling out a parameter estimate $\theta_{\text{est}} \in \Theta$, serving as a best guess of the true parameter value, and (ii) aims to provide a measure of confidence in the correctness of the identified point estimate, we will refer to this as the credibility [6, 53, 54, 55, 56]. In light of the above discussion, we must demand that a well-defined estimation theory is parameterization invariant. In particular, we impose two requirements. First, that the prescription for specifying an estimate should be consistent. The identified estimates should correspond to the same quantum state irrespective of the choice of parameterization. Mathematically, we demand that $\phi_{\text{est}} = \phi(\theta_{\text{est}})$. Second, that the credibility is a single unambiguous parameterization-invariant quantity.

2.5.2 Singling out a parameter estimate

First we consider the task of singling out a parameter estimate. As above, we consider a general estimation protocol consisting of n single-shot measurements. The full experiment generates a trajectory of n measurement outcomes, which are sampled from the ordered sequence of likelihood functions, that is

$$X_k \sim p_{\mathcal{X}_k}(x_k|\theta_{\text{true}}, \mathbf{x}_{k-1}), \quad \text{for } k = 1, \dots, n, \quad (2.27)$$

where the notation \sim indicates that the stochastic variable on the left is sampled according to the distribution on the right. Here, the likelihood functions are specified with respect to a true parameter value $\theta_{\text{true}} \in \Theta$, which is unknown to the experimenter. Note that since the likelihood function is inherently parameterization-invariant, we could equivalently specify the sampling with respect to the parameterization Φ , with the constraint that $\varphi_{\text{true}} = \varphi(\theta_{\text{true}})$.

In general, an estimate is computed as a function of the available measurement data. Any arbitrary estimator can be expressed as a mapping from the full data space to the parameter space:

$$\theta_{\text{est}} : \mathcal{X}_n \rightarrow \theta_{\text{est}}(\mathbf{x}_n) \in \Theta. \quad (2.28)$$

A common choice of estimator is the maximum-a-posterior (MAP) estimate [6, 47]. This is conventionally defined as the parameter value for which the posterior PDF, resulting from updating a prior PDF based on the acquired data, takes its maximum value. However, in adopting the MAP estimator we immediately run into conflict with the requirement of parameterization invariance. To see this, we recall from the invariance condition Eq. (2.26) that PDF on the parameter space transforms as

$$p_{\Theta}(\theta|\mathbf{x}_n) = \partial_{\theta}\phi p_{\Phi}(\phi|\mathbf{x}_n), \quad \text{for } \phi = \phi(\theta), \quad (2.29)$$

under a change of parameterization $\Theta \rightarrow \Phi$. The presence of the Jacobian factor $\partial_{\theta}\phi$, means that the MAP estimates computed by maximising the PDF on Θ and on Φ , might

be different depending on the specific parameterization employed. That is, we could find $\phi_{\text{MAP}} \neq \phi(\theta_{\text{MAP}})$, which breaks the invariance condition. The implication of this is that the MAP estimate of the true sample state depends on the choice of parameterization, or in other words, it depends on the whims of the person making the estimate. A situation that is untenable.

One might reasonably ask if this problem only arises because we consider the MAP estimator, and whether alternative well-defined estimators exist. Indeed the problem can be circumvented if we instead consider a maximum-likelihood (ML) estimator [6], in which we search for the parameter value that maximizes the sequential likelihood function $p_{\mathbf{x}_n}(\mathbf{x}_n|\theta)$ given the observed data. The fact that the ML estimate is parameterization-invariant follows directly from the parameterization-invariance of the likelihood function. The ML estimator is a common choice within the frequentist approach to estimation theory [6], and it has the feature, for better or worse, that it is independent of the choice of initial prior PDF.

2.5.3 Distance function and the mean-squared error

Although the use of the ML estimator is perfectly acceptable, we shall return to the problem of obtaining parameterization-invariant MAP estimates. First, however, we consider the second part of an estimation theory – the construction of a measure of credibility. The following analysis will allow us to correct the problems with the MAP estimator. If we consider the Θ parameterization, then a common credibility measure is the posterior mean-squared error defined as [47]

$$\text{MSE}[\theta_{\text{est}}](\mathbf{x}_n) := \int_{\Theta} d\theta p_{\Theta}(\theta|\mathbf{x}_n) \mathcal{D}_{\Theta}(\theta_{\text{est}}(\mathbf{x}_n), \theta)^2, \quad (2.30)$$

where we have introduced a distance function $\mathcal{D}_{\Theta}(\theta_{\text{est}}, \theta)$ gauging the distance, or deviation, between an adopted estimate $\theta_{\text{est}} \in \Theta$ and a parameter value $\theta \in \Theta$. A suitable distance function must be non-negative, symmetric in its arguments, it must vanish at $\theta_{\text{est}} = \theta$, it must increase monotonically away from this point and it must satisfy the triangle inequality [47]. Note that the MSE is defined with respect to the choice of estimator, which we indicate using square brackets, and that it is a function of the obtained measurement trajectory, i.e. it is a stochastic quantity.

In order for the adopted confidence measure to be well defined, and thus provide a single unambiguous credibility measure, it must satisfy the requirement of parameterization invariance. We can state this requirement as

$$\forall \phi : \phi = \phi(\theta), \quad \text{MSE}[\theta_{\text{est}}](\mathbf{x}_n) = \text{MSE}[\phi_{\text{est}}](\mathbf{x}_n), \quad (2.31)$$

where we use the convention that the choice of parameterization is indicated by the notation used for the estimator. From Eq. (2.26) we recall the invariance condition

$$\forall \phi : \phi = \phi(\theta), \quad d\theta p_{\Theta}(\theta|\mathbf{x}_n) = d\phi(\theta) p_{\Phi}(\phi(\theta)|\mathbf{x}_n), \quad (2.32)$$

ensuring that the MSE is defined with respect to a parameterization-invariant probability measure. The invariance condition makes it possible to express the requirement for Eq. (2.31) to hold, directly in terms of the distance function as

$$\forall \phi : \phi = \phi(\theta), \quad \mathcal{D}_{\Theta}(\theta_{\text{est}}, \theta) = \mathcal{D}_{\Phi}(\phi_{\text{est}}, \phi). \quad (2.33)$$

Hence the MSE is a parameterization-invariant confidence measure, whenever the distance function is a parameterization-invariant quantity, satisfying Eq. (2.33). Such a distance function can be understood as a distance directly between quantum states in the manifold. If we were to follow the conventional route to parameter estimation [6, 54, 55], we would adopt a distance function which is simply the absolute difference of the parameter values, that is

$$\mathcal{D}_\Theta(\theta_{\text{est}}, \theta) = |\theta_{\text{est}} - \theta|. \quad (2.34)$$

This choice satisfies all the requirements placed on a well-defined distance function. However, in general $|\theta_{\text{est}} - \theta| \neq |\phi_{\text{est}} - \phi(\theta)|$ and it is clear to see that the resulting MSEs cannot be equivalent. The question we must address, before we can compute an invariant credibility of a given estimate, is thus how to construct a parameterization-invariant distance function.

2.5.4 Metric structure and the geodesic distance

*This section, and its subsections, contains text and figures from Ref. [52].
Minor edits have been made to the text to conform to notation.*

The essential ingredient required for the development of a parameterization-invariant estimation theory on a smooth manifold is the introduction of a metric structure, i.e. an infinitesimal notion of length, on the manifold of quantum states [47, 57, 58]. A metric makes it possible to define the distance between states on the manifold, and more fundamentally, the idea of defining a PDF on a continuous parameter space, is not well-defined in the absence of such a metric structure [47]. Suppose for a moment that the manifold possesses a metric, denoted $\dot{\lambda}$. Then, it is always possible to construct a well-defined distance function, as the geodesic length – which we refer to as the metric-based distance – defined by [47, 57, 49]

$$\mathcal{D}_\Theta(\theta_0, \theta_1) := \left| \int_{\theta_0}^{\theta_1} d\theta \dot{\lambda}_\Theta(\theta) \right|, \quad (2.35)$$

where the integral is over the parameter domain $\theta \in [\theta_0, \theta_1] \subseteq \Theta$, and the quantity $d\lambda_\Theta(\theta) = d\theta \dot{\lambda}_\Theta(\theta)$ provides a parameterization-invariant integration measure, i.e. if $\phi = \phi(\theta)$ then it follows that $d\lambda_\Theta(\theta) = d\lambda_\Phi(\phi)$. The metric-based distance is a distance between states, and is thus a parameterization-invariant quantity — this is depicted in Fig. 2.2. The form of the geodesic length above is valid for one-parameter problems, more generally, defining the geodesic length involves a minimization over paths [47].

The above way of defining a distance function on the parameter space may still seem ambiguous unless a particular choice of metric can be justified. However, if we consider a measurement of the sample with POVM elements $\mathcal{M}(y)$ for $y \in \mathcal{Y}$, then a key insight of Bayesian information geometry [47, 57, 49] is that the likelihood associated with this measurement, induces a metric of the form $\dot{\lambda}_\Theta = \mathcal{F}_{\mathcal{M}, \Theta}^{1/2}$, where $\mathcal{F}_{\mathcal{M}, \Theta}$ is the Fisher information (FI) associated with the \mathcal{M} measurement [47, 57, 62]

$$\mathcal{F}_{\mathcal{M}, \Theta}(\theta) := \int_{\mathcal{Y}} d\mathbb{P}_0(y) p_{\mathcal{Y}}(y|\theta) (\partial_\theta \log p_{\mathcal{Y}}(y|\theta))^2. \quad (2.36)$$

Throughout, we refer to the \mathcal{M} measurement as the reference measurement. From the point of view of Bayesian probability theory, the FI metric is the unique Riemannian

metric, up to a choice of reference measurement, which is both parameterization invariant and which satisfies any other invariance property of the likelihood function [47, 63]. According to Chentsov's theorem, any other monotonic metric on the parameter space corresponds to a Fisher information metric up to a multiplicative constant [63, 64].

Quantum Fisher information metric

The choice of reference measurement represents a degree of freedom, i.e., we can define the metric-based distance, and the estimation theory itself, relative to an arbitrary reference measurement of the sample. A natural choice would be to maximize the distance $\mathcal{D}_\Theta(\theta_0, \theta_1)$ over all possible measurements, however this procedure does not generally yield a unique reference measurement. To see this, we consider the problem of maximizing the metric at a specific parameter value. This local maximization problem can be solved analytically, and yields a metric defined relative to a projective measurement of the symmetric logarithmic derivative $\mathcal{L}_\Theta(\theta)$ associated with the manifold of states¹ [48]. The symmetric logarithmic derivative (SLD) is defined implicitly by the relation [48]

$$\mathcal{L}_\Theta(\theta)\rho_\Theta(\theta) + \rho_\Theta(\theta)\mathcal{L}_\Theta(\theta) = 2\partial_\theta\rho_\Theta(\theta), \quad (2.37)$$

and is in general a function of the parameter value. The problem alluded to above is then that the SLD defines a natural reference measurement only if the eigenbasis of the SLD is parameter independent, i.e., the projectors must be parameter independent. Notice that this need not be the case for the SLD eigenvalues.

If we adopt the SLD reference measurement, then with this choice of reference we obtain the quantum Fisher information (QFI) metric, which gives the maximum likelihood-induced metric-based distance between sample states. The QFI metric is given by [48]

$$\sqrt{\mathcal{F}_{\mathcal{L},\Theta}(\theta)} = \text{Tr}[\mathcal{L}_\Theta(\theta)^2\rho_\Theta(\theta)]^{1/2}. \quad (2.38)$$

The resulting metric is equal to four times the Bures metric [62], and is thus directly related to the so-called fidelity between infinitesimally separated states in the manifold. If the family of states considered is a thermal state ensemble, then the metric-based distance is also called the thermodynamic length [49, 65, 66, 67].

Likelihood-induced Euclidean parameterization

Having equipped the manifold with a metric structure, and constructed a distance function, we return to the MSD, now defined with respect to an arbitrary reference measurement \mathcal{M} . For later convenience we define, implicitly, the function $\lambda_{\mathcal{M}}(\theta)$ as the inverse derivative of the FI metric associated with any measurement \mathcal{M} as

$$\partial_\theta\lambda_{\mathcal{M},\Theta}(\theta) := \mathcal{F}_{\mathcal{M},\Theta}^{1/2}(\theta). \quad (2.39)$$

Since the FI is non-negative it follows that the function $\lambda_{\mathcal{M},\Theta}$ is monotonically increasing. Furthermore, if we consider a change of parameterization $\Theta \rightarrow \Phi$, and note that under such a transformation the FI transforms as $\mathcal{F}_{\mathcal{M},\Theta}(\theta) = [\partial_\theta\phi]^2\mathcal{F}_{\mathcal{M},\Phi}(\phi)$, then we see that it follows directly from the definition of $\lambda_{\mathcal{M},\Theta}$ that it is a parameterization-invariant quantity, i.e., $\lambda_{\mathcal{M},\Theta}(\theta) = \lambda_{\mathcal{M},\Phi}(\phi)$ for $\phi = \phi(\theta)$. Given the $\lambda_{\mathcal{M},\Theta}$ function it is

¹We will provide a proof of this in Sec. 2.6.1.

always possible to express the distance Eq. (2.35) defined based on the FI metric in the Euclidean form, i.e.

$$\mathcal{D}_{\mathcal{M},\Theta}(\theta_{\text{est}}, \theta) = |\lambda_{\mathcal{M},\Theta}(\theta_{\text{est}}) - \lambda_{\mathcal{M},\Theta}(\theta)|, \quad (2.40)$$

where we indicate that the distance is defined with respect to the reference measurement \mathcal{M} . The Euclidean form of the distance follows directly from an application of the fundamental theorem of calculus for $\dot{\lambda}_{\mathcal{M},\Theta}(\theta) = \mathcal{F}_{\mathcal{M},\Theta}^{1/2}(\theta)$.

If we constrain ourselves to reference measurements for which the FI is non-vanishing (except perhaps at isolated points, e.g., the boundaries of the parameter domain), then $\lambda_{\mathcal{M},\Theta}$ itself constitutes a valid parameterization of the one-parameter family within the manifold. Referring to the associated parameter space as $\Lambda_{\mathcal{M}}$, it follows that when working in this parameterization the MSE takes the simple form

$$\text{MSE}[\lambda_{\text{est}}](\mathbf{x}_n) = \int_{\Lambda_{\mathcal{M}}} d\lambda p_{\Lambda_{\mathcal{M}}}(\lambda|\mathbf{x}_n) (\lambda_{\text{est}}(\mathbf{x}_n) - \lambda)^2, \quad (2.41)$$

where $\lambda_{\text{est}} = \lambda(\theta_{\text{est}})$, we have made use of Eq. (2.26), i.e., the invariance property of the posterior PDF. In the above we do not indicate explicitly that the MSE is defined with respect to a reference measurement. Furthermore we drop the subscript \mathcal{M} when referring to the λ parameter itself. Note that the $\Lambda_{\mathcal{M}}$ -parameterization is special in that it is associated with a FI equal to unity, i.e. a flat metric. The form of the MSE is then simply the conventional Euclidean error with respect to the parameterization $\Lambda_{\mathcal{M}}$.

2.5.5 Invariant MAP estimates

Having established a metric structure, we now return to the problem of defining MAP estimates satisfying the requirement of parameterization invariance. The treatment follows that of Ref. [47]. The problem with our previous attempt was that we neglected the fact that a PDF is always defined with respect to an underlying reference measure. We denote this by $\mathbb{U}_{\mathcal{M}}$ in anticipation of the fact that the reference measure is defined with respect to a reference measurement. The role of the reference measure is to assign an invariant measure of length to a given parameter interval. If we consider an arbitrary probability measure $\mathbb{P}_{\mathcal{S}}$ on the manifold of states \mathcal{S} , we can define the invariant PDF, denoted $\tilde{p}_{\mathcal{S},\mathcal{M}}$, with respect to $\mathbb{U}_{\mathcal{M}}$ via the Radon-Nikodym derivative [47, 35]. If we focus on the coordinate system Θ , then this takes the form

$$\tilde{p}_{\mathcal{S},\mathcal{M}}(\theta|D) := \frac{d\mathbb{P}_{\mathcal{S}}(\theta|D)}{d\mathbb{U}_{\mathcal{M}}(\theta)}, \quad (2.42)$$

where we use D to indicate available information, e.g. $D = \{\mathbf{x}_n, I_0\}$. In light of the discussion of the previous sections, we know that a suitable reference measure can be written in terms of the Fisher information metric:

$$d\mathbb{U}_{\mathcal{M}}(\theta) = d\theta \mathcal{F}_{\mathcal{M},\Theta}^{1/2}(\theta). \quad (2.43)$$

Note that, even though we express the reference measure in the Θ parameterization, it is a parameterization-invariant quantity. As mentioned above, the reference measure assigns an invariant measure of length to a parameter interval. If we recall that the FI metric provides a measure of the distinguishability of infinitesimally separated states,

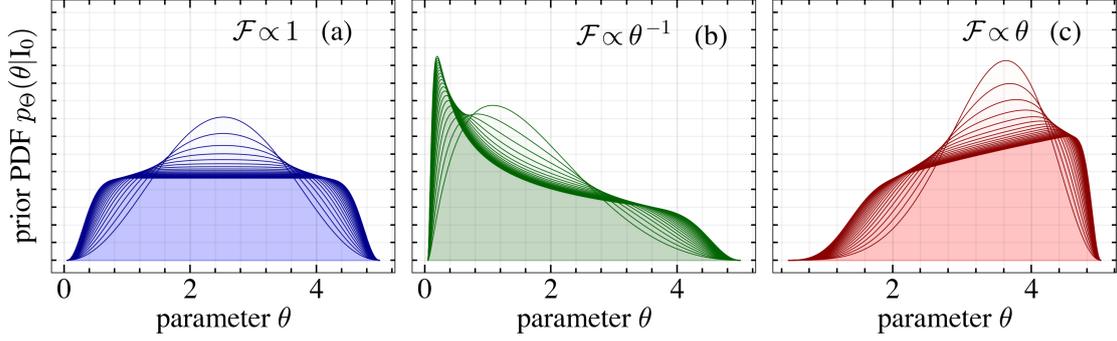


Figure 2.3: Illustration of the smoothed Jeffrey's prior (Eqs. (2.48) and (2.49)) for $\theta \in [0, 5]$ and $\alpha \in [-20, 0]$. (a) Plot for $\mathcal{F}_{\mathcal{M}}(\theta) \propto 1$, as α decreases we approach a flat prior distribution across the parameter domain. (b) Plot for $\mathcal{F}_{\mathcal{M}}(\theta) \propto 1/\sqrt{\theta}$, as α decreases we approach a $1/\sqrt{\theta}$ prior distribution across the parameter domain. (c) Plot for $\mathcal{F}_{\mathcal{M}}(\theta) \propto \theta$, as α decreases we approach a $\sqrt{\theta}$ prior distribution across the parameter domain. For a curved metric the prior density does not approach a constant across the parameter domain because the metric length is not uniform.

then we see that $\mathbb{U}_{\mathcal{M}}$ assigns length based on the ability of the reference measurement to resolve, or distinguish, the underlying quantum states.

Now, recall that the probability measure $\mathbb{P}_{\mathcal{S}}$ itself can also be written in terms of a PDF in the Θ parameterization, namely $\mathbb{P}_{\mathcal{S}}(\theta|D) = d\theta p_{\Theta}(\theta|D)$. Upon substitution we obtain an explicit form of the invariant PDF given by

$$\tilde{p}_{\mathcal{S},\mathcal{M}}(\theta|D) = \frac{p_{\Theta}(\theta|D)}{\mathcal{F}_{\mathcal{M},\Theta}^{1/2}(\theta)}. \quad (2.44)$$

Armed with the invariant PDF we can straightforwardly define an invariant MAP estimate. We simply have to define the invariant MAP estimate in terms of the invariant posterior PDF [47]. This can be expressed in a simple form if we consider the change of parameterization $\Theta \rightarrow \Lambda_{\mathcal{M}}$. Since the FI metric equals one in the $\Lambda_{\mathcal{M}}$ parameterization, we find that

$$\tilde{p}_{\mathcal{S},\mathcal{M}}(\theta|D) = p_{\Lambda_{\mathcal{M}}}(\lambda(\theta)|D), \quad (2.45)$$

and we can thus express the MAP estimate as

$$\lambda_{\text{MAP}}(D) := \arg \max_{\lambda \in \Lambda_{\mathcal{M}}} p_{\Lambda_{\mathcal{M}}}(\lambda|D). \quad (2.46)$$

If we want the MAP estimate in the Θ parameterization, we apply the inverse mapping $\lambda_{\Theta}^{-1} : \Lambda_{\mathcal{M}} \rightarrow \Theta$, and obtain $\theta_{\text{MAP}}(D) = \lambda_{\Theta}^{-1}(\lambda_{\text{MAP}}(D))$. This equivalence can also be shown by direct computation of the MAP estimate in the Θ parameterization. The above expression for the MAP estimator is one example of the utility of working within the Euclidean parameterization $\Lambda_{\mathcal{M}}$. In the next chapter we will utilize this choice of parameterization to develop a number of key results from Euclidean estimation theory [6, 53, 54, 55, 56].

2.5.6 Selecting an initial prior probability

The application of the Bayesian inference framework relies on the ability to specify an initial prior PDF on a given parameter space. As mentioned in the Sec. 2.4 on Bayesian inference we suppose that an experimenter has certain background information, denoted I_0 . Before the Bayesian inference framework can be applied, there should be a prescription for how background information is mapped to a prior PDF.

Following E. T. Jaynes [53], the first step is to establish the form of a distribution reflecting complete ignorance. Intuitively, given no information to the contrary, we should assign equal probability to equal lengths on our manifold. In our case, the invariant measure of length is given by the measure $\mathbb{U}_{\mathcal{M}}$. Hence, a probability measure reflecting complete ignorance must take the form

$$\mathbb{P}_{\mathcal{S}}(\theta) \propto \mathbb{U}_{\mathcal{M}}(\theta). \quad (2.47)$$

This form of the prior probability is essentially equivalent to the so-called Jeffrey's prior [53, 68, 61]. Jeffrey's prior does not generally hold across the entire parameter space, and in most cases must be supplemented with an envelope function to ensure that the resulting measure is integrable. For the majority of the simulations performed in this thesis, we assume that sufficient background information exists to identify a bounded parameter interval $\theta \in [\theta_{\min}, \theta_{\max}]$. On this interval we adopt a smoothed version of Jeffrey's prior:

$$p_{\Theta}(\theta) = \sqrt{\mathcal{F}_{\mathcal{M},\Theta}(\theta)} \xi_{\Theta}(\theta), \quad (2.48)$$

where the function $\xi_{\Theta}(\theta)$ is an invariant envelope function, satisfying the condition $\xi_{\Theta}(\theta) = \xi_{\Phi}(\phi)$ under a change of parameterization $\Theta \rightarrow \Phi$ with $\phi = \phi(\theta)$. Inspired by the envelope employed by Yan et al. in Ref. [69] we adopt the envelope function:

$$\xi_{\Theta}(\theta) = \frac{1}{\mathcal{N}} \left(\exp \left[\alpha \sin^2 \left(\pi \frac{\lambda(\theta) - \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}} \right) \right] - 1 \right) \quad (2.49)$$

with a normalization factor given by

$$\mathcal{N} := (\lambda_{\max} - \lambda_{\min}) [\exp(\alpha/2) J_0(\alpha/2) - 1], \quad (2.50)$$

where J_0 denotes the modified Bessel function of the first kind, and $\lambda_{\min, \max} := \lambda(\theta_{\min, \max})$. The difference between the above envelope and the one employed by Yan et al. [69], is that we specify the function with respect to the Euclidean parameterization of the manifold of states. In Fig. 2.3 we illustrate the smoothed Jeffrey's prior for three different FI metrics. We observe that in the limit of $\alpha \rightarrow -\infty$, we approximate Jeffrey's prior across the full parameter domain. Finally, we note that the topic of finding suitable prior distributions based on available background information extends far beyond the rudimentary results discussed here. For an in-depth treatment see Ref. [53].

2.6 Properties of the Fisher information

The Fisher information associated with a given measurement is a key quantity in the theory of estimation. It serves to equip a manifold with a local notion of distance based on how well states on the manifold can be distinguished from the observed measurement statistics [62]. In this section, we outline some technical properties of the Fisher information.

2.6.1 Derivation of the quantum Fisher information

We first derive the form of the quantum Fisher information (QFI) following the derivation of Ref. [48]. Starting out from the definition of the Fisher information (Eq. (2.36)):

$$\begin{aligned}
\mathcal{F}_{\mathcal{M}}(\theta) &= \int d\mathbb{P}_0(y) \frac{(\partial_\theta \text{Tr}(\rho(\theta)\mathcal{M}(y)))^2}{\text{Tr}(\rho(\theta)\mathcal{M}(y))} \\
&= \int d\mathbb{P}_0(y) \frac{(\text{Re}[\text{Tr}(\mathcal{L}(\theta)\rho(\theta)\mathcal{M}(y))])^2}{\text{Tr}(\rho(\theta)\mathcal{M}(y))} \\
&\leq \int d\mathbb{P}_0(y) \frac{|\text{Tr}(\mathcal{L}(\theta)\rho(\theta)\mathcal{M}(y))|^2}{\text{Tr}(\rho(\theta)\mathcal{M}(y))},
\end{aligned} \tag{2.51}$$

where the second equality follows from the definition of the SLD, and the third inequality express the fact that the absolute value of a complex number is greater than the real part of the number. If we now use the fact that quantum states and POVMs are positive semi-definite operators, we can write

$$\rho(\theta)\mathcal{M}(y) = \sqrt{\rho(\theta)}\sqrt{\rho(\theta)}\sqrt{\mathcal{M}(y)}\sqrt{\mathcal{M}(y)}. \tag{2.52}$$

Substituting this above, making use of the cyclic property of the trace, and applying the Cauchy-Schwarz inequality [1], we obtain the inequality:

$$\begin{aligned}
\mathcal{F}_{\mathcal{M}}(\theta) &\leq \int d\mathbb{P}_0(y) \frac{\text{Tr}(\rho(\theta)\mathcal{M}(y)) \text{Tr}(\mathcal{L}(\theta)\rho(\theta)\mathcal{L}(\theta)\mathcal{M}(y))}{\text{Tr}(\rho(\theta)\mathcal{M}(y))} \\
&= \text{Tr}\left(\mathcal{L}(\theta)\rho(\theta)\mathcal{L}(\theta) \int d\mathbb{P}_0(y)\mathcal{M}(y)\right) \\
&= \text{Tr}(\mathcal{L}(\theta)\rho(\theta)\mathcal{L}(\theta)) \\
&:= \mathcal{F}_{\mathcal{L}}(\theta),
\end{aligned} \tag{2.53}$$

where we employed the normalization requirement on the POVM, and the last equality defined the quantum Fisher information. As shown by Braunstein and Caves [48] the condition for saturating the bound in Eq. (2.53), comes down to choosing the POVM elements $\mathcal{M}(y)$ to be projectors onto the eigenstates of the SLD.

2.6.2 Additivity of independent measurements

An important property of the Fisher information is additivity. Given ν independent measurements, denoted $\mathcal{M}_\nu := \{\mathcal{M}_1, \dots, \mathcal{M}_\nu\}$ with

$$\mathcal{M}_k := \{\mathcal{M}_k(y) \text{ for } y_k \in \mathcal{Y}_k\}, \tag{2.54}$$

the associated total Fisher information is given by the sum of the individual Fisher informations. Defining $d\mathbb{P}_0(\mathbf{y}_\nu) = \prod_{k=1}^\nu d\mathbb{P}_0(y_k)$, this can be shown as follows:

$$\begin{aligned}
\mathcal{F}_{\mathcal{M}_\nu, \Theta}(\theta) &= \int_{\mathbf{y}_\nu} d\mathbb{P}_0(\mathbf{y}_\nu) p_{\mathbf{y}_\nu}(\mathbf{y}_\nu|\theta) [\partial_\theta \log p_{\mathbf{y}_\nu}(\mathbf{y}_\nu|\theta)]^2 \\
&= \int_{\mathbf{y}_\nu} d\mathbb{P}_0(\mathbf{y}_\nu) p_{\mathbf{y}_\nu}(\mathbf{y}_\nu|\theta) \left[\partial_\theta \sum_{k=1}^\nu \log p_{\mathcal{Y}_k}(y_k|\theta) \right]^2 \\
&= \sum_{k=1}^\nu \mathcal{F}_{\mathcal{M}_k, \Theta}(\theta) + \sum_{k,l=1}^\nu \int_{\mathcal{Y}_k \times \mathcal{Y}_l} d\mathbb{P}_0(y_k) d\mathbb{P}_0(y_l) \partial_\theta p_{\mathcal{Y}_k}(y_k|\theta) \partial_\theta p_{\mathcal{Y}_l}(y_l|\theta) \\
&= \sum_{k=1}^\nu \mathcal{F}_{\mathcal{M}_k, \Theta}(\theta),
\end{aligned} \tag{2.55}$$

where we have used that the total likelihood function is a product of the individual likelihood functions, and the final equality follows due to vanishing cross terms under the integral. Note that this property holds when the measurements are independent. In the next chapter we will discuss a generalization of the additivity property that holds for correlated measurements. Another important property of the Fisher information is convexity. However, we will not introduce this notion until it is needed in a later chapter.

2.7 Concluding remarks

In this chapter we have formulated the problem of identifying the true quantum state of a physical system, in the case that the state belongs to a smooth manifold of states, as a problem of statistical inference. We argued that in order for it to be possible to meaningfully talk about the credibility of a given estimate, the manifold must be equipped with a metric structure. Such a metric structure is in general defined relative to a reference measurement of the system, and it induce a distance measure between states in the manifold, based on how well the reference measurement can distinguish the system states.

For each reference measurement, it is possible to associate a Euclidean parameterization of the manifold. In the next chapter we will focus on this Euclidean parameterization. In particular, we will be able to apply standard results on the fundamental limits to estimation precision, which have been developed in the case of Euclidean parameter spaces.

Chapter 3

Estimation Theory: Precision Bounds and Adaptive Strategies

3.1 Introduction

In this chapter we consider a Euclidean parameter space $\Lambda \subseteq \mathbb{R}$, in which the parameter space is equipped with a flat metric structure. Given the Euclidean parameterization we can develop some key results of standard estimation theory [6, 53, 54, 55, 56]. In particular we can investigate lower bounds on the attainable credibility of an estimate, and construct a methodology for adaptive measurement design. In light of the discussion of the preceding section, we know that such a Euclidean theory is always defined relative to a reference measurement, e.g. a projective measurement of the symmetric logarithmic derivative. However, this fact will not be indicated explicitly in this chapter.

The structure of the chapter is as follows: In Sec. 3.2 we recap the introduction a posteriori mean-squared error in the previous chapter. In Sec. 3.3 we discuss lower bounds, in particular we focus on the posterior Cramér-Rao bound and the conditions under which it is tight. In Sec. 3.4 we describe the averaged error and averaged bounds. In Sec. 3.5 we make use of the averaged quantities to build strategies for adaptive measurement design. The majority of the material presented here is well-known [6, 53, 54, 55, 56, 69, 70]. The one minor exception is Sec. 3.5 on adaptive measurement design. Although similar results have been reported in the literature [71, 72], the specific analysis presented here, and the idea of measurement design based on the van Trees bound, seems novel.

3.2 Posterior mean-squared error

We first recap the basic idea of classical parameter estimation theory, and consider a general estimation protocol consisting of n single-shot measurements. The full experiment gives a trajectory \mathbf{x}_n of n measurement outcomes, which are generated via stochastic sampling from the ordered sequence of likelihood functions

$$X_k \sim p_{\mathcal{X}_k}(x_k | \lambda_{\text{true}}, \mathbf{x}_{k-1}), \quad \text{for } k = 1, \dots, n, \quad (3.1)$$

with respect to a true parameter value $\lambda_{\text{true}} \in \Lambda$. The true parameter value is unknown to the experimenter, and is only accessible via the observed measurement record. A

classical estimation theory, at least when considering point estimation, is (i) a prescription for singling out a parameter estimate $\lambda_{\text{est}} \in \Lambda$, serving as a best guess of the true parameter value, and (ii) aims to provide a measure of the credibility of the identified point estimate. An estimate is computed as a function of the available measurement data. Any arbitrary estimator can be expressed as a mapping from the full data space to the parameter space

$$\lambda_{\text{est}} : \mathcal{X}_n \rightarrow \lambda_{\text{est}}(\mathbf{x}_n) \in \Lambda, \quad (3.2)$$

where we recall that the full data space is denoted by $\mathcal{X}_n := \mathcal{X}_1 \times \dots \times \mathcal{X}_n$. An estimator can in principle be any arbitrary prescription preferred by the experimenter. However, a minimal requirement of asymptotic consistency is typically adopted, demanding that in the limit of infinite measurement data the estimator must converge to the true parameter value [73, 74]. A common choice of asymptotically consistent estimator, which is suitable given that Λ is a Euclidean space, is the MAP estimate, defined as the parameter value for which the posterior PDF, resulting from updating a prior PDF based on the acquired data, takes its maximum value.

Given an estimator, the next problem is to introduce a measure of confidence in the correctness of the identified point estimate. When the parameter space Λ is a Euclidean space, it is common to quantify the credibility of an estimate by the posterior mean-squared error (MSE) defined as [6]:

$$\text{MSE}[\lambda_{\text{est}}](\mathbf{x}_n) := \int_{\Lambda} d\lambda p_{\Lambda}(\lambda|\mathbf{x}_n) (\lambda_{\text{est}}(\mathbf{x}_n) - \lambda)^2, \quad (3.3)$$

where we recall that $p_{\Lambda}(\lambda|\mathbf{x}_n)$ denotes the posterior PDF on the parameter space Λ . Note that the MSE is a function of the acquired measurement data, i.e. it is a stochastic quantity. Furthermore, it is defined with respect to a specific choice of estimation function. We indicate this using square brackets. Lastly, the MSE depends on the choice of the initial prior PDF $p_{\Lambda}(\lambda)$. We will never indicate this dependence explicitly.

3.2.1 Minimal mean-squared error estimator

Although alternatives exist, the MSE is a typical choice of credibility measure, in large part due to its convenient mathematical properties. Given the MSE measure of confidence, it is a natural question to find the specific estimator for which the MSE takes its minimal value. It is a well-established fact that the minimal mean-squared error (MMSE) estimator, denoted $\bar{\lambda}$, is given by the average parameter value with respect to the posterior PDF [6]

$$\bar{\lambda}(\mathbf{x}_n) := \int_{\Lambda} d\lambda p_{\Lambda}(\lambda|\mathbf{x}_n)\lambda. \quad (3.4)$$

This result can be derived by considering the first and second variational derivatives of the MSE with respect to the estimator. Requiring that the first variation vanish gives the posterior mean, and showing that the second variation is positive establishes the posterior mean as a minimum. Evaluating the MSE using the MMSE estimator as the estimator of choice, gives a lower bound on the MSE for an arbitrary choice of estimator

$$\text{MSE}[\lambda_{\text{est}}](\mathbf{x}_n) \geq \text{MSE}[\bar{\lambda}](\mathbf{x}_n) := \text{MMSE}(\mathbf{x}_n). \quad (3.5)$$

Here the posterior MMSE on the right-hand side, is simply the posterior variance of the variable $\lambda \in \Lambda$ over the posterior distribution $p_{\Lambda}(\lambda|\mathbf{x}_n)$. When working with the MMSE

estimator, it follows directly from Eq. (3.5) that the MSE equals the MMSE. Thus, the bound in Eq. (3.5) is tight.

In this thesis, we shall primarily consider the MAP estimator and the MMSE estimator as our estimators of choice. One might reasonably ask if there is any reason for considering the MAP estimator, given that the MMSE estimator is always attributed maximum credibility. Essentially this comes down to the fact that in adopting the MSE we are implicitly assuming that the posterior PDF can be well-captured by its mean and variance. If the MAP and MMSE estimators are radically different, then this can be taken as an indication that adopting the MSE as a measure of confidence is not suitable.

3.3 Posterior Cramér-Rao bound

Basic computational convenience, or potential insight into fundamental limitations on estimation precision, motivates the study of lower bounds on the MMSE. Ideally, for a lower bound to be informative it must be computationally tractable and it should be tight, i.e. the MMSE should saturate the bound. An important family of lower bounds, known as the Weiss-Weinstein family [70], follow from various applications of the Cauchy-Schwarz inequality. Given two functions $f, g : \Lambda \rightarrow \mathbb{R}$ the Cauchy-Schwarz inequality states that

$$\langle f(\lambda)g(\lambda) \rangle^2 \leq \langle f(\lambda)^2 \rangle \langle g(\lambda)^2 \rangle, \quad (3.6)$$

where $\langle \cdot \rangle$ denotes an average over a PDF on the parameter space Λ [59]. In particular, the MMSE satisfies the posterior Cramér-Rao bound (CRB) [69, 70]. To derive this bound we first define the quantity

$$B(\mathbf{x}_n) := \int_{\Lambda} d\lambda \sqrt{p_{\Lambda}(\lambda|\mathbf{x}_n)} (\bar{\lambda}(\mathbf{x}_n) - \lambda) \times \sqrt{p_{\Lambda}(\lambda|\mathbf{x}_n)} \partial_{\lambda} \log p_{\Lambda}(\lambda|\mathbf{x}_n) \quad (3.7)$$

where we implicitly assume differentiability of the posterior PDF. The quantity $B(\mathbf{x}_n)$ is intentionally defined to motivate an application of the Cauchy-Schwarz inequality. Before getting to that, we first integrate $B(\mathbf{x}_n)$ by parts yielding

$$B(\mathbf{x}_n) = 1 + \{(\bar{\lambda}(\mathbf{x}_n) - \lambda)p_{\Lambda}(\lambda|\mathbf{x}_n)\}_{\lambda \in \mathcal{B}(\Lambda)}, \quad (3.8)$$

where $\mathcal{B}(\Lambda)$ denotes the boundaries of the parameter space, and the brace notation denotes a boundary term, i.e. the difference between the expression evaluated at the upper and lower boundaries. The boundary term above vanishes if the posterior PDF satisfies two conditions:

$$p_{\Lambda}(\lambda|\mathbf{x}_n) = 0, \quad \text{for } \lambda \in \mathcal{B}(\Lambda), \quad (3.9)$$

$$\lambda p_{\Lambda}(\lambda|\mathbf{x}_n) = 0, \quad \text{for } \lambda \in \mathcal{B}(\Lambda). \quad (3.10)$$

When these two conditions are satisfied it follows that $B(\mathbf{x}_n) = 1$. Notice that if these two conditions are true for the initial prior PDF $p_{\Lambda}(\lambda)$, then they will also hold for the full sequence of posterior PDFs.

Throughout this thesis we adopt Eq. (3.9) and Eq. (3.10) as constraints on the class of initial prior PDFs considered. If we return to the definition of $B(\mathbf{x}_n)$, Eq. (3.7), and apply the Cauchy-Schwarz inequality, we obtain the CRB:

$$\text{MMSE}(\mathbf{x}_n) \geq \mathcal{Q}(\mathbf{x}_n)^{-1} := \text{CRB}(\mathbf{x}_n), \quad (3.11)$$

where we have defined the posterior information [47]

$$\mathcal{Q}(\mathbf{x}_n) := \int_{\Lambda} d\lambda p_{\Lambda}(\lambda|\mathbf{x}_n) (\partial_{\lambda} \log p_{\Lambda}(\lambda|\mathbf{x}_n))^2. \quad (3.12)$$

Notice that the CRB is only well-defined for a given measurement trajectory if the associated posterior information is non-zero. From a computational point of view it is by no means obvious that the CRB should be any less demanding to compute than the MMSE itself, and indeed this is not the case in general. Nonetheless, the CRB serves as a starting point from which we can derive computationally more tractable lower-bounds, and as we will argue below, it is in principle tight for a wide class of likelihood functions.

3.3.1 Saturability of the CRB

One of the advantages of starting with the CRB is that we can write down a condition for when it is a tight bound. The condition for saturating the CRB is the same as the condition under which the Cauchy-Schwarz inequality becomes an equality. For our purposes, this condition can be written as [69, 70]

$$\mathcal{Q}(\mathbf{x}_n)(\bar{\lambda}(\mathbf{x}_n) - \lambda) = \partial_{\lambda} \log p_{\Lambda}(\lambda|\mathbf{x}_n). \quad (3.13)$$

If we substitute this condition into Eq. (3.7), we immediately see that the CRB becomes an equality. Integrating the above condition, and performing a few straightforward manipulations, allows us to conclude that a posterior PDF saturating the CRB has the form

$$p_{\Lambda}(\lambda|\mathbf{x}_n) = b(\mathbf{x}_n) \exp \left[-\frac{\mathcal{Q}(\mathbf{x}_n)}{2} (\lambda - \bar{\lambda}(\mathbf{x}_n))^2 \right], \quad (3.14)$$

where we have the normalization factor

$$b(\mathbf{x}_n)^{-1} := \int_{\Lambda} d\lambda \exp \left[-\frac{\mathcal{Q}(\mathbf{x}_n)}{2} (\lambda - \bar{\lambda}(\mathbf{x}_n))^2 \right]. \quad (3.15)$$

If the posterior information $\mathcal{Q}(\mathbf{x}_n)$ is sufficiently large, and the estimate $\bar{\lambda}(\mathbf{x}_n)$ is not close to the boundary of the parameter space Λ , then the above posterior PDF approximately takes the form of a normal distribution. Alternatively, this normal form is obtained exactly if the parameter space is the entire real line.

Attaining a posterior PDF of the form of Eq. (3.14) requires two things: (i) a likelihood function belonging to the so-called exponential model family, and (ii) a prior PDF which is conjugate to the likelihood function [6, 70]. The exponential model family includes most of the common distributions e.g. the Poisson, Bernoulli, Gamma, Gaussian, χ -squared and binomial distributions [6], and a conjugate prior refers to a prior such that the posterior belongs to the same family of PDFs. If we consider a single-shot measurement giving an outcome $x \in \mathcal{X}$, an exponential family model is characterized by three functions: $h(x) \geq 0$ and $\boldsymbol{\eta}(\lambda), \mathbf{t}(x) \in \mathbb{R}^2$, such that

$$p_{\mathcal{X}}^{\text{exp}}(x|\lambda) := h(x) \exp [\boldsymbol{\eta}(\lambda)^T \mathbf{t}(x) - A(\lambda)], \quad (3.16)$$

$$A(\lambda) := \int_{\mathcal{X}} dx h(x) \exp [\boldsymbol{\eta}(\lambda)^T \mathbf{t}(x)]. \quad (3.17)$$

Different choices of the three functions gives rise to different distributions within the exponential family. The prior PDF that is conjugate to the exponential-model likelihood function, can be expressed, given a constant $\gamma \in \mathbb{R}$ and $\boldsymbol{\mu} \in \mathbb{R}^2$, as follows

$$p_{\Lambda}(\lambda; \gamma, \boldsymbol{\mu}) := \kappa(\gamma, \boldsymbol{\mu}) \exp [\boldsymbol{\eta}(\lambda)^T \boldsymbol{\mu} - \gamma A(\lambda)], \quad (3.18)$$

$$\kappa(\gamma, \boldsymbol{\mu})^{-1} := \int_{\Lambda} d\lambda \exp [\boldsymbol{\eta}(\lambda)^T \boldsymbol{\mu} - \gamma A(\lambda)]. \quad (3.19)$$

Computing the posterior PDF resulting from this prior and the exponential-family likelihood function, gives a distribution of the form $p_{\Lambda}(\lambda; \gamma + 1, \boldsymbol{\mu} + \mathbf{t}(x))$, which, as expected, belongs to the same family of PDFs as the conjugate prior. The requirement for saturating the CRB is then that the PDF $p_{\Lambda}(\lambda; \gamma + 1, \boldsymbol{\mu} + \mathbf{t}(x))$ takes the form of Eq. (3.14). This condition can be written as

$$\boldsymbol{\eta}(\lambda)^T (\boldsymbol{\mu} + \mathbf{t}(x)) - (\gamma + 1)A(\lambda) = -\frac{\mathcal{Q}(x)}{2}(\lambda - \bar{\lambda}(x))^2, \quad (3.20)$$

which automatically implies $b(x) = \kappa(\gamma + 1, \boldsymbol{\mu} + \mathbf{t}(x))$. This condition would have to be checked in specific cases, but serves to show that there exist a very wide class of models belonging to the exponential family, for which the CRB is a tight bound.

3.3.2 Asymptotic statistics

Neither the MSE itself, nor the CRB, can be computed without an explicit expression for the posterior PDF. In general, it is not feasible to work with an analytic description of the posterior PDF, and one must resort to numerical investigations. Remarkably, however, it is often possible to provide good asymptotic approximations to the posterior PDF, and the accuracy of such approximations increases as the size n of the data-sample increases. The study of these approximations is called asymptotic statistics. The key result of this theory, which will be of relevance here, is that of local asymptotic normality [73, 74].

Before we can present the explicit approximation scheme, we must discuss the nature of the parametric model under study. Recall that we generally consider a stochastic variable X_j at the j th measurement step, and that a specific realization $x_j \in \mathcal{X}_j$ of the stochastic variable is sampled from the likelihood function $p_{\mathcal{X}_j}(x_j | \lambda_0, \mathbf{x}_{j-1})$, with respect to a true parameter value $\lambda_0 \in \Lambda$. In our context, the conditioning on the past measurement trajectory is to be understood as the ability to adapt the measurement based on the data accumulated in the past. More specifically, we assume that the experimenter has available the posterior PDF resulting from the measurement data \mathbf{x}_{j-1} , and that an adaptive strategy exists to optimize the expected performance of the measurement with respect to the parameter estimation task, based on the information represented by $p_{\Lambda}(\lambda | \mathbf{x}_{j-1})$.

The results of asymptotic statistics do not apply directly to adaptive strategies. However, progress can be made if we formally partition our estimation protocol into an adaptive and a non-adaptive phase, i.e. we assume that there is a measurement step, denoted k_c , after which no adaptation takes place. The measurements associated with steps $j > k_c$ are then independent and identical. Within this context, the key corollary of local asymptotic normality that we will apply here, is the Bernstein-von Mises theorem [73, 74].

Bernstein-von Mises theorem:

Consider the likelihood $p_\Lambda(x|\lambda, \mathbf{x}_{k_c})$ for $\lambda \in \Lambda$ and $x \in \mathcal{X}$. Let X_{k_c+1}, \dots, X_n be independently sampled from $p_\Lambda(x|\lambda_0, \mathbf{x}_{k_c})$ for some $\lambda_0 \in \Lambda$. We assume that (i) λ_0 is an interior point of Λ , (ii) the likelihood function is twice differentiable around λ_0 , and (iii) the Fisher information associated with the likelihood function is non-zero and finite around λ_0 . Then given that the initial prior PDF assigns non-zero probability to the region around λ_0 , it follows that

$$p_\Lambda(\lambda|\mathbf{x}_n) \rightarrow \mathcal{N}(\lambda; \bar{\lambda}(\mathbf{x}_n), \mathcal{F}_{\Pi_n}^{-1}(\bar{\lambda}(\mathbf{x}_n))), \quad \text{as } n \rightarrow \infty, \quad (3.21)$$

where \mathcal{N} denotes a normal PDF, with a mean given by the MMSE estimator and a variance given by the inverse Fisher information associated with the measurement protocol $\Pi_n := \{\Pi_1, \dots, \Pi_n\}$ where $\Pi_k := \{\Pi_k(x) \text{ for } x \in \mathcal{X}_k\}$. Furthermore, the MMSE estimator is asymptotically consistent and tends to the parameter value λ_0 . For completeness we recall the definition of the Fisher information:

$$\mathcal{F}_{\Pi_n}(\lambda) = \int_{\mathcal{X}_n} d\mathbb{P}_0(\mathbf{x}_n) p_{\mathcal{X}_n}(\mathbf{x}_n|\lambda) (\partial_\lambda \log p_{\mathcal{X}_n}(\mathbf{x}_n|\lambda))^2, \quad (3.22)$$

associated with the sequential likelihood function of the full estimation protocol. Note that we are essentially assuming that any adaptation of the measurement can be restricted to an initial non-asymptotic regime, and that this adaptation becomes irrelevant in the asymptotic regime.

Hence, the posterior PDF is asymptotically normal. We mention that if condition (i) in the above theorem is not valid, one can still derive an asymptotic approximation. In this case it takes the form of a normal PDF with possible gamma function corrections [75]. Given that we assume a prior PDF satisfying the constraints of Eq. (3.9) and Eq. (3.10), we are in effect assuming that condition (i) is always satisfied. Lastly, we note that although we formally partitioned the estimation protocol into an adaptive and a non-adaptive phase, it is often the case that the adaptive strategy employed converges on a specific measurement design. In this case, the non-adaptive phase is effectively reached without the need for imposing a cutoff.

If we substitute the normal form of the posterior PDF into the posterior information Eq. (3.12), we find that the posterior information tends asymptotically to the Fisher information with respect to the MMSE estimate, that is

$$\mathcal{Q}(\mathbf{x}_n) \rightarrow \mathcal{F}_{\Pi_n}(\bar{\lambda}(\mathbf{x}_n)), \quad \text{as } n \rightarrow \infty. \quad (3.23)$$

Note that although we assumed an initial prior PDF satisfying the two constraints Eq. (3.9) and Eq. (3.10), such that the boundary term $B(\mathbf{x}_n)$ equals unity, this property of the boundary term is true asymptotically provided that the Bernstein-von Mises theorem holds, regardless of the initial prior PDF. An important implication of the Bernstein-von Mises theorem is that in the asymptotic limit the MMSE saturates the CRB

$$\text{MMSE}(\mathbf{x}_n) \rightarrow \mathcal{F}_{\Pi_n}(\bar{\lambda}(\mathbf{x}_n))^{-1}, \quad \text{as } n \rightarrow \infty. \quad (3.24)$$

Note that we cannot say anything about the rate of convergence to the CRB, and it might well be the case that all results based on the asymptotic behaviour are irrelevant

for all practical purposes. Lastly, we note that Eq. (3.24) coincides with the form of the Cramér-Rao bound studied within the frequentist approach to estimation theory [6, 76]. Here we have seen how this form emerges as an asymptotic approximation within the Bayesian framework. This insight is of course not novel [70].

3.4 Bayesian mean-squared error

The various estimators described above allows one to analyse a specific experimental data set. In themselves, however, these tools do not provide a direct guide on how to enhance the design of the measurement protocol to improve the expected performance with respect to the parameter estimation task. In other words, the tools developed so far allows us to extract an estimate and assign a credibility, but there is no direct guide on how to design the measurement to optimize the credibility we expect to assign. Thus, before we can discuss measurement design, we must construct a measure of the expected credibility resulting from a given measurement protocol. Throughout, the approach taken to evaluate the expected performance of an estimation protocol is the following stochastic simulation algorithm:

Stochastic simulation algorithm:

1. A parameter value λ_0 is sampled from the initial prior PDF $p_\Lambda(\lambda_0)$.
2. n measurement outcomes are generated sequentially via stochastic sampling from the sequence of correlated likelihood functions:

$$X_k \sim p_{\mathcal{X}_k}(x_k | \lambda_0, \mathbf{x}_{k-1}) \quad \text{for } k = 1, \dots, n. \quad (3.25)$$

3. The set of posterior PDFs, and the associated trajectory of posterior MSEs, are computed from the data using Bayes theorem.
4. The first three steps of the algorithm are repeated a large number of times (ν), and the generated data is stored at each iteration.

The outcome of the above stochastic simulation is a set of ν posterior MSE trajectories. By the law of large numbers the average computed over measurement trajectories converges to the likelihood-averaged posterior MSE averaged over the initial prior PDF [59]. This averaged quantity is typically referred to as the Bayesian mean-squared error [6, 55, 56]. We will denote it as $\overline{\text{MSE}}$. It is defined as

$$\begin{aligned} \overline{\text{MSE}}[\lambda_{\text{est}}](n) &:= \int_{\Lambda} d\lambda_0 p_\Lambda(\lambda_0) \int_{\mathcal{X}_n} d\mathbb{P}_0(\mathbf{x}_n) p_{\mathcal{X}_n}(\mathbf{x}_n | \lambda_0) \text{MSE}[\lambda_{\text{est}}](\mathbf{x}_n) \\ &= \int_{\mathcal{X}_n} d\mathbb{P}_0(\mathbf{x}_n) p_{\mathcal{X}_n}(\mathbf{x}_n) \text{MSE}[\lambda_{\text{est}}](\mathbf{x}_n), \end{aligned} \quad (3.26)$$

where the second equality follows from Eq. (2.22) defining the evidence function $p_{\mathcal{X}_n}(\mathbf{x}_n)$. We will use the bar to distinguish between posterior quantities, such as the MSE, from Bayesian quantities, such as the $\overline{\text{MSE}}$. Note that the computed MSE trajectories provide information beyond the $\overline{\text{MSE}}$. In particular, the set of trajectories also provide insight into stochastic fluctuations around the $\overline{\text{MSE}}$, and in principle we could consider statistics beyond the average value.

In adopting the $\overline{\text{MSE}}$ as a quantifier of the expected performance, we are in effect modelling the true parameter value as a stochastic variable, which is sampled from the initial prior PDF. This is the reason that the resulting error is typically referred to as the *Bayesian* mean-squared error. Historically, the paradigm of Bayesian parameter estimation [55, 56] referred to the scenario in which both the data and the parameter to be estimated are considered to be stochastic variables. The initial prior then represents a true distribution according to which a value of the parameter is sampled in each realization of the experiment. Here, we take the initial prior to reflect our initial ignorance of the deterministic parameter. The $\overline{\text{MSE}}$ then gives the MSE that we expect to obtain on average, given our ignorance, if we assume that the true parameter value is sampled from the initial prior PDF.

Lastly we note that although the $\overline{\text{MSE}}$ is a deterministic quantity, it is still defined with respect to a specific choice of estimator. From the definition of the $\overline{\text{MSE}}$, namely that it can be written as an average of the MSE, it follows directly that we have the lower bound

$$\overline{\text{MSE}}[\lambda_{\text{est}}](n) \geq \overline{\text{MSE}}[\bar{\lambda}](n) := \overline{\text{MMSE}}(n), \quad (3.27)$$

where $\bar{\lambda}$ denotes the MMSE estimator, and we have defined the Bayesian minimal mean-squared error ($\overline{\text{MMSE}}$). The above lower-bound is trivially tight if the MMSE estimator is employed. In the following we study the MMSE exclusively.

3.4.1 Total variance decomposition

Conceptually, and for practical reasons when considering adaptive estimation protocols, it is desirable to provide a sequential decomposition of the $\overline{\text{MMSE}}$. This can indeed be done if we utilize the law of total variance. For a detailed discussion of this law see Ref. [36] or the brief discussion in Appendix A. To achieve the aim of a sequential decomposition, we first note that from Eq. (2.22) defining the sequential evidence function, it follows that

$$\int_{\mathbf{x}_n} d\mathbb{P}_0(\mathbf{x}_n) p_{\mathbf{x}_n}(\mathbf{x}_n)(\cdot) = \int_{\mathcal{X}_1} d\mathbb{P}_0(x_1) p_{\mathcal{X}_1}(x_1) \int_{\mathcal{X}_2} d\mathbb{P}_0(x_2) p_{\mathcal{X}_2}(x_2|x_1) \dots \int_{\mathcal{X}_n} d\mathbb{P}_0(x_n) p_{\mathcal{X}_n}(x_n|\mathbf{x}_{n-1})(\cdot), \quad (3.28)$$

where the fat dot denotes the function to be averaged. More generally this statement is known as the law of total expectation, or the tower property [35, 36]. The tower property allows us to decompose the $\overline{\text{MMSE}}$ by iteratively applying the law of total variance. In general, this law states that the expectation of a conditional variance can be decomposed into the difference of the unconditional variance and the variance of the conditional expectation [36]. For our purposes, the law of total variance can be stated as

$$\int_{\mathcal{X}_n} d\mathbb{P}_0(x_n) p_{\mathcal{X}_n}(x_n|\mathbf{x}_{n-1}) \text{MMSE}(\mathbf{x}_n) = \text{MMSE}(\mathbf{x}_{n-1}) - \mathcal{K}_{\Pi_n}(\mathbf{x}_{n-1}) \quad (3.29)$$

where we have defined the added information:

$$\mathcal{K}_{\Pi_n}(\mathbf{x}_{n-1}) := \int_{\mathcal{X}_n} d\mathbb{P}_0(x_n) p_{\mathcal{X}_n}(x_n|\mathbf{x}_{n-1}) (\bar{\lambda}(\mathbf{x}_n) - \bar{\lambda}(\mathbf{x}_{n-1}))^2. \quad (3.30)$$

The added information gives the average reduction in the MMSE resulting from performing a measurement. Intuitively, we see that the added information is large if the set of

possible posterior MMSE estimates have a large spread around the prior MMSE estimate. The more we are likely to change the estimate based on the acquired data, the more information must have been added. By iteratively applying the tower property and the law of total variance, it follows straightforwardly that the $\overline{\text{MMSE}}$ can be written as:

$$\overline{\text{MMSE}}(n) = \overline{\text{MMSE}}(0) - \sum_{k=1}^n \int_{\mathbf{x}_{k-1}} d\mathbb{P}_0(\mathbf{x}_{k-1}) p_{\mathbf{x}_{k-1}}(\mathbf{x}_{k-1}) \mathcal{K}_{\Pi_k}(\mathbf{x}_{k-1}), \quad (3.31)$$

where $\overline{\text{MMSE}}(0)$ is simply the variance of the initial prior PDF, and we recall that \mathbf{x}_0 is an empty vector such that $\int d\mu_0(\mathbf{x}_0) p_{\mathbf{x}_0}(\mathbf{x}_0) = 1$. The above expression shows that the $\overline{\text{MMSE}}$ can be written as the difference between the initial prior variance and a sum of evidence-averaged added information terms. It provides a sequential decomposition of the $\overline{\text{MMSE}}$ in the sense that each additional measurement adds an added information term. A decomposition of the $\overline{\text{MMSE}}$ similar to the one above was also considered in Ref. [72], where it was used to define a Bayesian analogue of the SLD.

3.4.2 Bayesian Cramér-Rao bound

We now demonstrate that the $\overline{\text{MMSE}}$ satisfies a number of Bayesian lower bounds. First, recall Eq. (3.11) stating the CRB is satisfied by the MMSE. It then follows directly that the $\overline{\text{MMSE}}$ satisfies the analogous Bayesian Cramér-Rao bound [69, 70]. We denote this by $\overline{\text{CRB}}$. It is defined as

$$\overline{\text{MMSE}}(n) \geq \int_{\Lambda} d\lambda p_{\Lambda}(\lambda) \int_{\mathbf{x}_n} d\mathbb{P}_0(\mathbf{x}_n) p_{\mathbf{x}_n}(\mathbf{x}_n|\lambda) \mathcal{Q}(\mathbf{x}_n)^{-1} := \overline{\text{CRB}}(n). \quad (3.32)$$

Note that the $\overline{\text{CRB}}$ is a tight inequality if the CRB itself is tight, i.e. if the posterior distribution is given by Eq. (3.14). As was also pointed out in the case of the CRB, it is not obvious that it is any less demanding to compute the $\overline{\text{CRB}}$, than it is to compute the $\overline{\text{MMSE}}$ itself. Indeed this is not the case in general. Note, however, that it is possible to compute the $\overline{\text{CRB}}$ using the same stochastic simulation algorithm that is used to compute the $\overline{\text{MMSE}}$. Our main reason for considering the $\overline{\text{CRB}}$ is that we can obtain an asymptotic approximation of the $\overline{\text{MMSE}}$ via the Bernstein-von Mises theorem. In particular, from Eq. (3.24) we have

$$\begin{aligned} \overline{\text{MMSE}}(n) &\rightarrow \int_{\Lambda} d\lambda p_{\Lambda}(\lambda) \int_{\mathbf{x}_n} d\mathbb{P}_0(\mathbf{x}_n) p_{\mathbf{x}_n}(\mathbf{x}_n|\lambda) \mathcal{F}_{\Pi_n}(\bar{\lambda}(\mathbf{x}_n))^{-1}, \quad \text{as } n \rightarrow \infty, \\ &= \int_{\Lambda} d\lambda p_{\Lambda}(\lambda) \mathcal{F}_{\Pi_n}(\lambda)^{-1}, \quad \text{as } n \rightarrow \infty, \end{aligned} \quad (3.33)$$

where the second equality should be understood as an asymptotic convergence and follows from the fact that the MMSE estimator converges to the true parameter value. The $\overline{\text{MMSE}}$ thus tends asymptotically to the prior average of the inverse Fisher information associated with the measurement protocol. This result is intuitive. Consider a Bayesian scenario in which the true parameter value is sampled according to the prior PDF. At the trajectory level the MMSE converges to the inverse Fisher information at the sampled parameter value. Averaging over trajectories is then equivalent to averaging over the prior PDF.

3.4.3 Van Trees bound

Assuming that the Fisher information is efficiently computable, we have at this point derived efficient asymptotic bounds, Eq. (3.24) and Eq. (3.33), on the MMSE and the $\overline{\text{MMSE}}$ respectively. Crucially, these bound are asymptotically tight, i.e. tight for sufficiently large n . However, with the exception of direct simulation, we have at this point no way of determining what constitutes sufficiently large n .

Partial progress on this problem can be made if we consider the so-called van Trees bound [54, 55] on the $\overline{\text{MMSE}}$. We denote this bound by $\overline{\text{VTB}}$. The $\overline{\text{VTB}}$ follows directly from the $\overline{\text{CRB}}$ via an application of Jensen's inequality. If Y is a stochastic variable drawn from a probability measure and f is a convex function, then Jensen's inequality states that $\langle f(Y) \rangle \geq f(\langle Y \rangle)$, where $\langle \cdot \rangle$ denotes an averaging over the probability measure [59]. Based on Jensen's inequality, and since $f(Y) = Y^{-1}$ is a convex function, we can derive the $\overline{\text{VTB}}$

$$\begin{aligned} \overline{\text{CRB}}(n) &= \int_{\mathbf{x}_n} d\mathbb{P}_0(\mathbf{x}_n) p_{\mathbf{x}_n}(\mathbf{x}_n) \mathcal{Q}(\mathbf{x}_n)^{-1} \\ &\geq \left[\int_{\mathbf{x}_n} d\mathbb{P}_0(\mathbf{x}_n) p_{\mathbf{x}_n}(\mathbf{x}_n) \mathcal{Q}(\mathbf{x}_n) \right]^{-1} := \overline{\text{VTB}}(n). \end{aligned} \quad (3.34)$$

The $\overline{\text{CRB}}$ saturates the $\overline{\text{VTB}}$ if and only if the posterior information is a linear function. For example, if the posterior information becomes a constant independent of the measurement trajectory [59]. Straightforward manipulations allow us to rewrite the $\overline{\text{VTB}}$ as follows

$$\begin{aligned} \overline{\text{VTB}}(n)^{-1} &= \int_{\mathbf{x}_n} d\mathbb{P}_0(\mathbf{x}_n) p_{\mathbf{x}_n}(\mathbf{x}_n) \mathcal{Q}(\mathbf{x}_n) \\ &= \int_{\mathbf{x}_n \times \Lambda} d\mathbb{P}_0(\mathbf{x}_n) d\lambda p_{\Lambda}(\lambda) p_{\mathbf{x}_n}(\mathbf{x}_n | \lambda) \\ &\quad \times (\partial_{\lambda} \log p_{\Lambda}(\lambda) + \partial_{\lambda} \log p_{\mathbf{x}_n}(\mathbf{x}_n | \lambda))^2 \\ &= \int_{\mathbf{x}_n \times \Lambda} d\mathbb{P}_0(\mathbf{x}_n) d\lambda p_{\Lambda}(\lambda) p_{\mathbf{x}_n}(\mathbf{x}_n | \lambda) \\ &\quad \times ((\partial_{\lambda} \log p_{\Lambda}(\lambda))^2 + (\partial_{\lambda} \log p_{\mathbf{x}_n}(\mathbf{x}_n | \lambda))^2), \end{aligned} \quad (3.35)$$

where the first equality is the definition, the second equality follows from substituting the definition of the posterior information (Eq. (3.12)) and applying Bayes' theorem, and the third equality follows from cross-terms vanishing under the integral. If we separate the last expression into two terms and perform the integrals, we arrive at the final form of the $\overline{\text{VTB}}$

$$\overline{\text{VTB}}(n)^{-1} = \mathcal{Q}_{\text{prior}} + \int_{\Lambda} d\lambda p_{\Lambda}(\lambda) \mathcal{F}_{\mathbf{\Pi}_n}(\lambda), \quad (3.36)$$

where the initial posterior, or prior, information $\mathcal{Q}_{\text{prior}} := \mathcal{Q}(\mathbf{x}_0)$ quantifies the information encoded in the prior PDF, and the second term is the Fisher information averaged over the prior PDF. The partial progress alluded to at the beginning of this paragraph is then as follows: The number of measurements, n , must be large enough that the information acquired via the measurement, represented by the averaged Fisher information term, is greater than the initial posterior information $\mathcal{Q}_{\text{prior}}$.

Working with the $\overline{\text{VTB}}$ is often computationally advantageous over working with the $\overline{\text{CRB}}$. This is due to the fact that the $\overline{\text{VTB}}$ is expressed directly in terms of the Fisher

information, which is computationally less demanding than the posterior information. However, this comes at the price that the $\overline{\text{VTB}}$ is not in general a tight inequality. If we take Eq. (3.33), and recall Jensen's inequality, we see that the $\overline{\text{VTB}}$ is asymptotically tight if the Fisher information tends to a constant across the width of the prior PDF. Note, however, that the $\overline{\text{VTB}}$ is not tight in general, not even asymptotically.

Additivity of the adaptive Fisher information

The last result we need, before we can turn to the question of designing optimal measurements, is a generalized version of the additivity property of the Fisher information. In particular, we need the fact that the full Fisher information can be re-written as

$$\begin{aligned}
\mathcal{F}_{\Pi_n}(\lambda) &= \int_{\mathbf{x}_n} d\mathbb{P}_0(\mathbf{x}_n) p_{\mathbf{x}_n}(\mathbf{x}_n|\lambda) \left(\frac{\partial \log p_{\mathbf{x}_n}(\mathbf{x}_n|\lambda)}{\partial \lambda} \right)^2 \\
&= \int_{\mathbf{x}_n} d\mathbb{P}_0(\mathbf{x}_n) p_{\mathbf{x}_n}(\mathbf{x}_n|\lambda) \sum_{k=1}^n \left(\frac{\partial \log p_{\mathcal{X}_k}(x_k|\lambda, \mathbf{x}_{k-1})}{\partial \lambda} \right)^2 \\
&= \sum_{k=1}^n \int_{\mathbf{x}_k} d\mathbb{P}_0(\mathbf{x}_k) p_{\mathbf{x}_k}(\mathbf{x}_k|\lambda) \left(\frac{\partial \log p_{\mathcal{X}_k}(x_k|\lambda, \mathbf{x}_{k-1})}{\partial \lambda} \right)^2 \\
&= \sum_{k=1}^n \int_{\mathbf{x}_{k-1}} d\mathbb{P}_0(\mathbf{x}_{k-1}) p_{\mathbf{x}_{k-1}}(\mathbf{x}_{k-1}|\lambda) \mathcal{F}_{\Pi_k}(\lambda, \mathbf{x}_{k-1}),
\end{aligned} \tag{3.37}$$

where the second equality follows from writing the sequential likelihood function as a product and noticing that cross-terms vanish, the third equality follows from moving the averaging within the sum, and the fourth equality follows from the decomposition $p_{\mathbf{x}_n}(\mathbf{x}_n|\lambda) = p_{\mathbf{x}_{k-1}}(\mathbf{x}_{k-1}|\lambda) p_{\mathcal{X}_k}(x_k|\lambda, \mathbf{x}_{k-1})$. Furthermore, we again note that \mathbf{x}_0 denotes an empty vector such that $\int_{\mathbf{x}_0} d\mathbb{P}_0(\mathbf{x}_0) p_{\mathbf{x}_0}(\mathbf{x}_0|\lambda) = 1$, and $\mathcal{F}_{\Pi_k}(\lambda, \mathbf{x}_{k-1})$ is the Fisher information associated with the likelihood function $p_{\mathcal{X}_k}(x_k|\lambda, \mathbf{x}_{k-1})$. Notice that if we are considering a non-adaptive scenario, then the above expression becomes the standard additivity property of the Fisher information.

Similarly to be $\overline{\text{MMSE}}$ we can obtain a sequential decomposition of the $\overline{\text{VTB}}$. If we take the above generalized additivity property of the full Fisher information, and substitute it into the $\overline{\text{VTB}}$, we obtain the following expression

$$\begin{aligned}
\overline{\text{VTB}}(n)^{-1} &= \mathcal{Q}_{\text{prior}} + \sum_{k=1}^n \int_{\mathbf{x}_{k-1}} d\mathbb{P}_0(\mathbf{x}_{k-1}) p_{\mathbf{x}_{k-1}}(\mathbf{x}_{k-1}) \\
&\quad \times \int_{\Lambda} d\lambda p_{\Lambda}(\lambda|\mathbf{x}_{k-1}) \mathcal{F}_{\Pi_k}(\lambda, \mathbf{x}_{k-1}) \\
&= \mathcal{Q}_{\text{prior}} + \sum_{k=1}^n \int_{\mathbf{x}_{k-1}} d\mathbb{P}_0(\mathbf{x}_{k-1}) p_{\mathbf{x}_{k-1}}(\mathbf{x}_{k-1}) \bar{\mathcal{F}}_{\Pi_k}(\mathbf{x}_{k-1}),
\end{aligned} \tag{3.38}$$

where we have defined a prior-averaged Fisher information

$$\bar{\mathcal{F}}_{\Pi_k}(\mathbf{x}_{k-1}) := \int_{\Lambda} d\lambda p_{\Lambda}(\lambda|\mathbf{x}_{k-1}) \mathcal{F}_{\Pi_k}(\lambda, \mathbf{x}_{k-1}). \tag{3.39}$$

Similarly to the total variance decomposition of the $\overline{\text{MMSE}}$ given in Eq. (3.31), the above expression for the $\overline{\text{VTB}}$ provides us with a sequential decomposition, in which

each measurement performed adds a prior averaged Fisher information term. Lastly, we note that in the special case of a non-adaptive estimation protocol in which the same measurement Π is repeated n times, the $\overline{\text{VTB}}$ takes the form

$$\overline{\text{VTB}}(n)^{-1} = \mathcal{Q}_{\text{prior}} + n \int_{\Lambda} d\lambda p_{\Lambda}(\lambda) \mathcal{F}_{\Pi}(\lambda), \quad (3.40)$$

and thus reduce to evaluating the prior average of the Fisher information associated with the implemented measurement. We will discuss the restriction imposed by considering only non-adaptive estimation protocols further in chapter 5.

3.5 Adaptive measurement design

For adaptive estimation protocols, the measure of confidence in an estimate must be supplemented by a criterion for optimal measurement design. If we focus on the k th measurement step of the protocol, then a natural design criterion consists of identifying the measurement, represented by a POVM with elements $\Pi_k = \{\Pi(x) \text{ for } x \in \mathcal{X}_k\}$, for which the associated added information is as large as possible. Typically the optimization over measurements will be restricted to a constrained space of allowed measurements, \mathcal{C} . An example of such a constrained optimization space could be that the sample is probed using a quantum probe, and the measurement can be optimized with respect to interaction time, coupling strength, or the energy level structure of the probe. We can then formally express our first design criterion as the maximization problem

$$\Pi_{k,\text{opt}}^{(1)} := \arg \max_{\Pi_k \in \mathcal{C}} \mathcal{K}_{\Pi_k}(\mathbf{x}_{k-1}), \quad (3.41)$$

where we stress that the optimal measurement identified is a functional of the posterior PDF resulting from the past measurement trajectory \mathbf{x}_{k-1} . We will refer to this design strategy as $\overline{\text{MMSE}}$ optimization, as it corresponds to a direct sequential minimization of the $\overline{\text{MMSE}}$. Note that we can recover non-adaptive estimation protocols as a special case, by restricting the optimization space \mathcal{C} to consist of a single measurement.

3.5.1 Asymptotic expression for the added information

We now investigate the form that the added information takes in the asymptotic limit. In particular we assume that the Bernstein-von Mises theorem holds, such that the PDF $p_{\Lambda}(\lambda|\mathbf{x}_{k-1})$ is a narrow Gaussian around the MMSE estimator. We first recall the definition of the added information

$$\begin{aligned} \mathcal{K}_{\Pi_k}(\mathbf{x}_{k-1}) &= \int_{\mathcal{X}_k} d\mathbb{P}_0(x_k) p_{\mathcal{X}_k}(x_k|\mathbf{x}_{k-1}) (\bar{\lambda}(\mathbf{x}_k) - \bar{\lambda}(\mathbf{x}_{k-1}))^2 \\ &= \int_{\Lambda} d\lambda p_{\Lambda}(\lambda|\mathbf{x}_{k-1}) \int_{\mathcal{X}_k} d\mathbb{P}_0(x_k) p_{\mathcal{X}_k}(x_k|\lambda, \mathbf{x}_{k-1}) \\ &\quad \times (\bar{\lambda}(\mathbf{x}_k) - \bar{\lambda}(\mathbf{x}_{k-1}))^2, \end{aligned} \quad (3.42)$$

where the first equality is the definition of the added information in Eq. (3.30), and the second equality follows from the definition of the evidence function in Eq. (2.18). If we

focus on the difference of MMSE estimators, we find

$$\begin{aligned}\bar{\lambda}(\mathbf{x}_k) - \bar{\lambda}(\mathbf{x}_{k-1}) &= \int_{\Lambda} d\lambda' p_{\Lambda}(\lambda'|\mathbf{x}_{k-1}) \left(\frac{p_{\mathcal{X}_k}(x_k|\lambda', \mathbf{x}_{k-1})}{p_{\mathcal{X}_k}(x_k|\mathbf{x}_{k-1})} - 1 \right) \lambda' \\ &\approx (\partial_{\lambda} \log p_{\mathcal{X}_k}(x_k|\lambda, \mathbf{x}_{k-1})) \int_{\Lambda} d\lambda' p_{\Lambda}(\lambda'|\mathbf{x}_{k-1}) \lambda' (\lambda' - \lambda),\end{aligned}\quad (3.43)$$

where the first equality follows from the definition of the MMSE estimators in Eq. (3.4), and the second approximation follows by locally Taylor-expanding the integrand around a parameter value λ to first order. Evaluating this expression at $\lambda = \bar{\lambda}(\mathbf{x}_{k-1})$, assuming that the Bernstein-von Mises theorem in Eq. (3.21) provides an accurate description, and substituting back into the added information above, we obtain the expression

$$\mathcal{K}_{\Pi_k}(\mathbf{x}_{k-1}) = \frac{\mathcal{F}_{\Pi_k}(\bar{\lambda}(\mathbf{x}_{k-1}))}{[\mathcal{F}_{\Pi_{k-1}}(\bar{\lambda}(\mathbf{x}_{k-1}))]^2}, \quad \text{as } k \rightarrow \infty. \quad (3.44)$$

Since only the numerator depends on the implemented measurement, it follows that in the asymptotic limit, $\overline{\text{MMSE}}$ optimization comes down to maximization of the Fisher information associated with the implemented measurement around the MMSE estimator. This result is to be expected. However, it serves to illustrate that in the asymptotic limit, $\overline{\text{MMSE}}$ optimization becomes equivalent to CRB optimization.

3.5.2 Alternative optimization strategy

While $\overline{\text{MMSE}}$ optimization is well-motivated, as it corresponds to a direct minimization of the $\overline{\text{MMSE}}$, it is generally associated with a substantial computational overhead. As an alternative to the above design strategy, we could search for the POVM that maximizes the prior averaged Fisher information

$$\Pi_{k,\text{opt}}^{(2)} := \arg \max_{\Pi_k \in \mathcal{C}} \bar{\mathcal{F}}_{\Pi_k}(\mathbf{x}_{k-1}). \quad (3.45)$$

We will refer to this design strategy as $\overline{\text{VTB}}$ optimization, as it corresponds to a direct sequential minimization of the $\overline{\text{VTB}}$. In this case we are minimizing a lower-bound on the $\overline{\text{MMSE}}$ rather than the quantity itself. Computationally, $\overline{\text{VTB}}$ optimization is preferable to $\overline{\text{MMSE}}$ optimization. This is due to the fact that computing the added information associated with a given measurement requires computing the full posterior ensemble and the associated MMSE estimators. On the other hand, $\overline{\text{VTB}}$ optimization requires only the Fisher information, which in many cases is a simpler computational task. Furthermore, we note that in the limit of a narrow prior $p_{\Lambda}(\lambda|\mathbf{x}_{k-1})$, the $\overline{\text{VTB}}$ optimization strategy simplifies to a local maximization of the Fisher information around the MMSE estimator. This feature implies that although $\overline{\text{VTB}}$ optimization might differ from $\overline{\text{MMSE}}$ optimization in the non-asymptotic regime, they become equivalent in the asymptotic limit.

3.6 Concluding remarks

In this chapter we have developed the theory of precision bounds in the case of estimation of a Euclidean parameter. Furthermore, we have described strategies for single-shot

measurement design. Armed with these theoretical tools, we can explore fundamental limits to parameter estimation in specific parameter estimation problems.

In the next chapter we take the insights gained on invariant estimation theory, precision bounds and adaptive strategies, and apply them to the problem of quantum thermometry, i.e. temperature estimation in quantum systems.

Chapter 4

Temperature Estimation Based on the Thermodynamic Length

This chapter contains text and figures from Ref. [52]. We will explicitly indicate when a given section or figure is taken from Ref. [52].

4.1 Introduction

4.1.1 Motivation

This subsection contains text from Ref. [52].

Measuring the temperature of a physical system is a fundamental task in science and technology. At the micro- and nanoscale in particular, highly precise temperature measurements are essential for a large number of current experiments. Examples include real-time monitoring of temperature profiles within living organisms e.g., utilizing colour centers in nanodiamonds [8, 9, 10], the preparation of ultracold atoms in optical lattices, as well as mapping thermodynamic phase diagrams and exploring transport phenomena [11, 16, 18, 19, 14, 13, 20], and studies of quantum thermodynamic phenomena in microelectronic devices [25, 21, 22, 26, 23]. Temperature is not a directly measurable property of a system, and in contrast to e.g. interferometry, phase estimation, or electromagnetic-field sensing [4, 77], thermometry is further complicated by the fact that temperature is also not a Hamiltonian-encoded parameter. Rather, the temperature of a system is an entropic quantity which must be estimated indirectly from the statistical behaviour of a variable that can be observed directly. The purpose of the theory of quantum thermometry is both to guide the design of optimal measurement processes, i.e., building good thermometers in the quantum regime, and to optimally infer from the acquired measurement data the underlying temperature [45, 46].

The majority of previous works on quantum thermometric theory, with the notable exception of the recent studies [78, 79, 80], have focused on local point estimation [56] – termed for short the local paradigm – in which measurements are designed to detect small variations around a known temperature value [45, 46]. Within the local paradigm, the expected precision of a temperature estimate is typically quantified by the frequentist mean-square error, with the associated signal-to-noise ratio providing a meaningful notion of relative error [6]. Given that certain conditions are satisfied, e.g., that the temperature estimate is unbiased, the frequentist mean-square error is lower bounded,

and typically well approximated, by the so-called Cramér-Rao bound [62, 70, 48]. Furthermore, optimal measurements applicable in the asymptotic (large data set) regime can be identified via the local optimization of the Cramér-Rao bound.

Motivation for constructing a theory applicable beyond the local paradigm is twofold: (i) it is typically an unjustified assumption that the temperature to be estimated is known with sufficient precision a priori to justify working within the local paradigm, and (ii) the optimal measurement protocol generally depends on the prior temperature information, and cannot be identified via an optimization of the “asymptotic” Cramér-Rao bound. Avoiding the restrictions of the local paradigm, i.e., providing a general approach to quantifying thermometric performance, and designing optimal measurements, under conditions of non-negligible prior uncertainty, requires a Bayesian framework [61, 55].

In this chapter we study a theory of Bayesian quantum thermometry, applicable for any amount of prior information, which is based on the concept of a thermodynamic length [65, 66, 49, 67]. The basic idea is that a meaningful measure of thermometric precision should be based on the ability to distinguish states at different temperatures, i.e., colder from hotter, and should be independent of the particular parameterization of the states, e.g., temperature. This can be naturally achieved by introducing a distance function between the thermal states of the sample system considered. Such a distance is exactly the thermodynamic length between thermal states [65, 66, 49, 67]. An interesting implication of the proposed framework is that any meaningful definition of relative error, must be given with respect to the specific sample system considered. In particular, we find that the standard signal-to-noise ratio, defined in terms of the frequentist mean-square error, is only recovered as a meaningful relative error – within the local regime – when the considered sample system can be effectively modelled as an ideal heat bath.

4.1.2 Outline of chapter

The structure of the chapter is as follows: In Sec. 4.2 we introduce the canonical thermal state, derive the associated quantum Fisher information, and describe a possible statistical interpretation of the thermodynamic length. In Sec. 4.3 we consider perhaps the most common scenario of quantum thermometry [45, 46], namely temperature estimation of an ideal heat bath. We argue for a logarithmic form of the thermodynamic length, and thus connects our approach with previous results [81]. In addition we point out that an ideal heat bath is a suitable description of a non-interacting Bose-gas well above the critical temperature for condensation. In Sec. 4.4 we consider a specific example involving thermal spin-1/2 particles. In particular we compare two scenarios: one in which we are interested in thermometry of the thermal particles themselves, and one in which the particles are employed as equilibrium probes of an underlying heat bath. In Sec. 4.5 we first describe the relation between our approach to thermometry, based on the thermodynamic length, and the standard approach in which temperature differences are gauged by an absolute difference. We then illustrate the difference via two examples: adaptive probe thermometry of a Bose-Einstein condensate and local thermometry of a fermionic tight-binding chain. In particular we focus on the regime of low temperature.

4.2 Canonical thermal states

In a typical thermometry scenario, the sample system is associated with a Hamiltonian operator \hat{H} , and is taken to be in a thermodynamic equilibrium state [46, 45, 82]. The question of why a thermal equilibrium state is a suitable description was discussed in Sec. 1.2. We will not pursue the topic further here. Rather, the physical relevance of the family of thermodynamic equilibrium states is taken as a phenomenological fact. To be specific, we adopt the temperature parameterization $\Theta = \mathbb{R}_+$, and take the quantum state of the sample to be an element of the manifold of canonical thermal states

$$\hat{\rho}(\theta) = \frac{\exp\left(-\hat{H}/k_B\theta\right)}{\mathcal{Z}(\theta)}, \quad (4.1)$$

where k_B is the Boltzmann constant, and $\mathcal{Z}(\theta) = \text{Tr}\left[\exp\left(-\hat{H}/k_B\theta\right)\right]$ is the canonical partition function. In this section we will not indicate the choice of parameterization explicitly. This is because we exclusively consider the temperature parameterization. Note that the present setting is sufficiently general to account for a temperature-independent chemical potential. Thus, we can consider the replacement $\hat{H} \rightarrow \hat{H} - \mu\hat{N}$, where μ is the chemical potential and \hat{N} is the particle-number operator. The inclusion of the chemical potential term brings us into the manifold of grand-canonical thermal states. In a more general estimation setting, we could make it explicit that the Hamiltonian operator is generically a function of a number of parameters, e.g. coupling constants, and consider the joint estimation of the temperature and the Hamiltonian parameters. Here, however, we focus exclusively on the temperature, and take the sample Hamiltonian to be perfectly known.

4.2.1 Quantum Fisher information metric

Given the manifold of canonical thermal states, the first task we consider is to evaluate the associated QFI metric, which will make it possible to construct a metric-based distance function. When considering equilibrium thermodynamic states the metric-based distance is usually called the thermodynamic length [49, 67]. If we first compute the temperature derivative of the canonical thermal state, and utilize the fact that the Hamiltonian operator commutes with the state, then we find the relation

$$\partial_\theta \hat{\rho}(\theta) = \frac{\hat{H} - E(\theta; \hat{H})}{2k_B\theta^2} \hat{\rho}(\theta) + \hat{\rho}(\theta) \frac{\hat{H} - E(\theta; \hat{H})}{2k_B\theta^2}, \quad (4.2)$$

where for convenience we have defined the thermal energy by

$$E(\theta; \hat{H}) := \text{Tr}\left[\hat{\rho}(\theta)\hat{H}\right]. \quad (4.3)$$

If we recall the definition of the symmetric logarithmic derivative (SLD) in Eq. (2.37), then the above expression for the derivative allows us to straightforwardly identify the SLD associated with the canonical thermal state

$$\hat{\mathcal{L}}(\theta) = \frac{\hat{H} - E(\theta; \hat{H})}{k_B\theta^2}. \quad (4.4)$$

The measurement for which the associated Fisher information is given by the QFI, is a projective measurement of the eigenstates of the SLD. Based on the obtained expression for the SLD, we see that this corresponds to projectively measuring the energy of the sample system, i.e. the Hamiltonian eigenstates. Since the Hamiltonian operator is independent of the temperature, the projective energy measurement constitutes a valid, temperature independent, reference measurement. For a projective measurement of the sample-system energy, the associated QFI takes the form [83, 84]

$$\mathcal{F}_{\mathcal{L}}(\theta) = \text{Tr} \left[\hat{\mathcal{L}}(\theta)^2 \hat{\rho}(\theta) \right] = \frac{C(\theta; \hat{H})}{k_B \theta^2}, \quad (4.5)$$

where we have defined the heat capacity of the sample-system

$$C(\theta; \hat{H}) := \partial_{\theta} E(\theta; \hat{H}). \quad (4.6)$$

The second equality in the QFI above follows from substituting Eq. (4.4) for the SLD and a few straightforward manipulations. Here we express the heat capacity as the temperature derivative of the thermal energy, however, direct computation shows that it could equivalently be given in terms of the energy variance of the canonical thermal state. This property reveals that the best possible local distinguishability of canonical thermal states is intimately connected to the presence of fluctuations in the statistics associated with energy measurements.

4.2.2 Statistical interpretation of thermodynamic length

Given the QFI metric, we recall from Eq. (2.35) that the thermodynamic length between canonical thermal states at temperatures θ_0 and θ_1 takes the form

$$\begin{aligned} \mathcal{D}(\theta_0, \theta_1) &= \int_{\theta_0}^{\theta_1} d\theta \sqrt{\mathcal{F}_{\mathcal{L}}(\theta)}, \\ &= \int_{\theta_0}^{\theta_1} d\theta \sqrt{\frac{C(\theta; \hat{H})}{k_B \theta^2}}. \end{aligned} \quad (4.7)$$

We now understand this as a measure of the number of thermal fluctuations in the sample energy along the traversed path $\theta_0 \rightarrow \theta_1$ [49]. Note that the thermodynamic length is a dimensionless quantity, or in other words, it is a pure number. Before proceeding we should pause and ask if we can attach an operational meaning to this number, i.e. what does it mean to say that two canonical thermal states are separated by a given thermodynamic length.

We can gain some intuition about the statistical meaning of the thermodynamic length by considering the distinguishability of two probability distributions. Consider a parametric set of probability distributions $\mathcal{P} = \{\mathcal{P}_{\lambda} \text{ for } \lambda \in \Lambda\}$ on the data space \mathcal{X} associated with the stochastic variable X . Here we write \mathcal{P}_{λ} to refer to the full likelihood measure $d\mathbb{P}(x|\lambda)$ for all $x \in \mathcal{X}$ given the parameter value $\lambda \in \Lambda$. Furthermore, we shift to the Euclidean parameterization to simplify our analysis. If n measurement outcomes are sampled according to one of these probability distributions, say \mathcal{P}_{λ} , then the probability $\mathbb{P}_{\text{confusion}}$ that it will look as if the data was sampled according to a different distribution $\mathcal{P}_{\lambda'}$, will decrease exponentially in n [85, 86]

$$\mathbb{P}_{\text{confusion}}(\lambda, \lambda') = \exp[-n\mathcal{S}(\mathcal{P}_{\lambda'} \|\mathcal{P}_{\lambda})], \quad (4.8)$$

where we have introduced the relative entropy, or Kullback–Leibler divergence [86]

$$\mathcal{S}(\mathcal{P}_{\lambda'} \parallel \mathcal{P}_{\lambda}) := \int_{\mathcal{X}} d\mathbb{P}(x|\lambda') \log \frac{d\mathbb{P}(x|\lambda')}{d\mathbb{P}(x|\lambda)}, \quad (4.9)$$

between the considered probability distributions. Recall that $d\mathbb{P}(x|\lambda)/d\mathbb{P}(x|\lambda')$ denotes the Radon-Nikodym derivative [35], and we could equivalently write this as the ratio of PDFs. The above result holds to first order in the exponent, i.e. when the relative entropy is not too large [86]. The quantity $\mathbb{P}_{\text{confusion}}$ should be interpreted as the probability that an experimenter will accept a hypothesized distribution $\mathcal{P}_{\lambda'}$ given that the true distribution is \mathcal{P}_{λ} after n measurements.

In the case of thermometry, we know the measurements that optimally resolve temperature differences are projective measurements of the sample energy. We thus take the stochastic variable X to be the sample energy. Note, however, that the analysis below is valid for any arbitrary measurement. Consider a discrete ordered sequence of Euclidean parameter values $\lambda_j = \lambda_0 + j\delta\lambda$, where we have introduced the increment $\delta\lambda = (\lambda_1 - \lambda_0)/N$, and an index variable $j = 0, 1, \dots, N$ for some large number N . Suppose a hypothesis test consisting of νN measurements is performed at each of these increments, where ν is a positive integer, and define the total probability of confusion $\tilde{\mathbb{P}}_{\text{confusion}}$ as the probability that all of the intermediate hypothesized distributions are accepted

$$\begin{aligned} \tilde{\mathbb{P}}_{\text{confusion}}(\lambda_0, \lambda_1) &:= \prod_{j=0}^{N-1} \mathbb{P}_{\text{confusion}}(\lambda_j, \lambda_{j+1}) \\ &= \exp \left[-\nu N \sum_{j=0}^{N-1} \mathcal{S}(\mathcal{P}_{\lambda_{j+1}} \parallel \mathcal{P}_{\lambda_j}) \right], \end{aligned} \quad (4.10)$$

where the second equality follows from substituting Eq. (4.8). The quantity $\tilde{\mathbb{P}}_{\text{confusion}}$ is a measure of how likely we are to accept the string of intermediate hypothesized distributions connecting one parameter value with another, given N tests and νN measurements at each test. For a vanishing increment, $N \rightarrow \infty$, it can be shown that the relative entropy takes the form [63]

$$\mathcal{S}(\mathcal{P}_{\lambda_{j+1}} \parallel \mathcal{P}_{\lambda_j}) = \frac{\delta\lambda^2}{2}, \quad \text{as } N \rightarrow \infty. \quad (4.11)$$

If we substitute this expression into the total probability of confusion, make use of the fact that $N\delta\lambda$ is simply the thermodynamic length, and shift into the temperature parameterization, then we obtain a lower bound on the total probability of confusion in terms of the squared thermodynamic length

$$\tilde{\mathbb{P}}_{\text{confusion}}(\theta_0, \theta_1) = \exp \left[-\frac{\nu}{2} \mathcal{D}(\theta_0, \theta_1)^2 \right], \quad (4.12)$$

which holds when N is sufficiently large that Eq. (4.11) is valid. Hence, the total probability of confusion is given by a quantity vanishing exponentially in the squared thermodynamic length between the endpoint temperatures. This expression is sufficient if we want to have a heuristic interpretation of the statistical meaning of the thermodynamic length. Given a certain value of the thermodynamic length we would need to perform on the order of $\mathcal{D}(\theta_0, \theta_1)^{-2}$ hypothesis test ν , in order to allow for a total probability of confusion significantly smaller than unity.

4.3 The thermalizing channel

This section contains text and figures from Ref. [52]. The text contains minor edits to conform to notation.

A common scenario in quantum thermometry, is that of a quantum probing-system subject to a thermalizing channel, i.e. a completely-positive and trace-preserving map taking any initial state of the probing-system to a thermal state at the channel temperature. A thermalizing channel can in general be modelled as induced by a sample-system that is effectively an infinitely large heat reservoir. Here we model such an ideal heat reservoir by a heat capacity that, either approximately or by definition, equals a constant value, denoted $k_B\mathcal{V}$, across the range of relevant temperatures. In this case the sample energy is a linear function of the temperature, and we can evaluate the thermodynamic length analytically to find:

$$\mathcal{D}(\theta_0, \theta_1) = |\log(\theta_1/\theta_0)|, \quad (4.13)$$

where for convenience we set $\mathcal{V} = 1$. The form of the mean-square error (MSE) resulting from this specific distance function is called the mean-square logarithmic error (MSLE). The MSLE can be adopted whenever it can be assumed that the manifold of thermal states is generated via a weak coupling to an infinite heat reservoir. In practice this assumption might break down at low temperatures, and more fundamentally there are cases where thermal behaviour cannot be linked to an infinite heat reservoir, e.g. sub-system thermalization described by the eigenstate thermalization hypothesis [87, 43].

The MSLE was recently proposed by Rubio et al. [78] as a suitable credibility measure, in the special case of a reference measurement for which the associated likelihood function satisfies the *scale-invariance* property [53, 61]:

$$p_{\mathcal{Y}}(y|\theta) = \frac{g(y/k_B\theta)}{\int_{\mathcal{Y}} dy' g(y'/k_B\theta)}, \quad (4.14)$$

where y denotes the outcome from a projective measurement of the sample energy and $g(y/k_B\theta)$ is a function only of the dimensionless ratio $y/k_B\theta$. This scale-invariance property of the likelihood function is satisfied for sample-systems with a constant density of states, or equivalently, a constant heat capacity.

When considering the MSLE, it follows that the reference QFI takes the form $\mathcal{F}_{\mathcal{L}}(\theta) = \theta^{-2}$, and we then find that the associated asymptotic $\overline{\text{CRB}}$ (Eq. (3.33)), for n measurements represented by the set of POVMs Π_n , is given by

$$\overline{\text{CRB}}(n) \rightarrow \int_{\Theta} d\theta \frac{p(\theta)}{\theta^2 \mathcal{F}_{\Pi_n}(\theta)}, \quad \text{as } n \rightarrow \infty. \quad (4.15)$$

The quantity $\theta^2 \mathcal{F}_{\Pi_n}(\theta)$ provides an upper bound on the signal-to-noise ratio within the frequentist estimation paradigm [51, 88]. This shows that when the sample can be modelled as an ideal heat reservoir, the standard notion of relative error is recovered in the local limit where $p(\theta)$ is sharply peaked. However, our analysis also points out that the standard relative error is not suitable unless the sample-system has an approximately constant heat capacity across the range of relevant temperature. This condition typically breaks down at sufficiently low temperatures [51].

A specific system approximately realizing the above assumptions on an ideal heat reservoir is a Bose gas at a fixed density well above the critical temperature [51]. As an

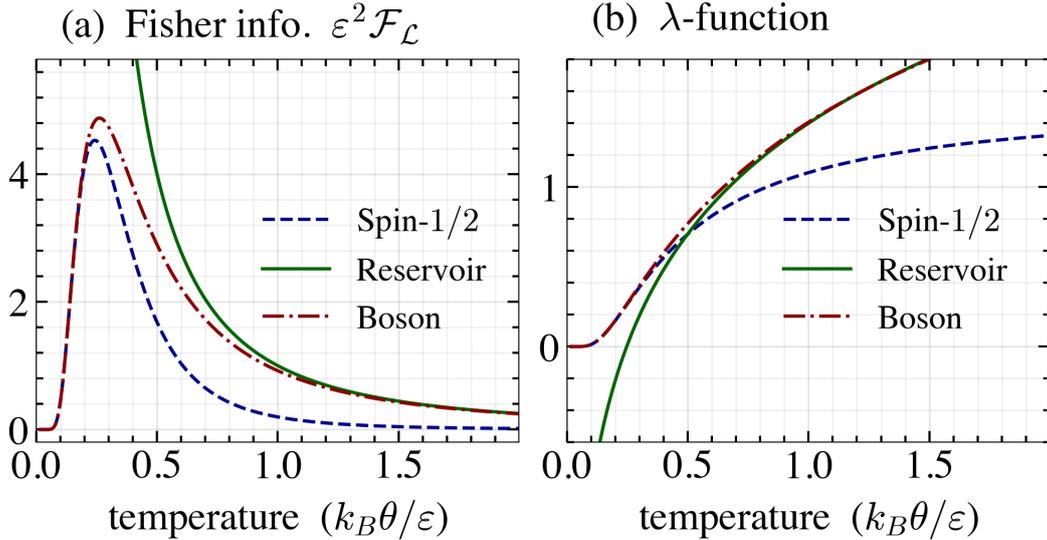


Figure 4.1: *This figure and caption is adopted from Ref. [52].* (a) Plot of the QFI for a spin-1/2 particle (dashed blue line), a bosonic mode (dashed-dotted red line), and for the ideal heat reservoir (solid green line). (b) Plot of the λ -functions (Eq. (2.39)) for a spin-1/2 particle (dashed blue line), a bosonic mode (dashed-dotted red line), and for the ideal heat reservoir (solid green line).

illustration we consider a gas of bosonic modes with energy gap ε . In this case the QFI per mode is given by [45]

$$\mathcal{F}_{\mathcal{L},\text{boson}}(\theta) = \frac{\varepsilon^2/k_B^2\theta^4}{4 \sinh^2(\varepsilon/2k_B\theta)}. \quad (4.16)$$

For this QFI the associated λ -function (see (2.39)) can be given an analytic expression

$$\lambda_{\text{boson}}(\theta) = -\log [\tanh(\varepsilon/4k_B\theta)], \quad (4.17)$$

which implies that $\Lambda_{\mathcal{L},\text{boson}} = [0, +\infty)$. The bosonic QFI and the associated λ -function are shown in figure 4.1 (a) and figure 4.1 (b), together with the corresponding quantities for an ideal heat reservoir. We observe that in the limit where the temperature is large compared to the boson energy gap, the bosonic modes approximate the ideal heat reservoir. This suggests that we can generically represent an ideal thermalizing channel by a collection of low-frequency bosonic modes.

4.4 Non-interacting spin one-half particles

This section contains text and figures from Ref. [52]. The text and figure contains minor edits to conform to notation.

We now consider N non-interacting spin-1/2 particles, or qubits, with identical energy gaps ε . The spin-1/2 particles are in a canonical thermal state, and as above we take the θ parameterization to be the temperature. We are going to compare and contrast two scenarios: (i) the spin-1/2 particles are employed as equilibrium thermometers of

an underlying heat reservoir, e.g. the particles could model impurities within an ultra-cold gas, mapping motional information of gas atoms onto the quantum-spin state [20, 89], and (ii) the spin-1/2 particles themselves constitute the sample-system of interest, e.g., we are interested in thermometry of the spin degrees of freedom of the ultracold atoms themselves [14, 90].

In case (i), as argued above, the MSLE is the suitable measure of confidence, and the QFI metric takes the form $\mathcal{F}_{\mathcal{L},\text{res.}}(\theta) = \theta^{-2}$. In case (ii) we take the MSE resulting from adopting the QFI metric of a thermal spin-1/2 particle as our reference – we denote this MSD in this section¹ — it takes the form [51]

$$\mathcal{F}_{\mathcal{L},\text{spin}}(\theta) = \frac{\varepsilon^2/k_B^2\theta^4}{4 \cosh^2(\varepsilon/2k_B\theta)}, \quad (4.18)$$

which, we recall, corresponds to a projective energy measurement of a single spin-1/2 particle. When referring to the MSD below, we refer to the spin-1/2 particle QFI metric. For the spin-1/2 reference, the inverse derivative function $\lambda(\theta)$ can be obtained analytically, it takes the form

$$\lambda_{\text{spin}}(\theta) = \pi - 2 \arctan(e^{\varepsilon/2k_B\theta}), \quad (4.19)$$

and thus $\Lambda_{\mathcal{L},\text{spin}} = [0, \pi/2]$. In figure 4.1(a) we plot the spin-1/2 QFI and compare it with the heat reservoir reference, and in figure 4.1(b) we plot the associated λ -functions. We observe that for temperature around $k_B\theta/\varepsilon \approx 0.3$ the two λ -functions exhibit a similar gradient, however away from this temperature regime the specific thermometric scenario considered plays a role.

The estimation strategy employed consists of projectively measuring the energy of a subset of μ particles. For thermalized spin-1/2 particles, projective energy measurements maximize the associated FI for all temperatures [83], and it follows that the FI associated with this measurement can be expressed as

$$\mathcal{F}_{\text{II}}(\theta) = \mu \mathcal{F}_{\mathcal{L},\text{spin}}(\theta). \quad (4.20)$$

In figure 4.2(a), figure 4.2(b) and figure 4.2(c) we show stochastic simulations of measurement trajectories sampled according to three different true temperatures. In all cases we plot the MMSLE estimator and the MMSD estimator, and note that only negligible differences exist between these. This feature is to be expected as the respective λ -functions rapidly become approximately constant across the posterior, when this is the case the estimator is simply the maximum-a-posterior temperature, i.e., the temperature at which the posterior takes its maximum value, independently of the specific λ -function.

Although the various temperature estimates are only negligibly different, the confidence assigned to the estimates depend on the thermometric scenario. In figure 4.2(e) and 4.2(f) we show the $\overline{\text{MSLE}}$ and the $\overline{\text{MSD}}$ respectively as a function of the subset size μ , and compare these with the associated $\overline{\text{VTB}}$ and the associated asymptotic $\overline{\text{CRB}}$. The $\overline{\text{MSLE}}$ and the $\overline{\text{MSD}}$ are evaluated using the corresponding smoothed Jeffrey's priors shown in figure 4.2(d). In the case of the $\overline{\text{MSLE}}$ we do not observe convergence to the

¹This section has appeared in Ref. [52] in which the MSD label was used. Here we keep this notation, and emphasize that it refers to the specific MSE resulting from the thermodynamic length with respect to the spin-1/2 particle QFI metric.

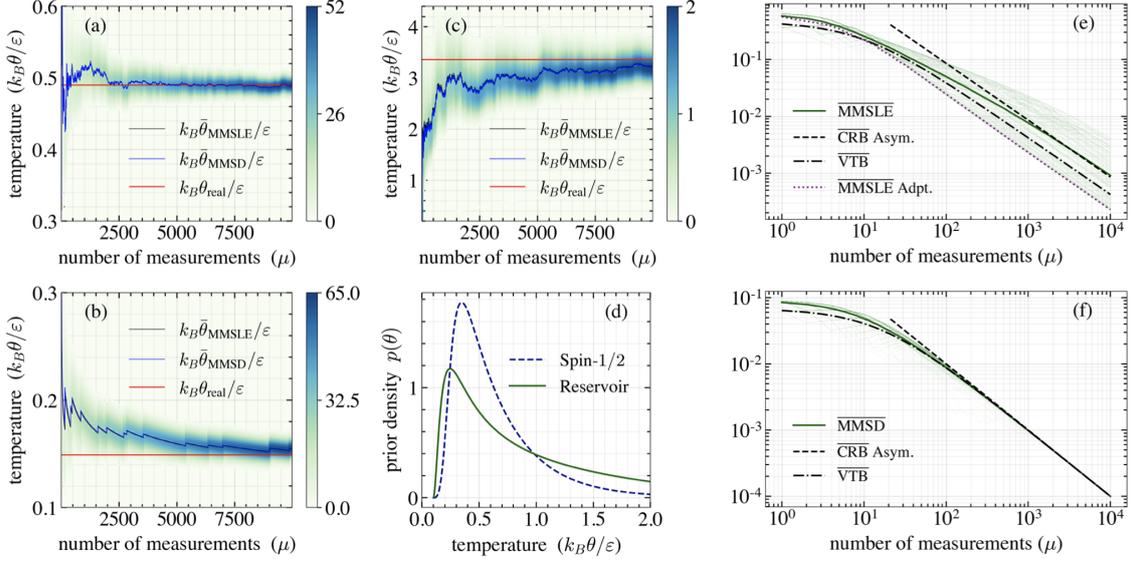


Figure 4.2: *This figure and caption is adopted from Ref. [52].* (a,b,c) Simulated stochastic measurement trajectories sampled according to a specific true temperature (red horizontal line), the posterior probability density function is shown as a colour map. Both the MMSLE and the MMSD estimators are shown. (d) The smoothed Jeffrey's prior (Eqs. (2.48) and (2.49)) for $\alpha = -2.5$ and $k_B\theta \in [\varepsilon/10, 5\varepsilon]$, corresponding to the spin-1/2 particle and the reservoir reference metric. (e) Simulation of the $\overline{\text{MSLE}}$ for thermometry employing free spin-1/2 particles. The simulation is performed by averaging over 250 stochastic trajectories where the measurement data is generated with respect to a true temperature sampled according to the prior density. The $\overline{\text{MSLE}}$ is compared with the $\overline{\text{CRB}}$, convergence is observed for $\nu \gtrsim 10^3$, and we note that the generated trajectories do not converge to the mean asymptotically. Furthermore, we plot the $\overline{\text{VTB}}$ and observe that this is not a tight bound. Lastly, we plot the adaptive $\overline{\text{MSLE}}$. This is obtained by adapting the energy gap ε in each repetition of the measurement. Here, we employ a $\overline{\text{VTB}}$ optimization strategy. (f) Simulation of the $\overline{\text{MSD}}$ for thermometry of free spin-1/2 particles. The simulation is performed by averaging over 250 stochastic trajectories where the measurement data is generated with respect to a true temperature sampled according to the prior density. The $\overline{\text{MSD}}$ is compared with the $\overline{\text{VTB}}$ and convergence is observed for $\nu \gtrsim 10^2$, furthermore we observe that the generated trajectories converge to the average.

$\overline{\text{VTB}}$. This is to be expected since $\mathcal{F}_{\mathcal{L},\text{res.}}/\mathcal{F}_{\Pi}$ is generally not constant across the domain of the employed prior. This feature also means that we do not observe a convergence to be asymptotic $\overline{\text{CRB}}$ at the trajectory level, i.e. we observe fluctuations around the average. For the $\overline{\text{MSD}}$ we observe rapid convergence to the $\overline{\text{VTB}}$, and trajectory level convergence to the asymptotic $\overline{\text{CRB}}$. This is due to the fact that $h_{\mathcal{L},\text{spin}}/h_{\Pi} = 1/\mu$ is a constant independent of the temperature.

When considering the spin-1/2 particles as thermometers of a underlying heat reservoir, it is sensible to consider adaptation of the energy gap ε to optimize the thermometric performance. That is we consider a protocol consisting of thermalizing a spin-1/2 particle and projectively measuring the energy. Based on the observed outcome, the

energy gap of the particle employed in the following measurement is adapted. In figure 4.2(e) we also show the adaptive $\overline{\text{MSLE}}$, which is obtained via a $\overline{\text{VTB}}$ optimization strategy of the spin-1/2 energy gap ε in each iteration of the protocol. Recall that by a $\overline{\text{VTB}}$ optimization strategy we refer to a minimization of the $\overline{\text{VTB}}$ with respect to the energy gap ε in each iteration. In this case we observe a rapid convergence to the optimized local result, i.e., the adaptive $\overline{\text{MSLE}} \sim \max_{\theta}[\mathcal{F}_{\mathcal{L},\text{res.}}/\mathcal{F}_{\Pi}]$. Furthermore, the adaptive $\overline{\text{MSLE}}$ converges to a lower bound on the set of MSLE trajectories. This is to be expected as the energy gap can be tuned to the optimal ratio of the energy gap to the temperature sampled from the prior.

4.5 Difficulty of thermometry at low temperatures

In this section we present two further examples illustrating the application of the developed framework for thermometry. In particular, we study the regime of low temperatures. First we contrast the conventional approach to temperature estimation with our approach based on the thermodynamic length. We then study two examples: 1) Probe thermometry of a Bose-Einstein condensate and 2) local thermometry of a fermionic tight-binding chain.

4.5.1 Relation to conventional thermometry

In this section we discuss the relation between our approach to temperature estimation, based on the thermodynamic length, and the more conventional approach in which the distance between temperatures is gauged by the absolute difference. Our main focus is on the regime of low temperatures. In the literature on quantum thermometry [45, 46] it is common to analyse the performance of an estimation protocol in terms of the frequentist mean-squared error attributed to a computed temperature estimate

$$\Delta^2 \tilde{\theta}_n(\theta) := \int_{\mathbf{x}_n} d\mathbb{P}_0(\mathbf{x}_n) p(\mathbf{x}_n|\theta) (\tilde{\theta}(\mathbf{x}_n) - \theta)^2. \quad (4.21)$$

In what follows we will refer to this as the absolute error on a temperature estimate [51]. This quantity was discussed in Sec. 2.3, where we considered n independent repetitions of a measurement Π . Evaluating the absolute error requires specifying a temperature estimator. Regardless of the choice of estimator, however, the resulting error satisfies the frequentist Cramér-Rao bound

$$\Delta^2 \tilde{\theta}_n(\theta) \geq \frac{1}{n\mathcal{F}_{\Pi}(\theta)} + O(\langle \delta_n^3 \rangle), \quad (4.22)$$

where it should always be stressed that this lower bound is valid in the asymptotic limit, i.e. for sufficiently large n . If we adopt a maximal-likelihood estimator for the temperature, then the bound is tight in the asymptotic limit [6, 54]. This feature implies that the performance of a temperature-estimation protocol can typically be analysed by studying the associated Fisher information [45, 46].

In many cases the Fisher information associated with a given measurement is found to vanish as the temperature approaches absolute zero [45, 83], indicating that in this regime the observed measurement statistics become increasingly insensitive to the exact temperature. It is thus tempting to conclude that thermometry becomes increasingly

challenging in the low-temperature regime, and some authors even speak of a divergence of the absolute error [51, 88, 91]. A statement which should be understood in reference to the frequentist Cramér-Rao bound. Taken literally, such a statement is nonsensical. A more careful articulation of the intended message is as follows: As the temperature tends to absolute zero, we observe a divergence in the number of independent measurements, which must be performed for the frequentist Cramér-Rao bound to be valid, i.e. for higher-order correction terms to be negligible. In other words: As the observed measurement statistics become increasingly insensitive to the exact value of the temperature, there is a corresponding growth in the amount of measurement data which must be collected to enter the local regime of estimation. For brevity, and with the more careful statement implied, we will also speak of a diverging absolute error in what follows.

With this potential semantic confusion out of the way, we can focus on the problem of designing measurements retaining maximal temperature sensitivity deep into the low-temperature regime. However, in light of the previous chapters of this thesis, we must ask ourselves whether the absolute error on the temperature is a suitable measure of error. First of all it is not an invariant, and secondly its physical interpretation requires additional knowledge of the sample being measured. These concerns were exactly the motivation for considering the thermodynamic length, rather than the absolute temperature difference, when gauging the difference of two temperatures. In Sec. 2.3 we considered the estimation of an arbitrary bijective function of the temperature, and those results apply if we consider the estimation of the Euclidean parameter λ . In particular from Eq. (2.15) we obtain an invariant version of the frequentist Cramér-Rao bound:

$$\Delta^2 \tilde{\lambda}_n(\theta) \geq \frac{\mathcal{F}_{\mathcal{L}}(\theta)}{n\mathcal{F}_{\Pi}(\theta)} + O(\langle \delta_n^3 \rangle), \quad (4.23)$$

where we have used Eq. (2.39) relating the Euclidean parameter and the quantum Fisher information of the sample. Note that the bound depends directly on the specific sample-system via the Fisher information ratio. Expressing the error in terms of such a ratio is sensible, as the temperature sensitivity of a given measurement must be determined relative to the temperature sensitivity of the sample itself. For this reason, we will refer to the left-hand side above as the relative error.

In contrast to our usage here, the concept of a relative error on the temperature estimate, typically refers to the noise-to-signal ratio defined by $\theta^{-2}\Delta^2\tilde{\theta}_n(\theta)$ [45, 46]. From the frequentist Cramér-Rao bound on the absolute error, Eq. (4.22), we can lower bound the noise-to-signal ratio as:

$$\theta^{-2}\Delta^2\tilde{\theta}_n(\theta) \geq \frac{1}{n\theta^2\mathcal{F}_{\Pi}(\theta)} + \theta^{-2}O(\langle \delta_n^3 \rangle). \quad (4.24)$$

In the above we mentioned that the Fisher information associated with a given measurement typically vanishes as the temperature tends to absolute zero [45, 83], giving rise to a diverging lower-bound on the absolute error. Adopting the noise-to-signal ratio does not improve on this situation, as we merely face a more severe divergence. However, having put into question the suitability of the absolute error, we must also question the noise-to-signal ratio derived from it. A natural first question might be whether any scenario can be identified, in which the noise-to-signal ratio is recovered as a special case of the relative error. In Sec. 4.3 we discussed the problem of estimating the temperature

of an ideal heat bath, and argued that the quantum Fisher information associated with such a sample-system takes the form $\mathcal{F}_{\mathcal{L}} \propto \theta^{-2}$. Thus within this scenario, we see from Eq. (4.23) that the relative error reduces to the noise-to-signal ratio. This is encouraging as in dealing with a heat bath we found that the temperature is a scale parameter, implying that only the ratios of temperatures are meaningful, and such cases are exactly the ones in which a noise-to-signal ratio is well motivated [53, 78].

Recovering the noise-to-signal ratio in the case of temperature estimation of a heat bath, justifies our relative error as a generalized concept applicable beyond the heat bath scenario. Armed with this concept, we may return to a discussion of the challenges associated with thermometry in the low-temperature regime. The key quantity is the ratio of the quantum Fisher information to the Fisher information associated with the implemented measurement. From the definition of the quantum Fisher information it follows directly that:

$$\forall \theta \in \mathbb{R}_+, \quad \frac{\mathcal{F}_{\mathcal{L}}(\theta)}{\mathcal{F}_{\Pi}(\theta)} \geq 1, \quad (4.25)$$

where equality is attained if the sample energy is projectively measured. Hence, it is in principle possible to obtain a relative error which tends to a constant as the temperature approach absolute zero, avoiding any diverging behaviour. Achieving this requires that the Fisher information associated with the measurement, vanish no more rapidly than the quantum Fisher information of the sample itself. In the following two sections we will study two illustrative examples: 1) Probe thermometry of a non-interacting Bose gas below the critical temperature in which the relative error is found to improve as the temperature decrease, and 2) local thermometry of a Fermionic tight-binding chain in which the relative error diverge as the temperature decrease, albeit at a slower rate than the corresponding noise-to-signal ratio.

To conclude this section we point out the connection between the frequentist quantities discussed above, and the Bayesian mean-squared error. This connection is of course well-known [92]. From the definition of the Bayesian mean-squared error, Eq. (3.26), it follows directly that:

$$\overline{\text{MMSE}}(n) = \int d\theta p(\theta) \Delta^2 \bar{\lambda}_n(\theta), \quad (4.26)$$

where the relative error is now specified with respect to the MMSE estimator. Thus, the Bayesian mean-squared error is equal to the prior-averaged relative error. If we ask instead about the connection with the posterior mean-squared error of Eq. (3.3), then we mention that due to local asymptotic normality the asymptotic form of the posterior Cramér-Rao bound given in Eq. (3.24), exactly coincides with the frequentist Cramér-Rao bound of Eq. (4.23).

4.5.2 Thermometry of a Bose-Einstein condensate

As our first example we consider an adaptive thermometry protocol of a Bose-Einstein condensate (BEC). This example is interesting as it illustrates a setting in which thermometry becomes increasingly easy as the temperature decreases. The condensate consists of a large number N of non-interacting Bose particles of mass m in three dimensional space with a fixed density. As our reference measurement we consider a projective energy measurement of the full sample. This gives the quantum Fisher information metric $\mathcal{F}_{\mathcal{L}}(\theta) = C(\theta; \hat{H}_{\text{BEC}})/k_B\theta^2$, where $C(\theta; \hat{H}_{\text{BEC}})$ is the heat capacity

associated with the BEC Hamiltonian \hat{H}_{BEC} . The BEC is associated with a critical temperature θ_c for condensation at the fixed BEC density ϱ , given by [51]:

$$k_B\theta_c = \frac{2\pi\hbar^2}{m} \left(\frac{\varrho}{\zeta(3/2)} \right)^{2/3}, \quad (4.27)$$

where ζ denotes the Riemann Zeta function and \hbar is the reduced Planck constant. Below the critical temperature, the BEC heat capacity takes the form [51]:

$$C(\theta; \hat{H}_{\text{BEC}}) = \gamma k_B N \left(\frac{\theta}{\theta_c} \right)^{3/2}, \quad \text{for } \theta \leq \theta_c, \quad (4.28)$$

where $\gamma = 15\zeta(5/2)/4\zeta(3/2)$. Above the critical temperature the heat capacity rapidly tends to a constant value, and the BEC approximates an ideal heat reservoir. This case will be discussed further in the following chapter. Here we will focus on the regime below the critical temperature. From the BEC heat capacity, Eq. (4.28), we find that the inverse derivative function can be obtained analytically:

$$\lambda(\theta) = \frac{4}{3} \sqrt{\gamma k_B N} \left(\frac{\theta}{\theta_c} \right)^{3/4}, \quad \text{for } \theta \leq \theta_c. \quad (4.29)$$

We note that strongly interacting Bose gases have been cooled down to $\sim 0.02\theta_c$ [89]. Studying strongly interacting systems brings with it a host of complications, e.g. it might not be possible to analytically obtain the equation of state [16]. Therefore, here we focus on the non-interacting case, and investigate the performance of the adaptive protocols proposed in the previous chapter.

We consider a probe-based measurement, in which m interacting qubits, i.e. two-level quantum systems, are fully thermalized by the BEC. The Fisher information of the thermalized probe is maximized for a projective measurement of the probe energy [50], and can be written as $\mathcal{F}_{\Pi}(\theta) = C_{\Pi}(\theta; \mathbf{E})/k_B\theta^2$ where $\mathbf{E} = \{E_1, \dots, E_{2^m}\}$ denotes the energy eigenspectrum of the probe, and $C_{\Pi}(\theta, \mathbf{E})$ is the probe heat-capacity. The most general formulation of the optimization problem, assuming no restrictions on our design capabilities, is to optimize the measurement with respect to the eigenspectrum of the probe.

The number of energy levels grows geometrically in m , and the general optimization problem becomes intractable for moderately large m . To simplify the optimization problem we restrict the optimization space to that of locally optimal spectra, i.e. we consider the asymptotic CRB and find the energy-level structure minimizing this bound. This problem has been solved by Correa et al. [50], and the answer is that the optimal spectrum takes the form of an effective two-level system with a $2^m - 1$ degenerate excited state, and an energy gap $\hbar\omega_{\text{opt}} = x_{\text{opt}}\theta_{\text{true}}$, where x_{opt} is the solution to a transcendental equation. This will be discussed in greater detail in the following chapter. Furthermore, for this probe structure the heat capacity takes the form $C_{\Pi} = x_{\text{opt}}^2/4 - 1$, and is thus temperature independent. Given these restrictions, the free parameter to be optimized is the probe frequency ω , and we have a one-dimensional optimization problem. In restricting ourselves to this optimization space, we are assuming that the energy-level structure of the interacting probe system can be arbitrarily designed. The optimization space considered can be approximately realized for a spin-chain probe described by the longitudinal-field Ising model [80].

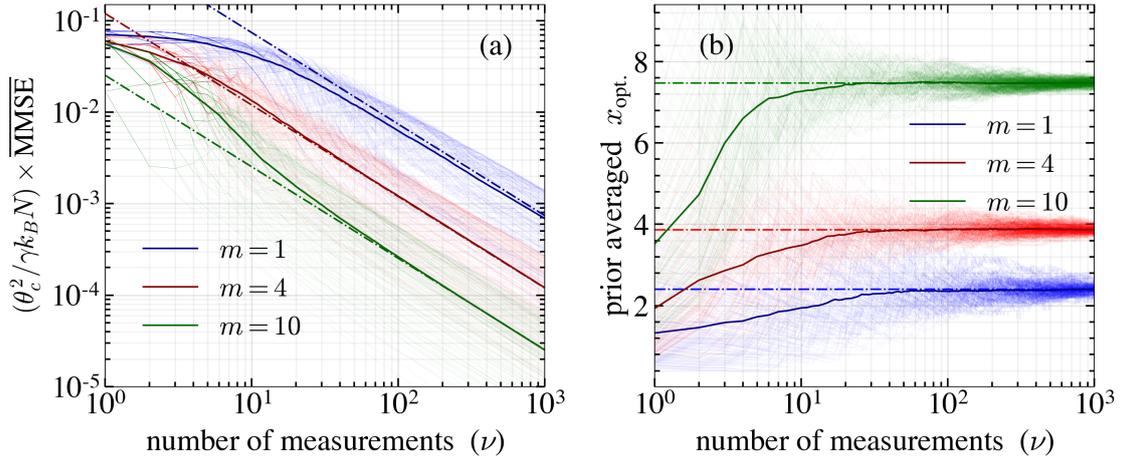


Figure 4.3: Stochastic simulation, averaged over 150 generated trajectories, of a thermalized probe measurement using $\sqrt{\text{VTB}}$ optimization of the probe frequency. The probe is an effective two-level system with a $2^m - 1$ degenerate excited state. For the performed simulations we employ the smoothed Jeffrey's prior (Eqs. (2.48) and (2.49)) with respect to the BEC QFI metric. We impose a boundary of $\theta/\theta_c \in [0.05, 0.95]$, and take $\alpha = -10$. (a) Plot of the simulated $\overline{\text{MMSE}}$ (solid lines) for probes build of m qubits. The semi-transparent lines gives the individual trajectories. The dot-dashed lines gives the asymptotic $\overline{\text{CRB}}$. (b) Plot of the simulated x_{opt} averaged over the prior. The optimal value can be obtained analytically [50], and are shown by horizontal dot-dashed lines. We observe that the $\sqrt{\text{VTB}}$ optimization strategy converges to the local optimum.

To further simplify our analysis we adopt an optimization strategy based on minimizing the van Trees bound ($\sqrt{\text{VTB}}$), or equivalently on maximizing the prior average of the quantity $\mathcal{F}_{\Pi}/\mathcal{F}_{\mathcal{L}}$. We consider ν repetitions of the probe-based measurement, and perform an optimization of the probe frequency in each repetition. Assuming that the adaptive strategy converges to the local optimum, we expect that, in the asymptotic limit, the MMSE is well approximated by the asymptotic CRB:

$$\frac{\text{CRB}(\mathbf{x}_{\nu})}{\gamma k_B N} \rightarrow \frac{1}{\nu(x_{\text{opt}}^2/4 - 1)} \left(\frac{\bar{\theta}(\mathbf{x}_{\nu})}{\theta_c} \right)^{3/2}, \quad \text{as } \nu \rightarrow \infty, \quad (4.30)$$

where \mathbf{x}_{ν} denotes the sequence of measurement outcomes, and $\bar{\theta}(\mathbf{x}_{\nu})$ is the associated MMSE estimator. We note that the CRB vanishes as $\bar{\theta}^{3/2}$ in the low-temperature limit. This indicates that thermometry becomes easier as the true temperature of the sample approaches zero. This conclusion is reached under the assumption that the probe frequency can be tuned to arbitrarily small values. In practice this is not the case, and the above result is only valid in the regime where the BEC temperature is large compared to the minimal attainable probe frequency.

In Fig. 4.3 (a) we show stochastic simulations of the adaptive thermometry protocol. We observe convergence of the $\overline{\text{MMSE}}$ to the $\overline{\text{CRB}}$. Furthermore, we see that generated trajectories do not converge to the mean, which reflects the $\bar{\theta}^{3/2}$ behaviour of the CRB. In Fig. 4.3 (b) we show stochastic simulations of the optimized probe-frequency, and observe convergence to the local optimum. Furthermore, we point out that convergence seems to require fewer adaptive steps ν as the number of qubits making up the probe in-

crease. This makes sense, as one will typically learn more from single-shot measurements based on a larger number of qubits.

4.5.3 Thermometry of a tight-binding chain

As our next example we consider thermometry of a fermionic tight-binding chain, and focus on the regime of low temperatures. This example is interesting as it illustrates a setting in which improving the attained thermometric precision becomes increasingly difficult as the temperature decreases. The chain is made up of N sites, and is found in a grand-canonical thermal state with a temperature-independent chemical potential μ . For convenience we adopt periodic boundary conditions for the chain. However, in the thermodynamic limit $N \rightarrow \infty$ this choice has negligible effect. As above, the Θ parameterization is taken to refer to the temperature of the chain. In diagonalized form the Hamiltonian of the tight-binding chain takes the simple form [51]:

$$\hat{H} = \sum_{k=1}^N \varepsilon_k \hat{C}_k^\dagger \hat{C}_k, \quad (4.31)$$

where we have defined the eigenenergies $\varepsilon_k = \varepsilon - 2t \cos(k2\pi/N)$ in terms of the on-site energy ε , which is the same for all sites, and the hopping strength t , quantifying the strength of the interaction between nearest-neighbour sites. Furthermore, \hat{C}_k (\hat{C}_k^\dagger) denotes an annihilation (creation) operator acting on the eigenmode of the chain with *quasi-momentum* k :

$$\hat{C}_k := \frac{1}{\sqrt{N}} \sum_{j=1}^N e^{2\pi i j k / N} \hat{c}_j, \quad (4.32)$$

where \hat{c}_j denotes the annihilation operator for a fermion at site j of the chain.

For simplicity, we consider the regime of half filling $\mu = \varepsilon$, and focus on the low-temperature regime $k_B\theta \ll t$. In this regime all the physics of the chain is determined by the partially filled eigenmodes with energies within $k_B\theta$ of the chemical potential μ . As our reference measurement we adopt a projective energy-measurement of the chain, and the associated metric is then the quantum Fisher information. For convenience we work with the QFI per site, which in the thermodynamic limit $N \rightarrow \infty$ takes the form [51]

$$\mathcal{F}_{\mathcal{L}}(\theta) = \frac{\pi k_B}{6t\theta}. \quad (4.33)$$

Note that this expression diverges as θ^{-1} in the low-temperature limit. This divergent behaviour holds as long as the temperature is much greater than the minimal energy gap of the chain $\Delta = 4\pi t/N$ (see Ref. [51]), and is thus a good approximation in the thermodynamic limit of $N \rightarrow \infty$. Given this QFI metric we can compute the inverse derivative function $\lambda(\theta)$ analytically:

$$\lambda(\theta) = \left(\frac{2\pi k_B \theta}{3t} \right)^{1/2}, \quad (4.34)$$

which indicates that the thermodynamic length between thermal states at different temperatures grows as the difference of the square roots of the respective temperatures. The λ -function is shown in Fig. 4.4 (d).

Following Ref. [51] we will consider a thermometry protocol employing a local measurement. In particular, the available measurement, denoted Π , is restricted to access only two neighbouring sites of the effectively infinite chain. This measurement is repeated ν times. The optimal POVM turns out to be a measurement of the occupation numbers of the modes $c_{\pm} = (c_1 \pm c_2)/\sqrt{2}$ of the accessible sites (here we adopt sites $j = 1$ and $j = 2$), and the quantum Fisher information associated with this optimal two-site measurement is given by [51]:

$$\mathcal{F}_{\Pi}(\theta) = \frac{\pi^4 k_B^2}{18(\pi^2 - 4)} \frac{(k_B \theta)^2}{t^4} + \mathcal{O}(k_B^6 \theta^4 / t^6), \quad (4.35)$$

where the $\mathcal{O}(\cdot)$ denotes higher-order terms in the temperature contributing negligibly in the low-temperature regime. Note that the Quantum Fisher information metric scales as $\sqrt{\mathcal{F}_{\mathcal{L}}(\theta)} \propto \theta^{-1/2}$, indicating that grand-canonical states of the full chain become more distinguishable as the temperature decreases. On the other hand, the Fisher information metric associated with the two-site measurement vanishes linearly in the temperature, i.e. $\sqrt{\mathcal{F}_{\Pi}(\theta)} \propto \theta$. This indicates that the two-site measurement becomes insensitive to the temperature in the low-temperature limit, even though the states of the full sample exhibit a growing sensitivity to temperature.

In the asymptotic limit $\nu \rightarrow \infty$ we expect the attained credibility to be well approximated by the asymptotic form of the CRB. Including only the leading term in temperature, the asymptotic CRB takes the form:

$$\begin{aligned} \text{CRB}(\mathbf{x}_{\nu}) &\rightarrow \frac{\mathcal{F}_{\mathcal{L}}(\bar{\theta}(\mathbf{x}_{\nu}))}{\nu \mathcal{F}_{\Pi}(\bar{\theta}(\mathbf{x}_{\nu}))}, \quad \text{as } \nu \rightarrow \infty \\ &= \frac{6(\pi^2 - 4)}{\nu \pi^3} \left(\frac{k_B \bar{\theta}(\mathbf{x}_{\nu})}{t} \right)^{-3}, \end{aligned} \quad (4.36)$$

where we recall that $\bar{\theta}(\mathbf{x}_{\nu})$ denotes the MMSE estimator resulting from the observed measurement data. Note that the CRB diverges as the third power of the temperature in the low-temperature limit. In the literature [45, 51, 88, 91], it is common to take such a diverging behaviour as a sign that the error in the attained temperature estimate diverges as the temperature tends to zero. In practice, it is never the case that the acquisition of measurement data results in a posterior distribution with a diverging variance. What is really going on is simply that the number of measurements ν we need to perform to reach the asymptotic regime diverges as the temperature tends to zero. We can obtain a rough estimate of the threshold data-size $\nu_{\text{th.}}$, by which we refer to the number of measurements that must be performed before we expect the CRB to be relevant, by comparing the $\overline{\text{CRB}}$ with the inverse prior information $\mathcal{Q}_{\text{prior}}^{-1}$. In particular, if we require equality of these two quantities we find the threshold:

$$\nu_{\text{th.}} := \mathcal{Q}_{\text{prior}} \int_{\Theta} d\theta p(\theta) \frac{\mathcal{F}_{\mathcal{L}}(\theta)}{\mathcal{F}_{\Pi}(\theta)}, \quad (4.37)$$

which is a functional of the initial prior PDF. Roughly speaking, this threshold reflects the number of measurements that must be performed before the asymptotic CRB is expected to be smaller than the CRB, when computed with respect to the initial prior PDF.

In Fig. 4.4(a,b,c) we show stochastic simulations of the posterior distribution and the MMSE estimator, resulting from optimal two-site measurements of the tight-binding

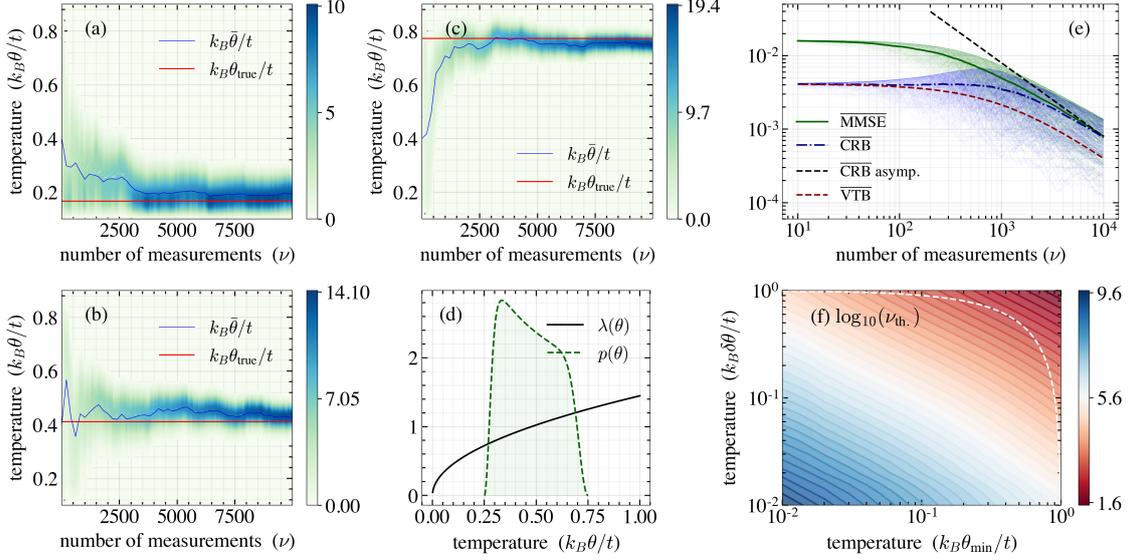


Figure 4.4: Thermometry of the tight-binding chain: (a,b,c) Simulated stochastic measurement trajectories sampled according to a specific true temperature (red horizontal line), the posterior probability density function is shown as a colour map, and the posterior mean is shown as the blue solid line. (d) Plot of the λ -function associated with the QFI of the full tight-binding chain, and of the smoothed Jeffrey's prior PDF (Eqs. (2.48) and (2.49)) employed for the performed simulations. The prior is taken to have support on the interval $k_B\theta/t \in [0.25, 0.75]$, and we take $\alpha = -10$. (e) Simulation of the $\overline{\text{MMSE}}$ via optimal measurements of two neighbouring sites. The simulation is performed by averaging over 250 stochastic trajectories where the measurement data is generated with respect to a true temperature sampled according to the prior density. The $\overline{\text{MMSE}}$ is compared with the $\overline{\text{CRB}}$ and its asymptotic approximation, convergence is observed for $\nu \gtrsim 3 \times 10^3$, and we note that convergence to the mean is not observed at the trajectory level (both MMSE and CRB trajectories are shown). Furthermore, we plot the $\overline{\text{VTB}}$ and observe that this bound is not even asymptotically tight. (f) Plot of the logarithm base 10 of the threshold data-sample-size (Eq. (4.37)) for $\alpha = -10$, and $\delta\theta := \theta_{\max} - \theta_{\min}$. The white dashed line indicates the limit $\theta_{\min} + \delta\theta = 1$.

chain at three different true temperatures. Both in terms of the convergence rate, and the attained posterior variance, thermometry becomes increasingly challenging as the true temperature decreases. This is to be expected as the temperature sensitivity of the two-site measurement, when quantified by the associated Fisher information, vanishes quadratically in the temperature. The initial prior PDF used for the simulations is shown in Fig. 4.4(d). In Fig. 4.4(e) we show stochastic simulations of the $\overline{\text{MMSE}}$ resulting from the optimal two-site measurement, and compare it with the $\overline{\text{CRB}}$, the asymptotic $\overline{\text{CRB}}$ and the $\overline{\text{VTB}}$. For the prior employed, we find that the $\overline{\text{VTB}}$ is not tight. This is to be expected since the quantity $\mathcal{F}_{\mathcal{L}}/\mathcal{F}_{\Pi}$ exhibits a θ^{-3} dependence on temperature across the domain of the prior. For the $\overline{\text{VTB}}$ to be tight, this should be a constant function. In Fig. 4.4(f) we show a contour plot of the threshold data-size ν_{th} as the prior domain is varied. Recall that this threshold is a measure of the number of measurements that must be implemented for the acquired data to impact our temperature estimate. We

observe that the threshold increases as the interval decreases ($\delta\theta \rightarrow 0$). This is intuitive. The better prior information, the more data is required to improve our information further. We also observe that as the lower boundary of the prior interval decreases, the threshold increase. This is a result of the Fisher information vanishing. As the temperature decreases the amount of extractable information decreases.

4.6 Concluding remarks

In this chapter we have explored quantum thermometry using an approach based on the thermodynamic length. In particular, we have highlighted the importance of taking into account the specific sample system under study. The key quantity, when it comes to the capacity, of a given measurement strategy Π , to distinguish thermal states of the sample being measured, is the associated Fisher information. In previous studies on quantum thermometry [83, 51, 91, 45, 46, 84], it was claimed that the quantum Fisher information associated with the sample provides an ultimate lower bound on the attainable temperature estimation precision via the frequentist Cramér-Rao bound on the absolute error. If we employ an unbiased estimator, $\bar{\theta}$, the bound is tight [6], and can be expressed as

$$\begin{aligned} n\Delta^2\bar{\theta}_n(\theta) &= 1/\mathcal{F}_\Pi(\theta) + O(\langle n\delta_n^3 \rangle) \\ &\geq 1/\mathcal{F}_\mathcal{L}(\theta) + O(\langle n\delta_n^3 \rangle). \end{aligned} \quad (4.38)$$

Based on the behaviour of the Fisher information, we can judge whether a given measurement strategy constitutes a good method for thermometry. In many cases it is found that the Fisher information vanish as the temperature tends to absolute zero. This feature has lead to the conclusion that low-temperature thermometry becomes increasingly difficult as we move into the ultracold regime. In this chapter we pointed out why a focus on the absolute error might be misguided. Instead it is suggested that the suitable measure of the attainable precision is the generalized relative error

$$\begin{aligned} n\Delta^2\bar{\lambda}_n(\theta) &= \mathcal{F}_\mathcal{L}(\theta)/\mathcal{F}_\Pi(\theta) + O(\langle n\delta_n^3 \rangle) \\ &\geq 1 + O(\langle n\delta_n^3 \rangle). \end{aligned} \quad (4.39)$$

As mentioned in the main text: Expressing the error in terms of the Fisher information ratio is sensible, as the temperature sensitivity of a given measurement strategy must be determined relative to the temperature sensitivity of the sample itself. In particular, it might be the case that the Fisher information \mathcal{F}_Π vanish as the temperature tends to zero. However, this would only lead to a “diverging” relative error, if the quantum Fisher information $\mathcal{F}_\mathcal{L}$ of the sample itself, vanish at a slower rate than \mathcal{F}_Π .

Conversely, the model under study might exhibit a classical phase transition at a certain critical temperature θ_c . In this case the quantum Fisher information $\mathcal{F}_\mathcal{L}$ might exhibit diverging behaviour around the critical temperature. When basing the analysis on the absolute error, a large QFI would signify a temperature which could be easily determined. When instead considering the relative error, we see that a large QFI might be associated with a large relative error, unless a measurement strategy can be devised exhibiting a similar diverging behaviour around the critical temperature.

In the next chapter, we take up the subject of attaining the fundamental bounds to precision via adaptive strategies, in the specific case of equilibrium probe thermometry. We then return to the problem of exploring the fundamental precision bounds in chapters 6 and 7.

Chapter 5

Fundamental Limits in Bayesian Thermometry and Attainability via Adaptive Strategies

This chapter is composed almost entirely of text and figures from Ref. [93]. In the appendix section, Sec. 5.6.1, Sec. 5.6.5 and Fig. 5.4 are new. Everything else has appeared in Ref. [93].

The chapter is self-contained. The notation does not correspond to the conventions of the previous chapters in all cases. In particular, what was called Bayesian credibility measures in previous chapters, is called expected credibility measures here. Furthermore, no mentioning is made of the reference measure, and we simply write $d\mathbb{P}_0(x) = dx$.

5.1 Introduction

Preparing quantum systems at low temperatures is an essential task for development of quantum technologies [94, 95, 96]. Measuring temperature precisely is necessary to validate cooling and ensure the performance of quantum protocols, and has been demonstrated in cutting-edge experiments [97, 20, 98, 89, 99, 100, 17, 101, 102]. However, measuring the temperature of quantum systems is often challenging. On the one hand, due to the scarcity of thermal fluctuations at such low temperatures, the relative statistical error on thermometry, i.e. the signal-to-noise ratio, can be exceptionally notable. On the other hand, the fragility of quantum systems, requires additional forward planning in order to minimise disturbance while maximising the information obtained. The theory of quantum thermometry is built to address this pivotal task [19, 46].

When quantum systems are used as thermometers or probes, they offer resources such as entanglement, coherence, and many-body interactions, that can improve precision [103, 50, 83, 104, 105, 106, 91, 107, 108, 109, 110, 19, 111, 112, 18, 113, 51, 88, 114]. To date, such enhancements have been developed in the context of local thermometry, aiming at designing a thermometer that detects the smallest temperature variations around a known temperature [19, 46]. In many practical situations, however, one might not know the temperature accurately beforehand. Rather, one has only limited prior knowledge about the temperature of the sample. Under such circumstances, Bayesian estimation is a more suitable approach to thermometry, and has been the subject of a

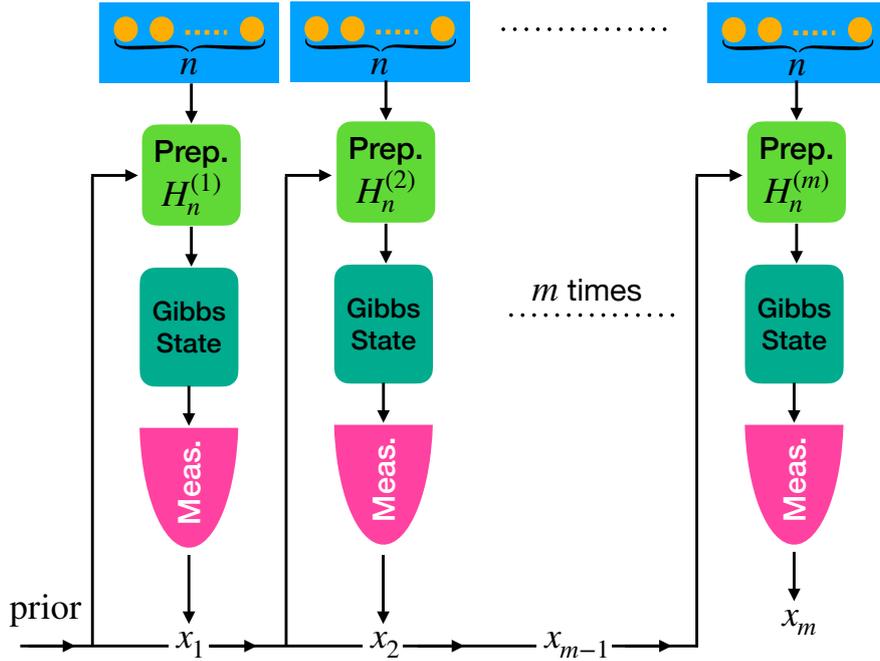


Figure 5.1: *This figure is taken from Ref. [93].* Schematic representation of the adaptive scenario. A total of N probes are used in groups of n to estimate the temperature of the sample, θ_0 . Initially, our prior temperature distribution is given by $p(\theta)$, according to which we choose the Hamiltonian of the first n probes to be $H_n^{(1)}$ that minimises the expected mean square logarithmic error. The probes interact and thermalise with the sample followed by an energy measurement, yielding an outcome, say x_1 . Our knowledge about the temperature will be reflected in the posterior distribution $p(\theta|x_1)$. This will be used as the prior for the second round—in order to find the optimal Hamiltonian $H_n^{(2)}$. This process is repeated $m = N/n$ times. In contrast, in the non-adaptive scenario the Hamiltonian is fixed $H_n^{(k)} = H_n \forall k$.

few recent studies [78, 79].

The goal of this work is to set the ultimate bounds of Bayesian equilibrium thermometry, and to develop adaptive strategies to saturate them. It is insightful to first recall analogous results in the local approach to equilibrium thermometry [19, 46]. In that case, for any unbiased estimator $\tilde{\theta}$ of the temperature θ_0 , the absolute error is inversely proportional to the heat capacity of the probe: $\Delta\tilde{\theta} \propto 1/C$ [84, 83, 50, 107]. For n -body probes, C can scale super-extensively with n in the vicinity of a critical point, with the ultimate bound $C \approx n^2/4$ [50, 115]—a quadratic scaling with the number of resources reminiscent of the Heisenberg scaling in quantum metrology [3]. Here, we show that similar bounds hold in the Bayesian approach, but that adaptive strategies are needed to saturate them in contrast to the local case. In fact, we prove that any non-adaptive strategy necessarily leads to $\Delta\tilde{\theta} \propto 1/n$ for sufficiently large n —i.e., a shot-noise-like scaling [3]—a no-go result that holds even when arbitrary control over the n -body probe Hamiltonian is allowed. This shows that adaptive measurement strategies are a crucial ingredient for optimal quantum thermometry whenever the temperature value is *a priori* not perfectly known.

5.2 Preliminaries and setup.

We consider estimation of temperature of a sample, denoted as θ_0 , given some prior distribution $p(\theta)$ reflecting our initial knowledge on θ_0 . We assume we have at our disposal N copies of a d -dimensional system that we use as probes. The temperature is inferred by letting the probes reach thermal equilibrium at θ_0 , and then measuring them—corresponding to the framework of equilibrium thermometry [19, 46]. In order to establish fundamental bounds, we assume full control on the Hamiltonian of the probes, and in particular the ability to make them interact. Therefore, alternatively one can think of a d^N -dimensional probe, which constitutes our resource.

The thermometry process is divided into m rounds, each involving $n = N/m$ probes. Every round consists of: (I) preparation of the n -body probe, (II) interaction with the sample and thermalisation, (III) measurement/data acquisition, and (IV) data analysis (see Fig. 5.1). In the first round, we start by engineering the Hamiltonian $H_n^{(1)}$ of the n -body probe into any desired configuration based on the prior distribution $p(\theta)$. That is, we arrange the energy distribution of the n -body probe to become most sensitive to the relevant temperature range. Next, in step (II), this n -body system is put in contact with the sample, reaching a Gibbs state $\omega_{\theta_0}(H_n^{(1)}) = \exp[-H_n^{(1)}/\theta_0]/Z$, with $Z = \text{Tr}(\exp[-H_n^{(1)}/\theta_0])$ the partition function. Then, in step (III), a measurement is performed that yields an outcome x_1 . We focus on energy measurements since they are optimal as the Gibbs state is diagonal in the energy basis. In the data analysis (step (IV)), the measurement outcome is used to obtain the posterior distribution through Bayes' rule:

$$p(\theta|x_1) = \frac{p(x_1|\theta)p(\theta)}{p(x_1)}, \quad (5.1)$$

where $p(x|\theta)$ is the likelihood function (which depends on the temperature *and* the Hamiltonian), $p(\theta)$ is the prior probability distribution on θ , and $p(x) = \int d\theta p(\theta) p(x|\theta)$ is the outcome probability. The next round proceeds in an analogous way, but replacing the prior $p(\theta)$ by $p(\theta|x_1)$ and $H_n^{(1)}$ by $H_n^{(2)}$. Likewise, in round $k > 1$, $p(\theta)$ is replaced by $p(\theta|\mathbf{x}_{k-1})$ with $\mathbf{x}_{k-1} \equiv \{x_{k-1}, \dots, x_2, x_1\}$ and $H_n^{(1)}$ is replaced by $H_n^{(k)}$. Such a strategy is adaptive since $H_n^{(k)}$ depends on \mathbf{x}_{k-1} . In contrast, a non-adaptive strategy satisfies $H_n^{(k)} = H_n \forall k$, where H_n is chosen according to the initial prior $p(\theta)$ only. At the end of the thermometry process (round m), the final estimate $\tilde{\theta}(\mathbf{x}_m)$ of θ_0 is computed.

In order to gauge the quality of the estimator, we need to introduce some figure of merit, or an error quantifier that describes how far $\tilde{\theta}$ is from θ_0 , on average. A natural measure which is suitable for equilibrium probes is the expected mean-square logarithmic error (EMSLE) (see [78] for justification and the accompanying paper [52] for a deeper analysis and generalisation)

$$\text{EMSLE} := \int d\theta p(\theta) \int d\mathbf{x}_m p(\mathbf{x}_m|\theta) \ln^2 \left[\frac{\tilde{\theta}(\mathbf{x}_m)}{\theta} \right], \quad (5.2)$$

with $d\mathbf{x}_m := dx_m \dots dx_1$. Moreover,

$$\tilde{\theta}(\mathbf{x}_m) = \exp \left[\int d\theta \frac{p(\theta)p(\mathbf{x}_m|\theta)}{p(\mathbf{x}_m)} \ln \theta \right], \quad (5.3)$$

is the optimal temperature estimator, i.e., it minimises the EMSLE [78].

We wish to find lower bounds for EMSLE, as well as optimal strategies to saturate them, for both adaptive and non-adaptive measurements. More precisely, our aim is to minimise EMSLE as a function of the number N of probes, with $N = mn$. We will pay particular attention to the relevant case where $m \gg 1$ is large (asymptotic regime) but n is limited due to e.g. experimental limitations on the amount of probes that can be collectively processed. In this case, we will focus on the scaling of EMSLE with n for a fixed but large m .

5.3 Main results

Our main results are (i) an ultimate precision limit for Bayesian thermometry that holds for both adaptive and non-adaptive strategies, which in principle allows for a quadratic (Heisenberg-like) scaling with n , (ii) a no-go theorem that forbids super-extensive scaling in any non-adaptive scenario, and (iii) an adaptive strategy that reaches the ultimate limit. These results are derived in what follows (some technical details are given in the Appendix).

Given the prior distribution $p(\theta)$, one may utilise the Van Trees inequality [54, 55] to construct a lower bound on the estimation error after m rounds of the protocol

$$\begin{aligned} \text{EMSLE}^{-1} &\leq Q[p(\theta)] \\ &+ \sum_{k=1}^m \int d\mathbf{x}_{k-1} p(\mathbf{x}_{k-1}) \int d\theta p(\theta|\mathbf{x}_{k-1}) C(\theta; H_n^{(k)}), \end{aligned} \quad (5.4)$$

where $p(\theta|\mathbf{x}_0) = p(\theta)$, $p(\mathbf{x}_0) = p(\mathbf{x}_0|\theta) = 1$, and $\int d\mathbf{x}_0 = 1$ are introduced to compress our notation. Here, $Q[p(\theta)]$ quantifies the prior information given by

$$Q[p(\theta)] := \int d\theta p(\theta) [1 + \theta \partial_\theta \log p(\theta)]^2, \quad (5.5)$$

while the second term takes into account the information acquired through all measurements. The heat capacity of the probe at round k of the measurement is denoted $C(\theta; H_n^{(k)})$, with the Hamiltonian $H_n^{(k)}$ designed according to the prior *and* the information acquired so far. Recall that, by definition, $C(\theta; H_n) := \partial_\theta E(\theta; H_n)$ where $E(\theta; H) = \text{Tr}[H\omega_\theta(H)]$ is the energy of the probe at thermal equilibrium. In order to bound Eq. (5.4), we first define the maximum of the integrand over $\{H_n^{(k)}\}_{k=1}^m$ for a given specific trajectory \mathbf{x}_m :

$$\begin{aligned} \Gamma(\mathbf{x}_m) &:= \max_{\{H_n^{(k)}\}_k} \sum_{k=1}^m \int d\theta p(\theta|\mathbf{x}_{k-1}) C(\theta; H_n^{(k)}) \\ &\leq \sum_{k=1}^m \int d\theta p(\theta|\mathbf{x}_{k-1}) C_D = mC_D \end{aligned} \quad (5.6)$$

where $C_D := \max_{H_n} C(\theta; H_n)$, i.e., the maximum heat capacity of an n -body probe, and in the last line we used that C_D is independent of θ (see [50] and the Appendix for an explicit expression for C_D). Furthermore, it is bounded by $C_D \leq \frac{n^2}{4} \log^2 d$, which is

quickly saturated as n is increased. Putting everything together, we obtain from (5.4) the following chain of inequalities:

$$\begin{aligned} \text{EMSLE}^{-1} &\leq Q[p(\theta)] + mC_D \\ &\leq Q[p(\theta)] + m\frac{n^2}{4}\log^2 d. \end{aligned} \quad (5.7)$$

This gives an ultimate bound on Bayesian thermometry [Result (i)], which both adaptive and non-adaptive strategies should respect. This bound implies that any Bayesian thermometry protocol is ultimately limited by a quadratic Heisenberg-like scaling.

The ultimate bound (5.7) becomes tight and can be saturated by adaptive strategies in the regime $m \gg 1$ (see results below). However, let us comment that any non-adaptive strategy fails to saturate it, and in fact EMSLE^{-1} can increase at most linearly with n [Result (ii)]

$$\text{EMSLE}^{-1} \stackrel{\text{non-adaptive}}{\leq} Q[p(\theta)] + f[p(\theta)]mn \log d, \quad (5.8)$$

This is rigorously proven in the supplementary materials, but let us here mention the intuition behind it. It has already been noted in the literature that engineered probes for thermometry show enhanced sensitivity only in a small temperature range Δ [50, 19, 116, 117, 80]. In fact, finite-size scaling theory tells us that if $C \propto n^{1+\alpha}$, then $\Delta \propto n^{-\gamma}$ with $\gamma \geq \alpha$ in order to ensure that the energy density of an equilibrium state remains finite [118]. This implies that, for any $p(\theta)$ with a finite width (independent of n), the figure of merit $\int d\theta p(\theta)C(\theta)$ entering into (5.4) can grow at most linearly with n for sufficiently large n . In other words, optimal n -body probes require priors with a width smaller than $O(1/n)$ to obtain super-linear scaling, and conversely a finite width in $p(\theta)$ will eventually kill any super-linear scaling. The no-go result (5.8) makes this intuition rigorous. We also note that (5.8) is a particular case of a more general family of bounds where the prior-dependent quantity $f[p(\theta)]$ can be replaced by other functionals which can become tighter for some $p(\theta)$; this is discussed in the Appendix.

The above reasoning also explains why adaptive protocols can potentially saturate (5.7). By updating the prior $p(\theta)$ to the posterior $p(\theta|\mathbf{x}_{j-1})$ in each step of the process ($j = 1, \dots, m$), the posterior can stay inside the optimal region for sufficiently large m , thus enabling super-linear thermometry precision. Furthermore, as a result of sharpening the prior knowledge, we are getting closer to a local thermometry scenario [119]. This suggests using optimal probes for local thermometry as an *ansatz* for the Bayesian thermometry with adaptive strategies. The optimal thermometer in the local scenario is an effective two-level system with $d^n - 1$ -fold degeneracy in the excited state, with an energy gap that depends both on n and the temperature [50]. For large n its heat capacity reads $C_D \approx (n^2/4)\log^2 d$, as already mentioned [115]. Inspired by this, at the k th round we restrict to the class of Hamiltonians $H_n^{(k)}$ with the aforementioned two-level structure, and tune the energy gap such that the EMSLE (5.2) is minimised. As we shall see in the example below, we are able to achieve a quadratic scaling with n and saturate (5.7) through this strategy [Result (iii)].

5.4 Case study

The results presented here are valid for a broad class of priors, but in what follows we stick to a specific choice in order to illustrate their usage. In any relevant application

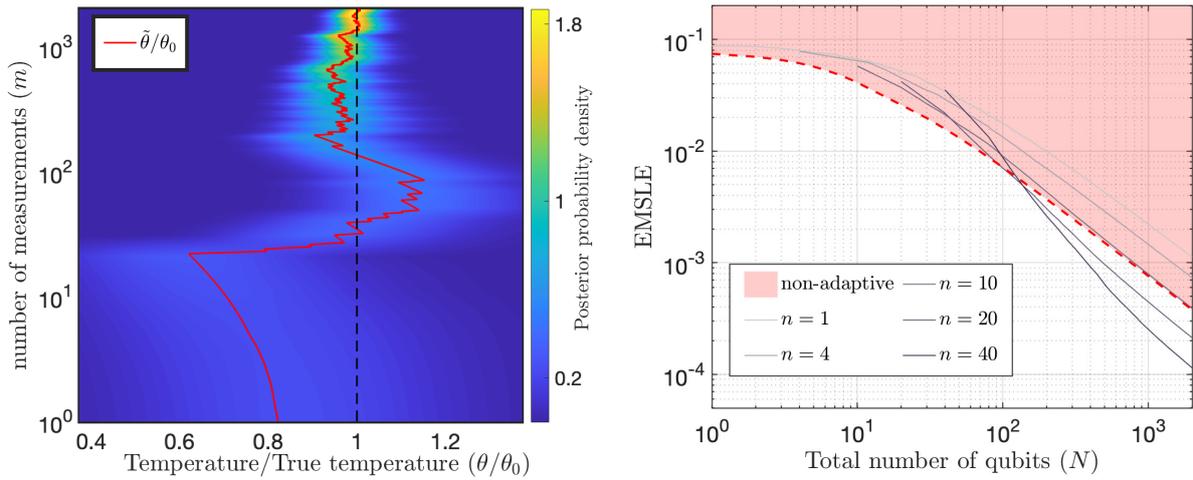


Figure 5.2: *This figure is taken from Ref. [93].* Left—Contour plot of the prior versus the number of simulation rounds m (logarithmic scale), and the ratio $\tilde{\theta}/\theta_0$ between the estimated $\tilde{\theta}$ and true temperature θ_0 . As m increases, the prior sharpens around the true temperature, and $\tilde{\theta}/\theta_0$ approaches unity. Here, we have set $n = 1$, $\alpha = 1$, $\theta_{\min} = 1$, and $\theta_{\max} = 10$ in arbitrary units. Right—Loglog plot of the expected mean square logarithmic error (EMSLE) attained by the adaptive strategy vs. the total number of qubits N . Dark solid lines represent different values of n . They show that, for sufficiently large N , the bigger n is the smaller the error can get. The red-dashed line is the (not necessarily tight) bound on non-adaptive strategies: only the shaded area can be achieved using non-adaptive protocols. One can cross the border with adaptive strategies for $n > 10$.

of thermometry, the temperature is known a priori to lie within a certain range, i.e., $\theta_{\min} \leq \theta_0 \leq \theta_{\max}$. We use a family of probability distributions that are suitable in this case and were proposed in [69]:

$$p(\theta) = \frac{1}{k_\alpha(\theta_{\max} - \theta_{\min})} \left[e^{\alpha \sin^2\left(\pi \frac{\theta - \theta_{\min}}{\theta_{\max} - \theta_{\min}}\right)} - 1 \right] \quad (5.9)$$

with

$$k_\alpha := e^{\alpha/2} I_0(\alpha/2) - 1, \quad (5.10)$$

where I_0 is the modified Bessel function of the first kind. In the limit $\alpha \rightarrow -\infty$ the above prior becomes a constant over the domain, while in the limit $\alpha \rightarrow 0$ we have $p(\theta) \propto \sin^2(2\theta)$.

The adaptive strategy works as follows. We consider as a resource N qubits, which are divided in m groups of n qubits. In each group, the n -qubit Hamiltonian is engineered to become a two-level system with degeneracy $(2^n - 1)$ and with a tunable gap ϵ (recall that this structure is optimal for local thermometry). In the first round, we tune the gap to $\epsilon^{(1)}$ to minimise the single shot EMSLE, that is we set $m = 1$ in (5.2). Then, we measure the energy of the system. Given the outcome x_1 is observed, we update the prior to $p(\theta) \rightarrow p(\theta|x_1)$, and implement the same procedure to choose $\epsilon^{(2)}$ in the second round (i.e., we minimise (5.2) replacing $p(\theta) \rightarrow p(\theta|x_1)$). This process is repeated until all probes are used.

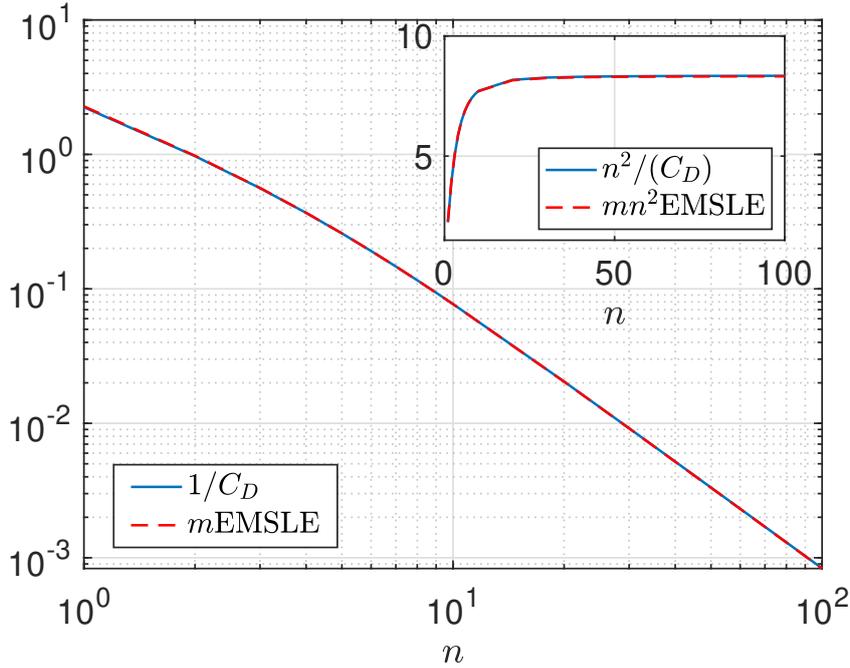


Figure 5.3: *This figure is taken from Ref. [93].* (Dashed red) Loglog plot of the normalised expected mean square logarithmic error (EMSLE) after m rounds of the adaptive scheme—for sufficiently large m —vs n . This shows that for large enough n the error vanishes quadratically with n , which can be better seen from the inset. (Blue) The minimum achievable EMSLE given by the r.h.s. of (5.7). The perfect agreement shows the efficiency of the proposed adaptive protocol.

In our simulations, we apply the adaptive process for a given θ_0 sampled from $p(\theta)$, which yields a trajectory as illustrated in the left panel of Fig. 5.2. We see that the prior peaks around the true temperature as m increases, and the estimated temperature gets closer to the true temperature, i.e., $\tilde{\theta}/\theta_0 \rightarrow 1$. The average over a large amount of trajectories enables us to compute EMSLE in Eq. (5.2) with high accuracy (in the numerical simulations, we consider $\mathcal{O}(1000/m)$ trajectories, which is enough to ensure convergence). This is used in the right panel of Fig. 5.3, where we plot EMSLE in the adaptive scenario for various values of n , benchmarked against the no-go bound for non-adaptive scenarios—only the shaded area can be accessed by non-adaptive strategies given any $n \leq N$. We see that as we increase n the error gets smaller for large enough N . In particular, there exist some threshold n (in this example $n > 10$) for which one can beat the no-go bound via adaptive strategies. As an example, given $N = 10^3$ and $\theta_{\max}/\theta_{\min} = 10$ in Eq. (5.9) (with $\alpha = 1$), adaptive strategies using $n \approx 10$ interacting qubits outperform arbitrary non-adaptive strategies, even ones involving full control of all the 10^3 qubits.

Finally, we ask whether the adaptive strategy can reach the Heisenberg-like scaling, $\text{EMSLE}^{-1} \propto mn^2$. To this aim, we study the behaviour of the error with the resources n for a sufficiently large number of repetitions m . The results are depicted in Fig. 5.3, where we see Eq. (5.7) is saturated and therefore the proposed adaptive scheme reaches the ultimate bound on thermometry.

5.5 Conclusions and future directions

We derived fundamental limitations of the Bayesian approach to equilibrium thermometry, which shows a Heisenberg-like quadratic scaling with the number of probes. We showed that non-adaptive strategies cannot saturate this bound and, in fact, are limited to shot-noise-like scaling whenever the initial prior is not sharp. We also constructed a simple adaptive protocol which can saturate the ultimate bound, thus highlighting the crucial role of adaptivity in quantum thermometry. This role is importantly different to the setting of Bayesian quantum phase-estimation protocols [120], in which the corresponding Heisenberg limit that applies to most general adaptive protocols [121] can be attained by resorting only to measurements being adaptively varied in between the phase-encoding channel uses [122]. In contrast, in equilibrium thermometry the form of probe states (Gibbs) and measurement (energy-basis) is fixed, and it is the thermalising channel (i.e. the probe Hamiltonian) that must be adaptively adjusted for the quadratic scaling indicated by the local approach [50] to become reachable.

While these results were obtained by taking the expected mean square logarithmic error as the precision quantifier, it is important to note that any reasonable error measure leads to the same main conclusions [78, 119]. The optimal strategy relies on using a highly degenerate two-level system as a probe, as in the local regime [50]. Although this is a very idealised Hamiltonian, Hamiltonians that achieve similar sensitive precision have been proposed in the low-temperature regime through a one-dimensional fermionic system [115], and in an upcoming work we will discuss a Hamiltonian probe reaching quadratic scaling with n based on the Ising model. We leave as interesting future work the establishment of similar Bayesian bounds in nonequilibrium thermometry [103, 123, 124] and critical metrology [125, 126].

5.6 Appendix

5.6.1 Van Trees inequality in equilibrium thermometry

The van Trees inequality was derived in chapter. 3. To derive the expression used in the main text, we simply have to evaluate the van Trees inequality (Eq. (3.38)) in the specific case $\lambda(\theta) = \log(\theta)$. This gives:

$$\text{EMSLE}^{-1} \leq Q[p(\theta)] + \sum_{k=1}^m \int d\mathbf{x}_{k-1} p(\mathbf{x}_{k-1}) \int d\theta p(\theta|\mathbf{x}_{k-1}) C(\theta; H_n^{(k)}). \quad (5.11)$$

where

$$Q[p(\theta)] = \int d\theta p(\theta) [1 + \partial_\theta \log p(\theta)]^2. \quad (5.12)$$

Our interest is in finding the probe Hamiltonian, for which the van Trees inequality is minimal. The optimal probe is then the one maximizing the averaged heat capacity

$$\Gamma(\mathbf{x}_{k-1}) := \max_{H_n^{(k)}} \int d\theta p(\theta|\mathbf{x}_{k-1}) C(\theta; H_n^{(k)}). \quad (5.13)$$

Note that Γ is in general a functional of the past measurement trajectory, i.e., the optimal probe structure depends on the prior knowledge of the parameter to be estimated. In the following sections we derive model-independent upper bounds on Γ .

5.6.2 Model-independent super-extensive upper bound on Γ

In this section we derive a super-extensive bound on $\Gamma(\mathbf{x})$. Starting with Eq. (5.13), we note that since the integrand is positive we can provide an upper bound by moving from a global maximization to a local maximization, i.e.

$$\Gamma(\mathbf{x}) \leq \int d\theta p(\theta|\mathbf{x}) \max_H C(\theta; H). \quad (5.14)$$

The problem of maximizing the heat capacity, over all possible probe Hamiltonians at a given temperature, has been solved by Correa et al. [50]. The solution can be formulated as the temperature-independent tight upper bound

$$C(\theta; H) \leq \left[\frac{\xi_D}{2} \right]^2 - 1, \quad (5.15)$$

where ξ_D is the solution to the transcendental equation

$$e^{\xi_D} = (D - 1) \frac{\xi_D + 2}{\xi_D - 2}. \quad (5.16)$$

This equation does not have a closed form solution. However, a general feature of the solution is that $\xi_D > \log(D - 1)$, and that ξ_D approach $\log(D - 1)$ from above as D becomes large. From this it follows that $\Gamma(\mathbf{x})$ satisfies the super-extensive upper bound

$$\Gamma(\mathbf{x}) \leq (\xi_D/2 - 1) (\xi_D/2 + 1) \int d\theta p(\theta|\mathbf{x}), \quad (5.17)$$

which grows super-extensively in $\log(D)$. If we average $\Gamma(\mathbf{x})$ over the past measurement trajectory we find

$$\int d\mathbf{x} p(\mathbf{x}) \Gamma(\mathbf{x}) \leq (\xi_D/2 - 1) (\xi_D/2 + 1) := C_D. \quad (5.18)$$

This bound is expected to be approximately tight in the limit where the prior is local with respect to the width of the heat capacity. As we will see in the next section, designing a probe with a critical heat capacity at a certain temperature, i.e. one attaining the maximal heat capacity, will result in the width of the heat capacity decreasing as $1/\log(D)$. We thus see that saturating the super-extensive bound requires a prior probability distribution confined to a domain $\theta \in [\theta_c - \Delta/2, \theta_c + \Delta/2]$ where θ_c is the critical temperature and $\Delta = 1/\log(D)$. As D increase this corresponds to an increasing amount of prior information.

5.6.3 Tight upper bound on the thermal energy density

In this section we want to derive an upper bound on the thermal energy at a given temperature for any probe structure, subject to the dimensionality constraint $\dim H = D$ on the considered probes. We will find that the thermal energy density is upper bounded by the temperature. Define the maximum thermal energy for any probe structure as

$$E_{\max}(\theta) := \max_H E(\theta; H), \quad (5.19)$$

$$E(\theta; H) := \text{Tr}\{H\omega(\theta; H)\}, \quad (5.20)$$

where $\omega(\theta; H)$ is a thermal state at temperature θ . We denote the energy eigenvalues of the probe Hamiltonian by $\{\varepsilon_l\}$, and for convenience set the ground-state energy to zero. If we take the derivative of the thermal energy, and equate to zero we obtain the condition

$$\varepsilon_l = \theta + E(\theta; H) := \varepsilon, \quad (5.21)$$

which implies a $D - 1$ degeneracy in the first excited state. Evaluating the above condition for this probe structure leads to a transcendental equation for ε/θ which can be solved. The result is the temperature-dependent upper bound

$$E(\theta; H) \leq \theta \mathcal{W}_D \quad (5.22)$$

$$\mathcal{W}_D := W\left(\frac{D-1}{e}\right), \quad (5.23)$$

where W denotes the product logarithm, also called the *Lambert W* function. In the limit of large D the behaviour of the product logarithm is such that \mathcal{W}_D tends asymptotically to $\log(D)$ from below. We stress that the above bound on the thermal energy can be saturated by an effective two level probe with a $D - 1$ degenerate excited state, and a temperature-dependent energy gap.

5.6.4 Extensive bounds for the non-adaptive scenario

We start with the second term in Eq (4) of the main text. Since the Hamiltonian remains constant throughout the protocol, i.e., $H_n^{(k)} = H_n \forall k$, this term can be rewritten as

$$\begin{aligned} \bar{\Gamma} &:= \sum_{k=1}^m \iint d\theta d\mathbf{x}_{k-1} p(\theta) p(\mathbf{x}_{k-1}|\theta) C(\theta; H_n) \\ &= m \int d\theta p(\theta) C(\theta; H_n). \end{aligned} \quad (5.24)$$

Integrating by parts—recall that $C(\theta; H_n) = \partial_\theta E(\theta; H_n)$ —and maximising over H_n gives

$$\bar{\Gamma} \leq m \max_{H_n} \int d\theta [-\partial_\theta p(\theta)] E(\theta; H_n), \quad (5.25)$$

where we assumed that $p(\theta)E(\theta; H_n)$ is smooth and vanishes at the boundaries. By defining \mathcal{R} as the temperature domain where $\partial_\theta p(\theta) \leq 0$ we have

$$\begin{aligned} \bar{\Gamma} &\leq m \max_{H_n} \int_{\mathcal{R}} d\theta [-\partial_\theta p(\theta)] E(\theta; H_n) \\ &\leq m \int_{\mathcal{R}} d\theta [-\partial_\theta p(\theta)] \max_{H_n} E(\theta; H_n) \end{aligned} \quad (5.26)$$

To make further progress, we use the upper bound on the energy of an n -body system at thermal equilibrium (with total dimension $D = d^n$) that is given by Eq. (5.22):

$$\max_{H_n} E(\theta; H_n) \leq \theta \mathcal{W}_D \leq \theta n \log d \quad (5.27)$$

where the second equality is saturated as $n \gg 1$. Plugging these results back into Eq. (4) of the main text we obtain a no-go theorem for non-adaptive strategies [Result (ii)]

$$\text{EMSLE}^{-1} \stackrel{\text{non-adaptive}}{\leq} Q[p(\theta)] + f[p(\theta)]mn \log d, \quad (5.28)$$

where $f[p(\theta)] = \int_{\mathcal{R}} d\theta [-\partial_{\theta} p(\theta)]\theta$ is a functional of the prior. Crucially, the bound (5.8) implies that, even with arbitrary control over the n -body Hamiltonian, one cannot go above a linear scaling in n with non-adaptive strategies (compare with the general bound given by Eq. (7) of the main text).

Our alternative bound follows the exact same procedure, except we first recall that the thermal energy can be expressed as $\theta^2 \partial_{\theta} \Psi(\theta; H)$, where the Massieu potential reads $\Psi(\theta; H) := \log \mathcal{Z}(\theta; H)$ with $\mathcal{Z}(\theta; H)$ being the partition function of the probe. Starting from Eq. (5.24) and by performing twice integration by parts we get

$$\begin{aligned} \bar{\Gamma} &= m \int d\theta p(\theta) C(\theta; H_n) = -m \int d\theta [\partial_{\theta} p(\theta)] E(\theta; H) \\ &= m \int d\theta [\partial_{\theta} (\theta^2 \partial_{\theta} p(\theta))] \Psi(\theta; H), \end{aligned} \quad (5.29)$$

where again we take the vanishing and differentiability of the boundary terms in both integrations—that is $p(\theta)E(\theta; H_n)$ and $\theta^2 \partial_{\theta} p(\theta)\Psi(\theta; H)$ —as a restriction on the choice of parameterization. We can derive an upper bound on the optimal solution by noting that $\Psi(\theta; H) \geq 0$ —recall that the ground state energy is set to zero—and by introducing $\bar{\mathcal{R}} = \{\theta \mid \partial_{\theta} (\theta^2 \partial_{\theta} p(\theta)) \geq 0\}$. Then

$$\bar{\Gamma} \leq m \max_H \int_{\bar{\mathcal{R}}} d\theta [\partial_{\theta} (\theta^2 \partial_{\theta} p(\theta))] \Psi(\theta; H). \quad (5.30)$$

As the integrand is now positive we can maximize the Massieu potential locally. Since the logarithm is monotonically increasing in its argument, this corresponds to substituting the largest value of the partition function, i.e. the Hilbert space dimension. The bound then takes the form

$$\begin{aligned} \bar{\Gamma} &\leq m \log(D) \int_{\bar{\mathcal{R}}} d\theta [\partial_{\theta} (\theta^2 \partial_{\theta} p(\theta))] \\ &\leq m \log(D) \{\theta^2 \partial_{\theta} p(\theta)\}_{\bar{\mathcal{R}}} \\ &:= m \log(D) g[p(\theta)], \end{aligned} \quad (5.31)$$

where $g[p(\theta)]$ is a functional of the prior distribution but independent of the probe. This gives two complementary bounds on $\bar{\Gamma}$, i.e. one expressed in terms of $f[p(\theta)]$ as presented in the main text, and one in terms of $g[p(\theta)]$. Which of these two is tighter depends on the specific prior.

5.6.5 Model-independent bound for a Gaussian prior

For illustration purposes we will study the behaviour of the model-independent extensive bounds derived above for the specific case of a Gaussian prior over $\lambda(\theta) = \log(\theta)$. If we assume that the mean $\bar{\lambda}$ and the variance σ^2 are known, then the maximum entropy principle—see Ref. [53]—tells us that a normal distribution is a suitable prior

$$p(\theta; \bar{\theta}, \sigma) = \frac{1}{\sqrt{2\pi}\theta\sigma} e^{-\frac{1}{2\sigma^2}(\log \theta/\bar{\theta})^2}. \quad (5.32)$$

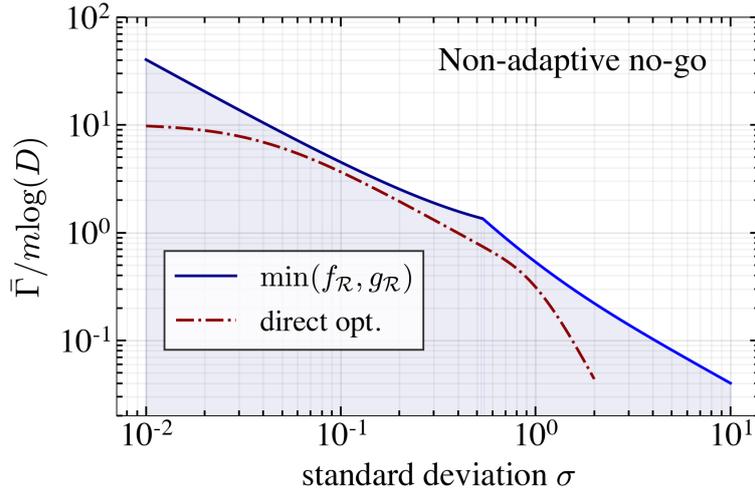


Figure 5.4: Illustration of non-adaptive no-go result for a Gaussian prior. (a) The model-independent extensive upper bound evaluated for a Gaussian prior on $\log(\theta)$. The red line gives the optimal $\bar{\Gamma}$ found through a direct numerical optimization of the probe structure in the case of an effective two-level probe consisting of $n = 60$ qubits.

Note that we are adopting a normal distribution over the parameter λ , this does not generally translate directly to a normal distribution over temperature.

We are going to compare the extensive bounds to a direct numerical optimization of $\bar{\Gamma}$, given a probe structure which is an effective two-level system. The solution to this optimization problem is simplified by the fact that the optimization of the probe structure will only depend on temperature ratios, i.e.

$$\begin{aligned} \bar{\Gamma} &= \max_H m \int d\theta p(\theta; \bar{\theta}, \sigma) C(\theta; H_n) \\ &= \max_H m \int d\theta p(\theta; \gamma \bar{\theta}, \sigma) C(\theta; H_n), \end{aligned} \quad (5.33)$$

where γ is a real positive constant. This shows that $\bar{\Gamma}$ is independent of the mean $\bar{\theta}$ and only depends on the variance σ . In the case of the Gaussian prior we can evaluate both the model-independent extensive bounds straightforwardly, the results are plotted in figure 5.4 alongside a direct maximization over the locally optimal probe structure, i.e. an effective two-level probe with a $D - 1$ degenerate excited state.

In the case of the Gaussian prior we can furthermore evaluate the van Trees inequality analytically. First of all we have $Q[p(\theta; \bar{\theta}, \sigma)] = \frac{1}{\sigma^2}$. Furthermore the model-independent extensive bound $f_{\mathcal{R}}$, which turns out to be relevant in the regime of small σ , can be evaluated. The resulting bound takes the form

$$\overline{\text{EMSLE}}^{-1} \leq \frac{1}{\sigma^2} + mn \log(d) \left[1 + \frac{1}{\sqrt{2\pi}\sigma} e^{-\sigma^2/2} \right]. \quad (5.34)$$

We then observe that in the limit of a prior distribution with a vanishing variance $\sigma \rightarrow 0$, the model-independent bound is dominated by the contribution from $Q[p(\theta; \bar{\theta}, \sigma)]$ unless $n \log(d)$ is on the order of σ^{-1} . This illustrated the general point that the extensive bound is relevant when the prior width exceeds the critical domain, i.e. $\sigma \gtrsim \log(D)^{-1}$.

Chapter 6

Tight bound on finite-resolution quantum thermometry at low temperatures

This chapter is composed entirely of text and figures from Ref. [88] © 2020 American Physical Society. The numerical simulations performed for the spin-boson model are in essence an extension of work initiated during my MSc studies and completed during the first year of my PhD studies. The work published in Ref. [127] © 2019 American Physical Society, was worked out fully during my PhD. The paper is reproduced in appendix B

The chapter is self-contained. In particular the notation does not conform to previous chapters. In addition the term relative error, refers to the noise-to-signal ratio in this chapter. This chapter takes a conventional approach to thermometry, i.e. one based on the absolute error and the associated noise-to-signal ratio. The technical results of the chapter are valid. However, the physical implications should be judged in light of the results of chapter 4.

6.1 Introduction

Sensitive measurements of temperature are essential throughout natural science and modern technology. Increasingly detailed studies of biological, chemical, and physical processes, the miniaturisation of electronics, and emerging quantum technology drive a need for new thermometry techniques applicable at the nanoscale and in regimes where quantum effects become important. Many new approaches are being developed [128, 19, 90, 129, 130, 8, 131, 25, 132, 133, 134, 98], however the fundamental limits to precision thermometry are not yet fully understood. Here, we determine a tight bound on the best possible precision with which temperature can be estimated in cold quantum systems, which accounts for limitations due to imperfect measurements.

The classical picture of thermometry is that of a thermometer which is brought into thermal contact with a sample. Observing the state of the thermometer after some time conveys information about the sample temperature. A similar picture can be applied in the quantum regime, where an individual quantum probe, e.g. a two-level

system, may interact with a sample system in a thermal state, and subsequently be measured to estimate the temperature. If the probe reaches thermal equilibrium with the sample, or a non-equilibrium steady state, optimal designs of the probe and of the probe-system interaction can be determined [50, 106, 19, 109, 82, 113]. Outside of the steady state regime, it was found that access to the transient probe dynamics may outperform the steady-state protocols [103, 135, 136], that dynamical control acts as a resource [137, 138, 139], and that thermometry can in some cases be mapped to a phase estimation problem [140, 141]. These findings have spurred further investigations into non-equilibrium thermometry [123, 142, 110].

Any thermometric technique will be subject to constraints due to finite measurement resolution. In the probe-sample picture, the size of the probe will limit the amount of information which can be extracted about the sample. More generally, any measurement on the sample, implemented using a finite-sized apparatus, comes with a lower bound on the attainable resolution of e.g. the system energy spectrum [143, 51, 144]. Similar restrictions apply in situations where measurements can be made on only part of a large sample [145, 91, 104], and clearly such finite-resolution constraints must play an important role in formulating fundamental bounds on the attainable thermometric sensitivity.

Here, we derive a bound on the precision scaling with temperature, as the temperature approaches zero, for thermometers with finite energy resolution. Our bound applies to any thermometric technique based on measurements which do not resolve the individual energy levels of the sample energy spectrum. We furthermore demonstrate that this scaling can be attained using a single-qubit probe, showing that the bound is tight. To derive our bound, we build upon the framework for finite-resolution quantum thermometry introduced by Potts, Brask, and Brunner in Ref. [51].

Our results also demonstrate that thermometry with a vanishing absolute error at low temperature is possible with finite resolution, answering an interesting question left open by previous work [51, 91, 105]. For systems with a heat capacity that vanishes at low temperatures, a property often included in the third law of thermodynamics, the relative error must diverge, regardless of the available resolution [51]. The absolute error may either also diverge, stay constant, or vanish, with the latter thus being the best behaviour one can hope for. However, for gapped systems, even the absolute error in any unbiased temperature estimate must diverge when the temperature becomes comparable to the gap [83]. A constant or vanishing absolute error, on the other hand, has been seen in gapless systems, when employing a measurement with a continuous outcome implying an infinite resolution [91]. Our results show that a vanishing absolute error may be obtained with a finite-resolution measurement having as little as two outcomes.

This paper is organized as follows. In Sec. 6.2 we introduce a general temperature estimation procedure, following [51], and discuss the fundamental precision bounds imposed by the third law of thermodynamics. In Sec. 6.3 we propose a finite-resolution criterion, and show how this criterion leads to a tight bound on the attainable precision. In Sec. 6.4 we generalize the framework to include noisy measurements, and finally in Sec. 6.5 we investigate a single-qubit thermometer coupled to a bosonic bath, showing that our bound can be saturated in a physical scenario. Our analytical results are supported by numerical simulations of the temperature estimation procedure.

6.2 Temperature estimation

We consider a quantum system described by the canonical thermal state

$$\rho_\beta = \exp[-\beta H] / \mathcal{Z}_\beta, \quad (6.1)$$

with H the Hamiltonian operator of the system, and $\mathcal{Z}_\beta \equiv \text{Tr} \{ \exp[-\beta H] \}$ the canonical partition function. The thermal state is parameterized by an inverse temperature $\beta = 1/k_B T$ where k_B is the Boltzmann constant. For convenience we adopt units in which $k_B = 1$, such that temperature has the units of energy. The task we consider is how to estimate the temperature T of the system. We remark that throughout we consider thermal states where the temperature does not itself fluctuate. However, since temperature is not directly measurable (it is not a quantum mechanical observable), there are fluctuations in any temperature estimate based on indirect measurements.

6.2.1 Quantifying the estimation precision

A general temperature estimation procedure consists of first performing a measurement on the system. The most general N -outcome measurement is represented by a positive-operator valued measure (POVM) with N elements Π_m . Such POVMs capture any possible measurement in quantum mechanics, including scenarios in which information is obtained through a probe interacting with the system, as well as those exploiting quantum coherence [131, 103, 135]. Each POVM element Π_m corresponds to a measurement outcome m which is observed with probability

$$p_{m;\beta} = \text{Tr} \{ \Pi_m \rho_\beta \}, \quad (6.2)$$

and the resulting probability distribution encodes the system temperature as a statistical parameter. The second step in estimating the temperature is to construct an estimator T_{est} . A general prescription for doing this does not exist [62]. However, it can be shown that for any unbiased estimator the variance is lower bounded through the Cramer-Rao inequality $\delta T_{\text{est}}^2 \geq 1/\nu \mathcal{F}_T$ [76], where ν is the number of independent measurement rounds and

$$\mathcal{F}_T := \sum_{m=1}^N p_{m;\beta} [\partial_T \ln p_{m;\beta}]^2, \quad (6.3)$$

is the Fisher information. We note that the Cramer-Rao inequality is asymptotically tight for Bayesian or maximum likelihood estimators [62]. Throughout, motivated by the Cramer-Rao inequality, we adopt the Fisher information as the quantifier of precision.

Identifying measurement strategies for which the temperature estimate can achieve minimal variance corresponds to maximizing the Fisher information over all possible measurements (POVMs). This results in a measurement-independent quantity, the quantum Fisher information \mathcal{F}_T^Q [48]. Within the canonical ensemble, it can be shown that a projective measurement of the system energy is optimal [83, 51]. The quantum Fisher information is then related to the variance of the system energy

$$T^4 \mathcal{F}_T^Q = \langle H^2 \rangle - \langle H \rangle^2, \quad (6.4)$$

where $\langle O \rangle = \text{Tr} \{ O \rho_\beta \}$. This expression provides a fundamental upper bound on the attainable value of the Fisher information for any measurement at any temperature. As a

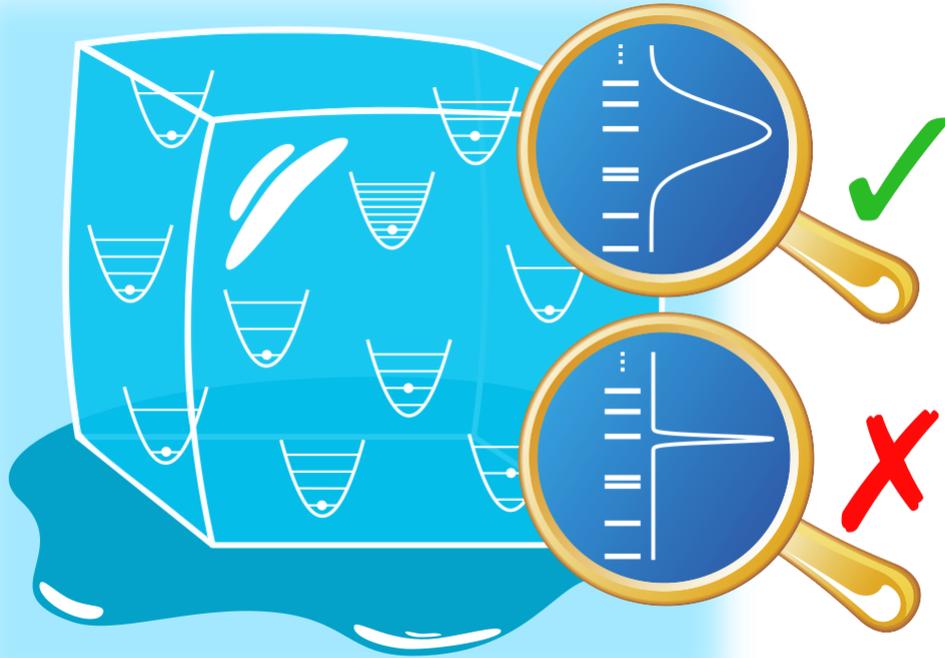


Figure 6.1: *This figure and caption is taken from Ref. [88] ©2021 American Physical Society. The chapter is self-contained.* Finite measurement resolution is interpreted as an inability to sharply distinguish between consecutive system energy eigenstates and results in a non-trivial constraint on the attainable thermometric precision. For a macroscopic system with an effectively continuous energy spectrum, any measurement is subject to finite resolution and thus limited by the bound in Eq. (6.27).

consequence of the third law of thermodynamics, or more explicitly the assumption that the heat capacity vanishes at zero temperature, the variance of the system energy must vanish at least quadratically in temperature as absolute zero is approached [51]. Hence it follows that $T^2 \mathcal{F}_T^Q$ must vanish in the low-temperature limit, and that the relative error $\delta T_{\text{est}}^2 / T^2$ must diverge by virtue of the Cramer-Rao inequality. This relation constitutes the ultimate bound on the optimal low-temperature scaling behaviour of the Fisher information, applicable for any system and for any measurement strategy.

6.2.2 Accounting for measurement limitations

In many settings of interest, it is not realistic to implement a projective measurement of the system energy. For instance, whenever the gaps in the energy spectrum are below the energy resolution of the available measurement [91], which happens, e.g., when the system is large enough to appear continuous while the measurement apparatus has a finite size, or whenever only a finite part of the full system can be interacted with within a finite time (see Fig. 6.1). Under such conditions of constrained experimental access,

it is useful to introduce the POVM energies [51]

$$E_{m;\beta} := \frac{1}{p_{m;\beta}} \text{Tr} \{ \Pi_m H \rho_\beta \}, \quad (6.5)$$

where $E_{m;\beta}$ may be interpreted as the best guess of the system energy before the measurement, given that outcome m was observed [51]. In the case of projective energy measurements on the system, the POVM energies coincide with the system energy eigenvalues. In general however, the POVM energies are temperature dependent.

For convenience we may identify a specific POVM energy $E_{0;\beta}$, defined as the smallest POVM energy in the low-temperature limit. We can then introduce the POVM energy gaps $\Delta_{m;\beta} := E_{m;\beta} - E_{0;\beta}$, which by definition are non-negative at low temperatures. In terms of these gaps, the Fisher information for a general measurement is given by

$$\mathcal{F}_T = \frac{\sum_m p_{m;\beta} \Delta_{m;\beta}^2 - (\sum_m p_{m;\beta} \Delta_{m;\beta})^2}{T^4}. \quad (6.6)$$

Similarly to the quantum Fisher information, the above expression takes the form of an energy variance. However for general measurements the energy spectrum of the system is replaced by the spectrum of POVM energies, and the Boltzmann probabilities associated with projective energy measurements are replaced by the POVM probabilities. These changes incorporate restrictions due to limitations of the measurement on top of those imposed by the system itself.

In investigating the scaling behaviour we are implicitly assuming that the Fisher information is a continuous function of temperature, which implies that the POVM energy gaps $\Delta_{m;\beta}$ must also be continuous functions. Following Ref. [51], we are going to study the scaling behaviour of the Fisher information when the POVM energy gaps have a well-defined power-series expansion in temperature around absolute zero

$$\Delta_{m;\beta} = \Delta_{m,0} + \sum_{k=1}^{\infty} \Delta_{m,k} \beta^{-k}. \quad (6.7)$$

By virtue of Weierstrass' approximation theorem, any continuous function can be approximated arbitrarily well by such a power series [146]. Note that this formulation does not exclude the case of projective energy measurements as this would be described by a series with only the constant term. For more general measurements, however, the expansion might contain non-zero higher-order coefficients.

Following Potts et al. [51] we can make use of the relation between the POVM energies and the associated probabilities (Eq. (6.5)) to write $\Delta_{m;\beta} = -\partial_\beta \ln p_{m;\beta} / p_{0;\beta}$. Given the power-series expansion of the POVM energy gaps, we can integrate this equation and express the ratio of the probabilities for outcomes m and 0 as

$$\frac{p_{m;\beta}}{p_{0;\beta}} = g_m e^{-\beta \Delta_{m,0}} \beta^{-\Delta_{m,1}} \prod_{k=1}^{\infty} e^{\Delta_{m,k+1} \beta^{-k}/k}, \quad (6.8)$$

where g_m is a temperature-independent integration constant. We stress that as a consequence of how we defined $E_{0;\beta}$, the probability $p_{0;\beta}$ is the largest probability at zero temperature and must be non-vanishing in this limit. We thus obtain an expression for the probability of obtaining outcome m given fully in terms of the expansion coefficients of the corresponding POVM energy gap (note that the explicit dependence on $p_{0;\beta}$ could be avoided by using the fact that the full distribution must be normalised).

6.2.3 Low-temperature scaling behaviour

The above model of limited measurements allows us to obtain, by substituting Eqs. (6.7) and (6.8) into Eq. (6.6), an expression for the Fisher information given fully in terms of the POVM energy gaps. Based on this, we can analyse the possible scaling behaviour of the Fisher information, as the system approaches zero temperature. First of all, we note that Eq. (6.6) can be rewritten as

$$\mathcal{F}_T = \frac{1}{2T^2} \sum_{m,n} p_{m;\beta} p_{n;\beta} (\beta\Delta_{m;\beta} - \beta\Delta_{n;\beta})^2. \quad (6.9)$$

Notice that all terms on the right-hand side are positive, and because of this the scaling behaviour of the Fisher information is determined by the term in the sum (or the set of terms) which vanishes least rapidly as the temperature goes to zero. We now consider the scaling that arises from different terms in Eq. (6.9). We focus on terms that result in sub-exponential scalings, referring the reader to Ref. [51] for a discussion of the remaining terms.

For convenience, we define the ground-state set of measurement outcomes, as those for which the probability of obtaining that outcome remains finite at zero temperature (note that the outcome $m = 0$ is in the ground-state set by definition). From Eq. (6.8), we see that formally this set can be defined as $\Omega = \{m \mid \Delta_{m,0} = \Delta_{m,1} = 0\}$. Now consider those terms in the Fisher information above where both outcomes belong to the ground-state set. To leading order in temperature, the contribution from these terms takes the form

$$\frac{1}{2T^2} \sum_{m,n \in \Omega} p_{m;\beta} p_{n;\beta} (\Delta_{m,j} - \Delta_{n,j})^2 T^{2(j-1)}, \quad (6.10)$$

where j labels the lowest order for which the expansion coefficient of any element in the ground-state set is non-zero ($j \geq 2$). These terms in the sum thus vanish at least quadratically, giving at best a constant contribution to the Fisher information. Notice that if the ground-state set contains only a single outcome ($m = 0$), then the contribution is identically zero.

Next we consider the terms in the Fisher information where one of the outcomes belong to the ground-state set but the other one does not. To this end, we define the set of outcomes $\tilde{\Omega} = \{m \mid \Delta_{m,0} = 0 \text{ and } \Delta_{m,1} \neq 0\}$, for which the associated probability vanish sub-exponentially as the temperature goes to zero. The set of outcomes $\tilde{\Omega}$ has an associated POVM energy coinciding with that of the ground-state set at zero-temperature, but exhibits a linear degeneracy splitting at finite temperature. To leading order in temperature, the contribution from the corresponding terms is

$$\frac{1}{T^2} \sum_{m \in \tilde{\Omega}} g_m \Delta_{m,1}^2 T^{\Delta_{m,1}}, \quad (6.11)$$

which vanishes at a rate determined by the the first-order expansion coefficients $\Delta_{m,1}$. It is straightforward to show that all other contributions vanish exponentially in the low-temperature limit.

The (sub-exponential) low-temperature behaviour of the right-hand side of the Fisher information (6.9), is generally given by the sum of Eq. (6.10) and Eq. (6.11). Which of

these two dominate depends on the smallest first-order expansion coefficient. If the set $\tilde{\Omega}$ is not empty, and at least one element in the set has a value $\Delta_{m,1} < 2(j-1)$, where j denotes the lowest order with non-vanishing expansion coefficient within the ground-state set, then the low-temperature behaviour of the Fisher information is captured by

$$\mathcal{F}_T = \sum_{m \in \tilde{\Omega}} g_m \Delta_{m,1}^2 T^{\Delta_{m,1}-2}. \quad (6.12)$$

In principle the first-order coefficient can take any positive value without violating the scaling bound imposed by the third law of thermodynamics (ensuring divergence of the relative error). Notice that even a divergent low-temperature behaviour of the Fisher information can in principle be realised, if $\Delta_{m,1}$ can take a value smaller than 2.

6.3 Scaling bound for large systems

In this section, we propose a finite-resolution criterion characterizing realistic measurements. We aim to capture any situation in which the available measurements cannot resolve the individual gaps in the system energy spectrum, which therefore appears continuous. Below, we make this statement precise. We then go on to show how this criterion leads to a lower bound on the first-order coefficient $\Delta_{m,1}$, constraining the low-temperature scaling of the error in any temperature estimation scheme. Furthermore we present an example of a measurement saturating the finite-resolution bound, showing that the bound is tight.

6.3.1 Finite-resolution criterion

In the regime where the system has an effectively continuous energy spectrum (as the measurement only resolves energy differences much larger than the gaps in the spectrum, it is convenient to work with the system density of states $\mathcal{D}(\epsilon) := \sum_k d_k \delta(\epsilon - \epsilon_k)$, where the sum is over distinct system energy eigenvalues and d_k is the corresponding degeneracy. Throughout, we adopt the convention that the smallest system energy eigenvalue is set to zero ($\epsilon_0 = 0$).

Now, we introduce a filtered density of states \mathcal{D}_m for each measurement outcome m , as the system density of states filtered through the corresponding POVM element

$$\mathcal{D}_m(\epsilon) := \sum_k d_k \delta(\epsilon - \epsilon_k) \text{Tr} \left[\Pi_m \frac{\mathbb{1}_{\epsilon_k}}{d_k} \right], \quad (6.13)$$

where $\mathbb{1}_{\epsilon_k}$ is the projection operator onto the eigenspace with energy ϵ_k . Notice that the sum of all the filtered densities of states adds up to the total density of states. Furthermore, we introduce the continuous filter function $f_m(\epsilon)$, formally defined by the values $f_m(\epsilon_k) = \text{Tr} [\Pi_m \mathbb{1}_{\epsilon_k} / d_k]$ and the straight-line segments connecting these values. In addition we note that the density of states can be expressed as the rate of change of the number of states with energy below ϵ $\sigma(\epsilon) = \sum_k d_k \theta(\epsilon - \epsilon_k)$, where θ denotes the Heaviside step function. Given these, the filtered density of states decompose into the product

$$\mathcal{D}_m(\epsilon) = f_m(\epsilon) \frac{d\sigma(\epsilon)}{d\epsilon}, \quad (6.14)$$

where the filter function fully characterizes the implemented measurement. Importantly we notice that the function $\sigma(\epsilon)$ is non-decreasing for all energies. If we compute the Laplace transform in β of the filtered density of states, the result takes the form of a Stieltjes integral over a measure given by $\sigma(\epsilon)$ [147]

$$\hat{\mathcal{D}}_m(\beta) := \int_0^\infty d\sigma(\epsilon) f_m(\epsilon) e^{-\beta\epsilon} = \mathcal{Z}_\beta p_{m;\beta}. \quad (6.15)$$

The last equality can be obtained directly from equation (6.2), and relates the Laplace transformed filtered density of states to the product of the probability and the canonical partition function. Notice that the measure $\sigma(\epsilon)$ is a discontinuous function of energy.

For macroscopic systems the measure can often be approximated by an effective continuous measure, when $\sigma(\epsilon)$ and $f_m(\epsilon)$ vary on widely separated energy scales. To see this, we first define the averaged measure with respect to an energy window ω by

$$\sigma_\omega(\epsilon) = \theta(\epsilon) \frac{1}{\omega} \int_{\epsilon-\omega/2}^{\epsilon+\omega/2} ds \sigma(s), \quad (6.16)$$

which for non-zero ω is a continuous function of energy except at $\epsilon = 0$, and which tends to a differentiable function of energy as ω is increased. The inclusion of the step function at zero energy is important if we are to capture the zero temperature limit correctly, since it ensures that the ground-state of the averaged model coincides with that of the exact model. For the purposes of low-temperature thermometry only the low-energy behaviour is of importance, and to leading order in energy we adopt an effective measure given by

$$d\sigma_\omega(\epsilon) = [d_{0;\omega}\delta(\epsilon) + \alpha_\omega\gamma_\omega\epsilon^{\gamma_\omega-1} + \mathcal{O}(\epsilon^{\gamma_\omega})] d\epsilon, \quad (6.17)$$

where $d_{0;\omega}$ is an effective ground-state degeneracy and $\alpha_\omega, \gamma_\omega$ are positive, real-valued constants. The coefficient $\gamma_\omega > 0$ characterizes the low-energy growth in the total number of states with energy less than ϵ .

If we compute the Laplace transform with respect to this averaged measure (which now takes the form of a standard Riemann integral) we obtain to leading order in energy

$$\begin{aligned} \hat{\mathcal{D}}_{m;\omega}(\beta) &= d_{0;\omega} f_m(0) \\ &+ \alpha_\omega \gamma_\omega \int_0^\infty d\epsilon \epsilon^{\gamma_\omega-1} f_m(\epsilon) e^{-\beta\epsilon}. \end{aligned} \quad (6.18)$$

The averaged measure tends to overestimate the number of low-energy states as ω is increased, however this effect becomes negligible in the limit $\omega \ll T$. Now if we assume that $f_m(\epsilon)$ does not vary significantly across an energy range ω , then $\hat{\mathcal{D}}_m(\beta)$ is well approximated by the averaged function $\hat{\mathcal{D}}_{m;\omega}(\beta)$. More quantitatively we can state this condition in the form of an inequality

$$\frac{|f_m(\epsilon + \omega) - f_m(\epsilon)|}{\omega} \ll \frac{1}{\omega}, \quad (6.19)$$

which bounds the rate of change of the filter function with energy. For macroscopic systems we can take the limit $\omega \rightarrow 0$, and in this case we are going to adopt the following *finite-resolution criterion (FRC)*:

FRC: *In the limit of a macroscopic system, the filter function $f_m(\epsilon)$ tends to a continuous, right-differentiable function of the system energy.*

This is nothing more than a restatement of equation (6.19) for vanishing ω , which restricts the rate of change of the filter function to a finite value. We note that at $\epsilon = 0$, the filter function may be discontinuous and Eq. (6.19) tends to the right derivative for $\omega \rightarrow 0$.

6.3.2 Finite-resolution bound

Having characterized what we mean by a finite-resolution measurement, we ask what the consequences of our finite-resolution criterion are for the behaviour of the POVM energy gaps in the macroscopic limit. By making use of equation (6.8), we obtain the relation (we now drop the dependence on the energy window ω and write simply d_0, α and γ)

$$\hat{\mathcal{D}}_m(\beta) = \hat{G}_m(\beta)\hat{\mathcal{D}}_0(\beta), \quad (6.20)$$

where for convenience we have defined the transfer function

$$\hat{G}_m(\beta) := g_m e^{-\beta\Delta_{m,0}} \beta^{-\Delta_{m,1}} \prod_{k=1}^{\infty} e^{\Delta_{m,k+1}\beta^{-k}/k}. \quad (6.21)$$

Now this is a relation at the level of the Laplace-transformed, filtered densities of states. We can obtain a relationship directly between the filtered densities of states by taking the inverse Laplace transform of both sides of Eq. (6.20). By applying the Laplace convolution theorem [148, 149], we derive the relation

$$\mathcal{D}_m(\epsilon) = \int_0^\epsilon ds G_m(\epsilon - s)\mathcal{D}_0(s). \quad (6.22)$$

We now focus on the specific case of $m \in \tilde{\Omega}$. For these outcomes, the inverse Laplace transform can be computed straightforwardly, and to leading order in energy we obtain

$$G_m(\epsilon) = \frac{g_m}{\Gamma(\Delta_{m,1})} \epsilon^{\Delta_{m,1}-1} + \mathcal{O}(\epsilon^{\Delta_{m,1}}), \quad (6.23)$$

where $\Gamma(\Delta_{m,1})$ denotes the Gamma function [148]. As we saw in the preceding section, the outcomes within $\tilde{\Omega}$ are exactly the ones with potential to provide optimal low-temperature scaling of the Fisher information.

Recall, that the reference outcome $m = 0$, was chosen such that the associated probability approaches a constant value at zero temperature. This implies that the overlap of the POVM element Π_0 with the system ground state is non-zero, and therefore $f_0(0)$ is non-zero. On the other hand for outcomes $m \in \tilde{\Omega}$ the probability vanishes in the low-temperature limit, implying a vanishing overlap $f_m(0) = 0$. Hence in this case we find from equations (6.18) and (6.22) that to leading order in energy

$$f_m(\epsilon) = \frac{g_m d_0 f_0(0)}{\alpha \gamma \Gamma(\Delta_{m,1})} \epsilon^{\Delta_{m,1}-\gamma} + \mathcal{O}(\epsilon^{\Delta_{m,1}+1-\gamma}). \quad (6.24)$$

Based on this expression we can infer constraints on the linear coefficient. First, the requirement that $f_m(0) = 0$ gives the weakest constraint $\Delta_{m,1} > \gamma$. This simply expresses the fact that the Fisher information is upper bounded by the the quantum Fisher

information, which scales as $T^{\gamma-2}$ for a density of states scaling as $\epsilon^{\gamma-1}$. Further, the finite-resolution criterion restricts the rate of change to be bounded, $\frac{d}{d\epsilon}f_m(\epsilon) < \infty$. This implies a tightened scaling bound

$$\Delta_{m,1} \geq 1 + \gamma, \quad \text{for } m \in \tilde{\Omega}. \quad (6.25)$$

Since $\gamma > 0$ by definition, this implies that the Fisher information must grow slower than $1/T$, i.e.,

$$\lim_{T \rightarrow 0} T \mathcal{F}_T = 0. \quad (6.26)$$

Further note that a diverging Fisher information in the low-temperature limit can only be realized through a $\sigma(\epsilon)$ that grows sub-linearly with energy, i.e., $\gamma < 1$. As an example of a system exhibiting such a sub-linear growth we mention systems of massive non-interacting particles at zero chemical potential [51]. For such systems $\gamma = 1/2$ for one-dimensional geometries.

By virtue of the Cramer-Rao bound, Eq. (6.26) implies that the absolute error (squared) must vanish more slowly than T

$$\lim_{T \rightarrow 0} \frac{\delta T_{\text{est}}^2}{T} = \infty. \quad (6.27)$$

The equivalent Eqs. (6.26) and (6.27) constitute the main result of our paper. They imply that for an effectively continuous spectrum, the low-temperature scaling of the precision is not only bounded by the third law, which demands a diverging relative error, but by a tighter bound. Interestingly, our bound still allows for a vanishing absolute error, a scenario that can be physically realized as illustrated below.

6.3.3 Proving tightness of bound

We now illustrate that the proposed finite-resolution bound is tight. Consider a binary measurement which resolves the system ground state exponentially well in the sense that it has POVM elements

$$\Pi_0 = e^{-\kappa H}, \quad \Pi_1 = \mathbb{1} - e^{-\kappa H}, \quad (6.28)$$

where $\kappa > 0$. Note that the overlap of Π_0 with the system energy eigenstates decays exponentially away from zero. This feature makes it straightforward to write down the filtered density of states. Focusing on $m = 1$ we find

$$\mathcal{D}_1(\epsilon) = [1 - e^{-\kappa\epsilon}] \mathcal{D}(\epsilon), \quad (6.29)$$

where nothing has been assumed about the form of the system density of states. We thus see that the corresponding filter function takes the form $f_1(\epsilon) = \kappa\epsilon + \mathcal{O}(\epsilon^2)$ to leading order in energy. If we adopt the density of states introduced in the preceding subsection, that is $\mathcal{D}(\epsilon) = d_0\delta(\epsilon) + \alpha\gamma\epsilon^{\gamma-1} + \mathcal{O}(\epsilon^\gamma)$, then upon comparison with equation (6.24) we find $\Delta_{1,1} = 1 + \gamma$. Hence the binary exponential resolution measurement saturates the finite-resolution bound.

For good measure we now show how the same conclusion can be derived directly from the probabilities. The probability of obtaining outcome $m = 0$ can be written in terms of the system partition function as

$$p_{0;\beta} = \mathcal{Z}_{\kappa+\beta} \mathcal{Z}_\beta^{-1}. \quad (6.30)$$

Substituting the probabilities $p_{0;\beta}$ and $p_{1;\beta} = 1 - p_{0;\beta}$ into the general form of the Fisher information (Eq. 6.3), one finds that

$$T^4 \mathcal{F}_T = \frac{\mathcal{Z}_{\kappa+\beta}}{\mathcal{Z}_\beta - \mathcal{Z}_{\kappa+\beta}} \left(\langle H \rangle_\beta - \langle H \rangle_{\kappa+\beta} \right)^2. \quad (6.31)$$

The partition function is given by the Laplace transform of the density of states, hence we find $\mathcal{Z}_\beta = d_0 \exp(\alpha\gamma\Gamma(\gamma)\beta^{-\gamma}/d_0)$ (in App. 6.7.1 we show how this form of the partition function describes a system of non-interacting bosonic modes). From this form of the partition function we can derive the low-temperature behaviour of the average energy

$$\langle H \rangle_\beta = \frac{\alpha\gamma^2\Gamma(\gamma)}{d_0} \beta^{-(1+\gamma)}. \quad (6.32)$$

If we substitute these into the above Fisher information, we find that to leading order in temperature (assuming that $\kappa/\beta \ll 1$)

$$\mathcal{F}_T = \alpha\kappa\gamma^2\Gamma(\gamma)(1+\gamma)^2 T^{\gamma-1} + \mathcal{O}(T^{2\gamma}), \quad (6.33)$$

which takes the form of Eq. (6.12) with $\Delta_{1,1} = 1 + \gamma$ and $g_1 = \alpha\kappa\gamma^2\Gamma(\gamma)$. Since γ can in principle take any positive value, the exponential-resolution measurement saturates the finite-resolution bound and asymptotically attains a Fisher information scaling as $1/T$ in the limit $\gamma \rightarrow 0$.

6.4 Generalization to noisy measurements

In this section we extend the thermometry framework above to include noisy measurements. As the framework is general, one might ask if noise effects are not already accounted for. The answer is that in principle noise effects are described. However, for some noisy measurements, the POVM energy gap does not have a Taylor expansion. While one may still approximate the energy gap by a polynomial, a physically appealing extension of the formalism allows for circumventing this approximation. We find that our bound given in Eq. (6.27) also holds for noisy measurements.

6.4.1 Noisy temperature measurements

To model noisy measurements, we consider the case where the observed outcomes m correspond to coarse graining over a fine-grained POVM with elements $\Pi_{m\mu}$. The probability of observing m is then

$$p_{m;\beta} = \sum_{\mu} p_{m\mu;\beta} = \sum_{\mu} \text{Tr} \{ \Pi_{m\mu} \rho_\beta \}. \quad (6.34)$$

Physically this could correspond to a measurement implemented using a sensor, where only a subset of the sensor degrees of freedom (or a subspace of the full sensor Hilbert space) is experimentally accessible. If we were to compute the Fisher information directly using the fine-grained distribution $p_{m\mu}$, we recover the noiseless results, and obtain an upper bound on the Fisher information computed from the coarse-grained distribution. This fact follows directly from the relation between the relative entropy of two

probability distributions differing by an infinitesimal temperature δT and the Fisher information

$$\mathcal{S}(p_T \| p_{T+\delta T}) = \mathcal{F}_T \delta T^2 \quad \text{as } \delta T \rightarrow 0. \quad (6.35)$$

Since the relative entropy is monotonically decreasing under coarse-graining [150], we conclude that noise always reduces the Fisher information.

The question we now address is, how it impacts the attainable scaling with temperature. Following the approach developed above, we introduce the fine-grained POVM energies

$$E_{m\mu;\beta} := \frac{1}{p_{m\mu;\beta}} \text{Tr} \{ \Pi_{m\mu} \rho_\beta \}, \quad (6.36)$$

which may be interpreted as the best guess of the system energy before the measurement, given the outcome (m, μ) [51]. For convenience we identify the smallest POVM energy in the low-temperature limit with the outcome $E_{00;\beta}$, and then define the fine-grained POVM energy gap $\Delta_{m\mu;\beta} := E_{m\mu;\beta} - E_{00;\beta}$, which by definition is non-negative at low temperatures. Modelling the fine-grained POVM energy gaps by a power-series expansion around zero temperature as in Eq. (6.7), we are led to a probability distribution identical to (6.8), but with m replaced by the compound index $m\mu$.

Since the Fisher information is not defined with respect to the fine-grained probabilities, but rather with respect to the coarse-grained probabilities, the relevant energies are the coarse-grained POVM energy gaps defined by

$$\Delta_{m;\beta}^{(c)} := \sum_{\mu} \frac{p_{m\mu;\beta}}{p_{m;\beta}} \Delta_{m\mu;\beta}. \quad (6.37)$$

In terms of these, the Fisher information can be written in the same form as the fine-grained Fisher information of Eq. (6.9), but with the fine-grained probability and the fine-grained POVM energy gaps replaced by their coarse-grained versions

$$\mathcal{F}_T = \frac{1}{2T^2} \sum_{m,n} p_{m;\beta} p_{n;\beta} \left(\beta \Delta_{m;\beta}^{(c)} - \beta \Delta_{n;\beta}^{(c)} \right)^2. \quad (6.38)$$

Notice that all terms in the sum are positive. Hence, the scaling behaviour of the Fisher information is determined by the term (or set of terms) which vanishes least rapidly as the temperature approaches zero.

From Eq. (6.37), we can anticipate that fine-grained energy gaps that have a Taylor expansion may result in coarse-grained gaps that do not. This may result in qualitatively different behaviour of the fine- and coarse-grained Fisher information. In particular, noise may render the scaling of the Fisher information worse. In appendix 6.7.2 we discuss in general terms how noise impacts the attainable Fisher information scaling. In particular, we show that the noise can never result in a better scaling for the Fisher information, implying that the bound given in Eq. (6.27) also holds for noisy measurements. Here we illustrate the effect of noise with an example.

6.4.2 Illustration of noisy measurement

A simple example illustrating noise is obtained by adding white noise to the binary, exponential-resolution measurement of Sec. 6.3.3. That is, we study a binary POVM

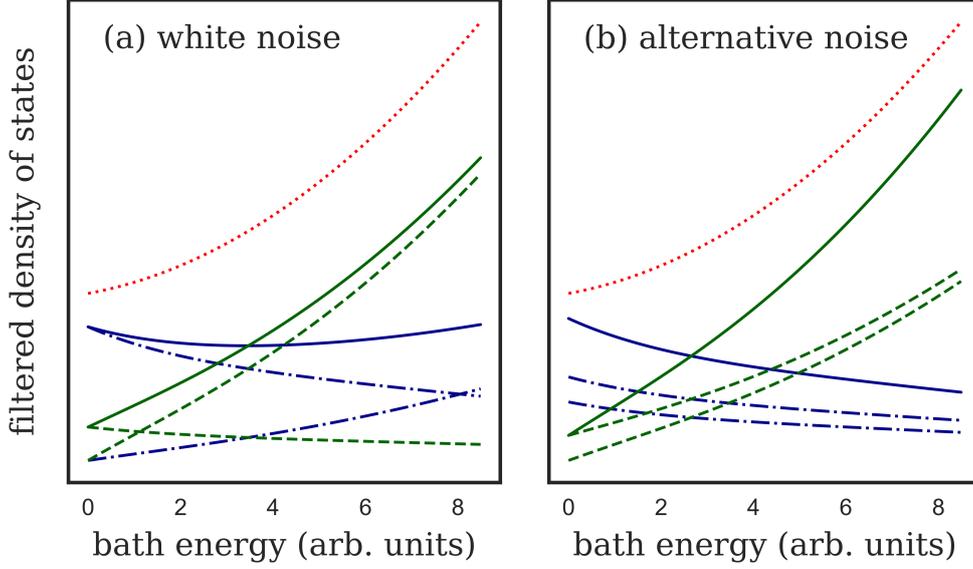


Figure 6.2: *This figure and caption is taken from Ref. [88] © 2020 American Physical Society. The chapter is self-contained.* Illustration of filtered density of state for a noisy binary exponential resolution measurement using $\mathcal{D}(\epsilon) = \mathcal{L}^{-1}[\exp(\alpha\beta^{-1})]$ (dotted red line) with $\alpha = 0.2$. (a) The white noise measurement corresponds to swapping the observed measurement outcomes with some probability, such that each coarse-grained outcome has contributions both from elements within and elements not within the ground-state set. The dashed green lines gives \mathcal{D}_{00} and \mathcal{D}_{01} (their sum is shown with the solid green line), and the blue dashed-dotted lines correspond to elements \mathcal{D}_{10} and \mathcal{D}_{11} (with their sum given by the solid blue line). (b) In App. 6.7.2 we show that an alternative noise model consists of a mixing of several similar measurement outcomes. In the specific case depicted here, the fine-grained outcomes to be summed are almost identical except for projecting onto slightly different energy distributions.

defined by $\Pi_0 = \eta \exp(-\kappa H) + (1 - \eta)\mathbb{1}/2$. To understand how this noise model arises from coarse graining a fine-grained measurement, we consider the fine-grained POVM

$$\begin{aligned} \Pi_{00} &= \frac{1 + \eta}{2} e^{-\kappa H}, & \Pi_{01} &= \frac{1 - \eta}{2} (\mathbb{1} - e^{-\kappa H}), \\ \Pi_{10} &= \frac{1 + \eta}{2} (\mathbb{1} - e^{-\kappa H}), & \Pi_{11} &= \frac{1 - \eta}{2} e^{-\kappa H}, \end{aligned} \quad (6.39)$$

such that $\Pi_0 = \Pi_{00} + \Pi_{01}$ and $\Pi_1 = \Pi_{10} + \Pi_{11}$. As in the noiseless case, we suppose that the average energy exhibits a power-law behaviour $\langle H \rangle_\beta = \alpha\beta^{-(1+\gamma)}$ at low temperatures in the macroscopic limit, with α and γ both positive. The corresponding partition function (at low temperatures) is then $\mathcal{Z}_\beta = \exp(\alpha\beta^{-\gamma}/\gamma)$. For the fine-grained measurement outcomes, we find that to leading order in temperature (assuming that $\kappa/\beta \ll 1$ and $\eta < 1$), the POVM energy gaps with respect to the reference $E_{00;\beta}$, take the form

$$\begin{aligned} \Delta_{00;\beta} &= \Delta_{11;\beta} = 0, \\ \Delta_{10;\beta} &= \Delta_{01;\beta} = (1 + \gamma)T + \mathcal{O}(T^{2+\gamma}). \end{aligned} \quad (6.40)$$

We see that the fine-grained measurement outcomes have an associated set of POVM energy gaps that have a Taylor series in the low-temperature limit. Furthermore, they exhibit a linear degeneracy splitting. It then follows from Eq. (6.12) that the Fisher information takes the form

$$\mathcal{F}_T = \alpha\kappa(1 + \gamma)^2 T^{\gamma-1} + \mathcal{O}(T^{2\gamma}), \quad (6.41)$$

which is equivalent to the noiseless form found above [cf. Eq. (6.33)]. Notice that when having access to the fine-grained distribution, both the POVM energies and the resulting Fisher information is independent of the parameter η quantifying the amount of white noise.

The picture changes when considering the coarse-grained energy gap (Eq. (6.37)). To leading order in temperature this is given by

$$\Delta_{1;\beta}^{(e)} = \frac{1 + \eta}{1 - \eta} \alpha\kappa(1 + \gamma) T^{2+\gamma} + \mathcal{O}(T^{3+2\gamma}). \quad (6.42)$$

Notice that in contrast to the fine-grained energy gaps, this coarse-grained gap does not have a Taylor expansion. Computing the Fisher information over the coarse-grained gaps and probabilities (making use of Eq. (6.38)) gives

$$\mathcal{F}_T = \frac{4\eta^2}{1 - \eta^2} (\alpha\kappa(1 + \gamma))^2 T^{2\gamma} + \mathcal{O}(T^{1+3\gamma}). \quad (6.43)$$

This example thus illustrates how noise can result in a coarse-grained gap that has no Taylor expansion and how this may result in a different (worse) scaling for the Fisher information at low-temperatures. Qualitatively we can understand the altered scaling by studying the coarse-grained filtered density of states. For the example considered here we have

$$\begin{aligned} \mathcal{D}_{00}(\epsilon) &= f_{00}(\epsilon) \mathcal{D}(\epsilon) = \frac{1 + \eta}{2} e^{-\kappa\epsilon} \mathcal{D}(\epsilon), \\ \mathcal{D}_{01}(\epsilon) &= f_{01}(\epsilon) \mathcal{D}(\epsilon) = \frac{1 - \eta}{2} (1 - e^{-\kappa\epsilon}) \mathcal{D}(\epsilon), \end{aligned} \quad (6.44)$$

and under coarse-graining these are added together. Notice that whereas the filter function $f_{01}(\epsilon)$ goes to zero as $\epsilon \rightarrow 0$, this is not true of $f_{00}(\epsilon) + f_{01}(\epsilon)$ (the same feature is found for the $m = 1$ outcomes). Hence in this case the noise removes outcomes from the set $\tilde{\Omega}$, resulting in the worse scaling (note that a vanishing filter function at $\epsilon = 0$ implies a vanishing probability at $T = 0$ and vice versa, cf. Eq. (6.15)). This effect is illustrated in Fig. 6.2a. In App. 6.7.2 we study an alternative noise model. In this model each coarse-grained outcome can be seen as the sum of several similar (in the sense of preparing similar energy distributions) fine-grained outcomes. This is illustrated in Fig. 6.2b.

The noisy framework put forward here shows that our finite-resolution bound, as well as the results of Ref. [51] apply for any POVM that can be written as a coarse graining over a fine-grained POVM which has a spectrum with a well defined Taylor series. As the coarse-grained POVM itself may not have a spectrum with a well defined Taylor series, this extends the applicability of the results of Ref. [51] (as long as we do not want to rely on approximate Taylor series in the spirit of the Weierstrass theorem).

6.5 Single-qubit probe

6.5.1 Measurement protocol

We now illustrate our results by considering temperature estimation of a system of non-interacting bosons using a single qubit as a probe. The system is described by a spectrum of single-particle energies ω_k (we take $\hbar = 1$). Consider the following measurement strategy: (i) first we initialise the probe qubit in its ground state $|0\rangle$, (ii) then an interaction is turned on between the probe and the system for a short time t , and (iii) we perform a projective measurement of the qubit energy. Given this protocol, the probability of finding the qubit in the excited state $|1\rangle$ is

$$p_{1;\beta} = \text{Tr} \left\{ \langle 0| U_t^\dagger |1\rangle \langle 1| U_t |0\rangle \rho_\beta \right\}. \quad (6.45)$$

We take the time-evolution operator U_t to be generated by a time-independent Hamiltonian

$$H = \sum_k \omega_k a_k^\dagger a_k + \frac{\Omega}{2} \sigma_z + H_{\text{int}}, \quad (6.46)$$

where a_k^\dagger, a_k denotes the bosonic creation and annihilation operators. The probe qubit is characterised by the three Pauli operators $\{\sigma_x, \sigma_y, \sigma_z\}$, and we take the probe energy to be proportional to the σ_z operator.

Computing the outcome probabilities requires specifying an interaction Hamiltonian and determining the resulting dynamics. This task is complicated by the fact that the low-temperature and short-time regime is generally not accessible via standard Markovian master equations [151, 19]. However, if the interaction time is sufficiently short we can make analytical progress by approximating the probability up to second order in t . In this case we find that

$$p_{1;\beta} = t^2 \text{Tr} \left\{ \langle 0| H_{\text{int}} |1\rangle \langle 1| H_{\text{int}} |0\rangle \rho_\beta \right\} + \mathcal{O}(t^4). \quad (6.47)$$

We consider a linear interaction Hamiltonian consisting of an excitation-preserving part and a non-excitation-preserving part. Introducing the raising and lowering operators $\sigma_\pm = \frac{1}{2}(\sigma_x \pm i\sigma_y)$ for the probe qubit, the interaction Hamiltonian takes the form

$$H_{\text{int}} = \sum_k g_k \left[\sigma_+ a_k + \sigma_- a_k^\dagger \right] + \sum_k \lambda_k \left[\sigma_- a_k + \sigma_+ a_k^\dagger \right], \quad (6.48)$$

where $\{g_k, \lambda_k\}$ are real-valued coupling coefficients. In the limit of a macroscopic system, these coupling coefficients are taken to approach continuous functions. Physically this means that the interaction cannot selectively probe an individual system mode (ensuring that the finite resolution criterion is satisfied).

Given H_{int} , it becomes straightforward to show from Eq. (6.47) that the excited-state probability at short times takes the form

$$p_{1;\beta} = t^2 \sum_k (g_k^2 + \lambda_k^2) n_\beta(\omega_k) + t^2 \sum_k \lambda_k^2, \quad (6.49)$$

where $n_\beta(\omega_k)$ denotes the Bose-Einstein distribution. We see that the probability consists of two contributions: a temperature-dependent term, in which the probability is directly related to the occupation of the bath modes, and a temperature-independent term. The presence of the temperature-independent term means that the probability of finding the probe qubit in the excited state is generally non-zero even at arbitrarily low temperatures. As in the example in Sec. 6.4.2, this prevents a scaling of the form of Eq. (6.12) and can be captured by our framework for noisy thermometry.

6.5.2 Excitation-preserving interaction

We now focus on the excitation-preserving case ($\lambda_k = 0$), and consider an interaction characterised by a continuous spectral density of the form

$$\rho(\omega) = \sum_k g_k^2 \delta(\omega - \omega_k) = 2\alpha\omega_c^{1-s}\omega^s e^{-\omega/\omega_c}, \quad (6.50)$$

where α is the dissipation strength, s is the ohmicity and ω_c is the cutoff energy [37, 38, 152, 151]. The sum in the excited-state probability (6.49) is then replaced by an integral, which can be solved analytically. In the low-temperature limit we find

$$p_{1;\beta} = 2\alpha (\omega_c t)^2 \Gamma(1+s) \left(\frac{T}{\omega_c}\right)^{1+s} + \mathcal{O}(T^{2+s}). \quad (6.51)$$

We see that this protocol gives a probability vanishing sub-exponentially as the temperature goes to zero, and comparing with the general expression Eq. (6.8), we see that to lowest order, the POVM gap scales as $\Delta_1 = (1+s)T$. The case of an excitation-preserving interaction can thus (for short time at least) be described within our noiseless thermometry framework.

From the value of the linear expansion coefficient, $\Delta_{1,1} = 1+s$, it follows that for ohmicity approaching zero, the finite-resolution bound $\Delta_{1,1} \geq 1$ is approached. The corresponding Fisher information scales as $\mathcal{F}_T \propto T^{s-1}$ and thus diverges for sub-Ohmic baths in the low-temperature limit. This serves as an illustration that the finite-resolution bound is in principle attainable via an excitation-preserving interaction in the short-time limit, and thus the bound is tight. Realising such an excitation-preserving interaction may however be challenging.

6.5.3 Excitation-non-preserving interaction

We now turn to the arguably more realistic excitation non-preserving case. The case $\lambda_k = g_k$ corresponds to the well-known spin-boson model [153, 37, 152, 38]. Adopting the same spectral density as above, the excited-state probability in this case takes the form

$$p_{1;\beta} = 4\alpha (\omega_c t)^2 \Gamma(1+s) \left(\frac{T}{\omega_c}\right)^{1+s} + 2\alpha (\omega_c t)^2 \Gamma(1+s) + \mathcal{O}(T^{2+s}). \quad (6.52)$$

In contrast to the excitation-preserving case, this probability does not in general correspond to the noiseless version of Eq. (6.8) since the POVM energy gap $\Delta_1 \propto T^{s+2}$, does not have a Taylor expansion for arbitrary s at low temperatures. However, as shown

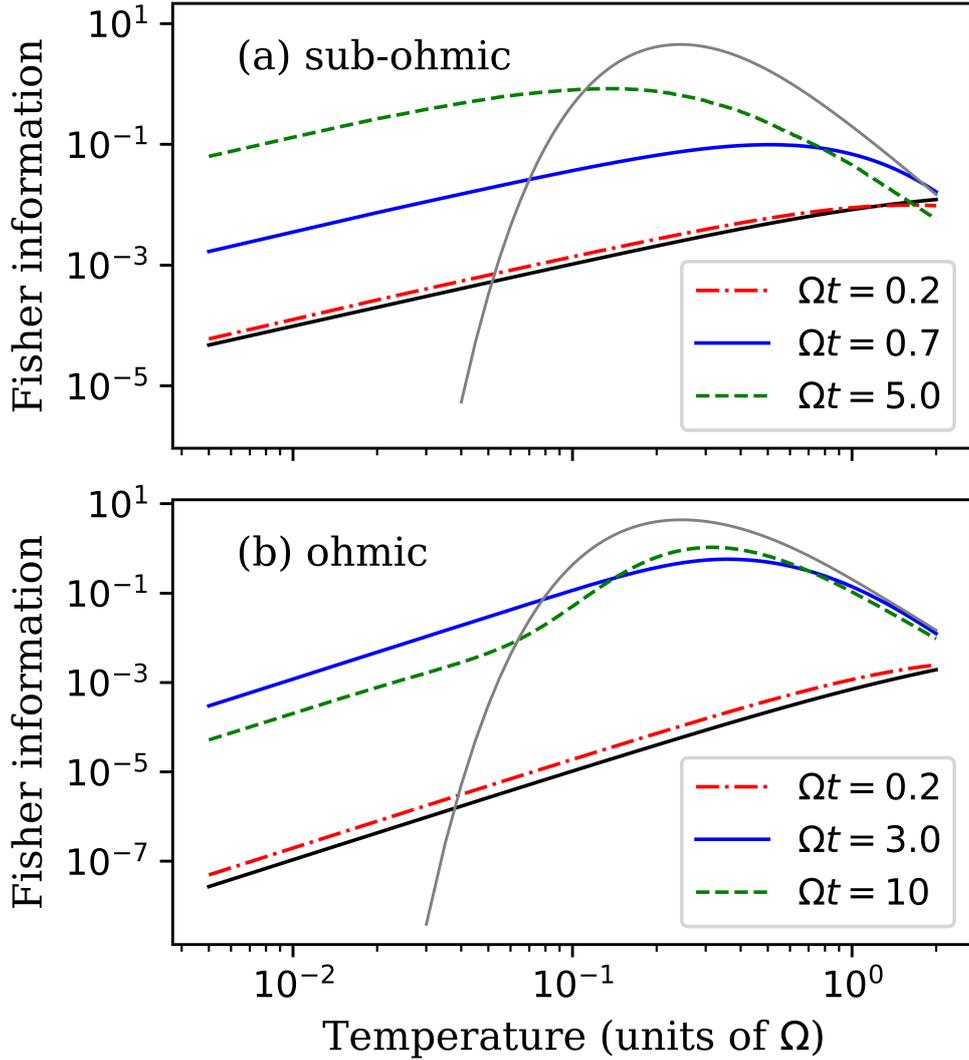


Figure 6.3: *This figure and caption is taken from Ref. [88] ©2020 American Physical Society. The chapter is self-contained.* Numerically computed Fisher information for (a) the sub-Ohmic ($s = 1/2$), and (b) the Ohmic ($s = 1$) spin-boson model, with $\delta t = 0.1/\Omega$, $\alpha = 0.1$ and $\omega_c = 10\Omega$. The solid black lines display the short-time analytical results at time $\Omega t = 0.2$, showing good agreement with the numerical simulations. In case (b) the simulations exhibit a quadratic temperature scaling at low temperatures, while in case (a) a linear scaling is obtained. The solid grey line gives the Fisher information obtained from the steady state of a secular Born-Markov master equation, which scales exponentially at low temperatures [51].

in App. 6.7.3, this scenario can be described using a fine-grained POVM with energy gaps that do have a Taylor expansion. Therefore, the scenario is captured by the noisy framework.

Given the probability (6.52), a short calculation shows that the Fisher information has a low-temperature scaling given by $\mathcal{F}_T \propto T^{2s}$. Again, this is in full agreement with the general noisy theory developed above. Within the spin-boson model, the Fisher

information thus vanishes quadratically for an Ohmic spectral density with $s = 1$, and linearly for a sub-Ohmic spectral density with $s = 1/2$.

To corroborate the analytical results based on the short-time approximation, we turn to a numerical simulation of the Fisher information for the spin-boson model. To perform the simulations we made use of the recently developed tensor-network TEMPO algorithm and its extension to multi-time measurement scenarios [154, 127]. Details of the simulations are provided in App. 6.7.4. Making use of this algorithm has the benefit that the temperature derivative of the excited state can itself be expressed as a tensor network and computed to the same level of accuracy as the probability itself.

Results for the Ohmic and the sub-Ohmic cases are shown in Fig. 6.3. Generally we find that the short-time approximation provides a good description of the observed scaling behaviour at sufficiently short times. Even more interesting we note that the scaling behaviour predicted within the short-time approximation is valid even at times well beyond the regime in which the short-time approximation is expected to hold ($\alpha\delta t^2\Gamma(1+s)\omega_c^2 \ll 1$). This indicates that the predicted precision scaling is experimentally relevant, even without the requirement of being able to probe the non-equilibrium qubit dynamics at very short-times. Notice that the low-temperature Fisher information tends to initially increase with time as information about the environment state is extracted by the qubit. After some time the low-temperature Fisher information starts to decrease. This can be understood as the qubit reaching a stationary state, such that a one-time measurement performed on the qubit can no longer probe the relaxation dynamics induced by the coupling with the thermal bath (see also [50, 19]).

Finally, we note that at sufficiently low temperatures the simulated Fisher information differs from the Markovian result, even for the rather weak coupling and long times considered here. A similar effect was observed in the context of temperature estimation via the Kubo-Martin-Schwinger-like relations obeyed by emission and absorption spectra of multichromophoric systems [155]. There it was pointed out that faithfully recovering the temperature from observed spectra requires taking into account system-environment correlations. This is true even at very low coupling strengths, where these correlations are generally weak.

6.6 Conclusion

In this paper we have discussed precision scaling for thermometry in cold quantum systems. In particular, we have investigated how finite measurement resolution, meaning that states that are close in energy cannot be perfectly distinguished, impacts the precision scaling. We have proposed a finite-resolution criterion characterising such measurements. Based on this, we derived a tightened bound on the scaling of the Fisher information. Furthermore, we showed that this bound is tight as it can be saturated via both an exponential resolution measurement as well as an excitation-preserving, single-qubit measurement on a sample of non-interacting bosons. We validated the approximations involved in demonstrating tightness for the single-qubit measurement by performing a numerical simulation of the sub-Ohmic spin-boson model. Here, we provided an illustration of a Fisher information scaling linearly with temperature. Interestingly, as far as we are aware, this is the best scaling which has been found in any concrete physical model subject to finite-resolution constraints.

6.7 Appendix

6.7.1 Density of states for a bosonic bath

Consider a collection of non-interacting bosonic modes with Hamiltonian $H = \sum_k \omega_k b_k^\dagger b_k$. The partition function of this system takes the form

$$\ln \mathcal{Z}_\beta = - \sum_k \ln [1 - e^{-\beta\omega_k}]. \quad (6.53)$$

In the continuum limit, the sum over modes can be approximated by the integral over a continuous density of modes $g(\omega)$

$$\ln \mathcal{Z}_\beta = - \int_0^\infty d\omega g(\omega) \ln [1 - e^{-\beta\omega}]. \quad (6.54)$$

Expanding the logarithm in powers of $e^{-\beta\omega}$ and re-scaling each term in the resulting series, we can write the above as

$$\ln \mathcal{Z}_\beta = \int_0^\infty d\omega \left[g(\omega) + \frac{g(\omega/2)}{2} + \dots \right] e^{-\beta\omega} \quad (6.55)$$

In the low-temperature limit, this integral is dominated by the low-energy part of the density of modes. If we assume that at low-energies the density of modes takes the form $g(\omega) = \alpha\omega^\gamma$, where α and γ are positive constants, then the integral takes the form

$$\begin{aligned} \ln \mathcal{Z}_\beta &= \alpha \left(\sum_{n=1}^{\infty} \frac{1}{n^{1+\gamma}} \right) \int_0^\infty d\omega \omega^\gamma e^{-\beta\omega} \\ &= \alpha \zeta(\gamma + 1) \Gamma(\gamma + 1) \beta^{-(1+\gamma)}, \end{aligned} \quad (6.56)$$

where ζ denotes the Riemann zeta function and Γ the gamma function. Thus

$$\mathcal{Z}_\beta = \exp \left(\alpha \zeta(\gamma + 1) \Gamma(\gamma + 1) \beta^{-(1+\gamma)} \right). \quad (6.57)$$

This expression has the general form used in the main text if we make the identification $1 + \gamma \rightarrow \gamma$.

6.7.2 Scaling behaviour for the noisy model

In the low-temperature limit, the dominant fine-grained probabilities are those with a vanishing zeroth-order coefficient in the POVM energy-gap expansion, and only coarse-grained probabilities containing contributions from such terms are relevant. For convenience we introduce two sets of fine-grained outcomes: First $\Omega_m = \{\mu \mid \Delta_{m\mu,0} = \Delta_{m\mu,1} = 0\}$, which is the set of fine-grained outcomes giving a non-vanishing contribution to the coarse-grained probability of obtaining outcome m . Second, $\tilde{\Omega}_m = \{\mu \mid \Delta_{m\mu,0} = 0 \text{ and } \Delta_{m\mu,1} \neq 0\}$, which is the set of fine-grained outcomes for which the contribution to the coarse-grained probability for m vanishes sub-exponentially. Lastly, to simplify the later discussion, we denote the specific outcome within $\tilde{\Omega}_m$ which realises the smallest value of the first-order coefficient by $\tilde{\mu}_m$.

We now note that if there exists some coarse-grained outcome m such that Ω_m is empty while $\tilde{\Omega}_m$ is non-empty, then the arguments presented for the noiseless case

also apply to the noisy case, and the same optimal scaling behaviour of the Fisher information can be attained. Thus, in this case, the noise is not detrimental for the scaling. On the other hand, if no such m exists, then we refer to *detrimental noise* (assuming that $\tilde{\Omega}_m$ is non-empty for at least one outcome). For detrimental noise we are then left with outcomes for which Ω_m is non-empty, while $\tilde{\Omega}_m$ may or may not be non-empty. We now show that detrimental noise results in a worse scaling compared to the noise-free scenario. This implies that our finite-resolution bound is also applicable to noisy measurements.

Consider the right-hand side of Eq. (6.38) for the case of detrimental noise. For terms where both $\tilde{\Omega}_m$ and $\tilde{\Omega}_n$ are empty, the scaling behaviour is identical with that of the corresponding noiseless terms (Eq. (6.10)), except that the noiseless coefficients of the POVM energy gap must be replaced by the coarse-grained version

$$\Delta_{m,j}^{(c)} \equiv \sum_{\mu \in \Omega_m} \frac{p_{m\mu;\beta}}{p_{m;\beta}} \Delta_{m\mu,j}. \quad (6.58)$$

If a coarse-grained second-order POVM energy gap exists (that is $\Delta_{m,2}^{(c)} - \Delta_{n,2}^{(c)} \neq 0$ for some m and n), then the same scaling behaviour of the Fisher information as given by Eq. (6.10) is attainable and this scaling is optimal (note that the probabilities considered here tend to nonzero constants at zero temperature). If a second-order gap does not exist, then the optimal scaling is instead provided by terms for which $\tilde{\Omega}_m$ is non-empty for some m . A straightforward calculation shows that the contribution from such terms takes the form

$$\frac{[g_{m\tilde{\mu}_m} \Delta_{m\tilde{\mu}_m,1}]^2}{\sum_{\mu \in \Omega_m} g_{m\mu}} T^{2\Delta_{m\tilde{\mu}_m,1}-2}, \quad (6.59)$$

which should be summed over all outcomes m for which both Ω_m and $\tilde{\Omega}_m$ are non-empty. Assuming that the finite-resolution criterion applies ($\Delta_{m\tilde{\mu}_m,1} \geq 1$), this contribution is at best constant. Hence under the conditions of finite resolution and detrimental noise, a diverging Fisher information is impossible.

As a second example of a noisy measurement we can consider the coarse-graining of a fine-grained measurement of the form

$$\begin{aligned} \Pi_{00} &= \frac{1}{2} e^{-\kappa H}, & \Pi_{01} &= \frac{1-\eta}{2} e^{-\kappa H}, \\ \Pi_{10} &= \frac{1}{2} (\mathbb{1} - e^{-\kappa H}), & \Pi_{11} &= \frac{1}{2} \mathbb{1} - \frac{1-\eta}{2} e^{-\kappa H}. \end{aligned} \quad (6.60)$$

This fine-grained model is illustrated in Fig. 6.2b. For this measurement we find $\Delta_{00;\beta} = \Delta_{01;\beta} = 0$ and

$$\begin{aligned} \Delta_{10;\beta} &\approx (1+\gamma)T + (1+\gamma)\alpha\kappa T^{2+\gamma} \\ \Delta_{11;\beta} &\approx (1+\gamma)\frac{\alpha\kappa}{\eta} T^{2+\gamma}. \end{aligned} \quad (6.61)$$

Hence, as in the previous example, the fine-grained measurement gives a Fisher information scaling as $T^{\gamma-1}$ to leading order, and the coarse-grained measurement gives a $T^{2\gamma}$ scaling,

$$\mathcal{F}_T = \frac{(2-\eta)(\alpha\kappa)^2}{\eta} (1+\gamma)^2 T^{2\gamma} + \mathcal{O}(T^{1+3\gamma}). \quad (6.62)$$

Thus the same scaling behaviour of the Fisher information is observed for this alternative example of a noisy model. Note that both models exhibit detrimental noise which results in the different scalings for the fine- and coarse-grained Fisher information.

6.7.3 The non-excitation-preserving interaction as a noisy POVM

From Eq. (6.47), we find that the POVM elements can be written as

$$\Pi_1 = t^2 \langle 0 | H_{\text{int}} | 1 \rangle \langle 1 | H_{\text{int}} | 0 \rangle, \quad (6.63)$$

and $\Pi_0 = 1 - \Pi_1$. In the thermal state under consideration, there are no coherences between different bosonic modes and there is no squeezing. Therefore, many terms in Eq. (6.63) do not contribute to the probabilities. Dropping these terms, we can write a slightly simpler POVM that results in the exact same probabilities, capturing the full effect of the measurement

$$\tilde{\Pi}_1 = t^2 \sum_k (g_k^2 + \lambda_k^2) a_k^\dagger a_k + t^2 \sum_k \lambda_k^2, \quad (6.64)$$

and $\tilde{\Pi}_0 = 1 - \tilde{\Pi}_1$. This POVM has an energy gap that has no Taylor expansion, scaling as T^{2+s} in the low temperature limit for $g_k = \lambda_k$ and the spectral density given in Eq. (6.50). We can however write the POVM in Eq. (6.64) as a coarse graining over the fine-grained POVM (note the similarity to Eq. (6.39))

$$\begin{aligned} \tilde{\Pi}_{11} &= \frac{1+\eta}{2} X, & \tilde{\Pi}_{10} &= \frac{1-\eta}{2} (1-X), \\ \tilde{\Pi}_{00} &= \frac{1+\eta}{2} (1-X), & \tilde{\Pi}_{01} &= \frac{1-\eta}{2} X, \end{aligned} \quad (6.65)$$

such that $\tilde{\Pi}_1 = \tilde{\Pi}_{11} + \tilde{\Pi}_{10}$ and $\tilde{\Pi}_0 = \tilde{\Pi}_{00} + \tilde{\Pi}_{01}$. Here we introduced

$$\eta = 1 - 2t^2 \sum_k \lambda_k^2, \quad (6.66)$$

and

$$X = \frac{t^2 \sum_k (g_k^2 + \lambda_k^2) a_k^\dagger a_k}{1 - 2t^2 \sum_k \lambda_k^2}. \quad (6.67)$$

The fine-grained POVM elements are of the same form as the POVM elements for the excitation-preserving case. Indeed, setting $\lambda_k = 0$, only $\tilde{\Pi}_{00}$ and $\tilde{\Pi}_{11}$ remain finite but do not change their form. We therefore find the same POVM gaps as for the excitation-preserving case

$$\Delta_{00} = \Delta_{10} = 0, \quad \Delta_{11} = \Delta_{01} = (1+s)T. \quad (6.68)$$

The Fisher information for the fine-grained POVM thus scales as T^{s-1} . The coarse grained POVM gap is determined by Eq. (6.37) and reads

$$\Delta_1 = \frac{p_{11}}{p_{11} + p_{10}} (1+s)T, \quad (6.69)$$

which scales as T^{s+2} for the scenario considered in the main text.

6.7.4 Tensor network simulation

The numerical simulations performed for the spin-boson model are in essence an extension of work initiated during my MSc studies and completed during the first year of my PhD studies. The work was published in Ref. [127] ©2019 American Physical society. This paper is reproduced in appendix B.

Here we provide details of the numerical methods behind the result shown in Fig. 6.3. We consider the ground state probability

$$p_{0;\beta}^{(k)} = \text{Tr} \left\{ \hat{P}_0 \mathcal{U}_{\delta t}^k \left[\hat{P}_0 \otimes \rho_\beta \right] \right\}, \quad (6.70)$$

where \hat{P}_0 is a projection operator onto the qubit ground state $|0\rangle$, and we have decomposed the unitary evolution into k -steps of duration δt . Furthermore we consider the spin-boson model

$$\hat{H} = \sum_k \omega_k \hat{a}_k^\dagger \hat{a}_k + \frac{1}{2} \Omega \hat{\sigma}_z + \frac{1}{2} \hat{\sigma}_x \sum_k g_k (a_k + a_k^\dagger). \quad (6.71)$$

The spin-boson model can be numerically simulated using recently developed tensor network methods [154, 127]. Taking each unitary step to be of a short duration we can make the approximation (Trotter-Suzuki decomposition)

$$U_{\delta t} = W_{\delta t/2} V_{\delta t} W_{\delta t/2} + \mathcal{O}(\delta t^3), \quad (6.72)$$

where $W_{\delta t} = \exp(-i\delta t(\hat{H} - \Omega\hat{\sigma}_z/2))$ describes the influence of the sample on the probe qubit, and $V_{\delta t} = \exp(-i\delta t\Omega\hat{\sigma}_z/2)$ describes the free evolution of the probe qubit. As the interaction term is diagonal in the eigenstates of the operator $\hat{\sigma}_x$, we can expand the ground state probability in terms of these eigenstates. This gives rise to a discrete Feynman-Vernon Influence functional, which can be summed analytically. The ground state probability then takes the form

$$p_{0;\beta}^{(k)} = \sum_{\{\alpha\}} \hat{P}_0^{\alpha_{2k+1}} \mathcal{V}_{\delta t}^{\alpha_{2k}\alpha_{2k+1}} \dots \mathcal{V}_{\delta t}^{\alpha_2\alpha_1} \hat{P}_0^{\alpha_0} \times \left[\prod_{i=1}^{2k} \prod_{j=1}^i \mathcal{A}_\beta^{\alpha_i\alpha_j} \right] \left[\prod_{l=0}^k \delta_{\alpha_{2l+1}, \alpha_{2l}} \right]. \quad (6.73)$$

where we have introduced a compound index $\alpha = (s, r)$ of spin-x eigenvalues, $\delta_{\alpha_i, \alpha_j}$ denotes the Kronecker delta function, $\hat{P}_0^\alpha = \langle s | \hat{P}_0 | r \rangle$, and \mathcal{V} are the Liouville operators representing the free dynamics of the ancilla qubit

$$\mathcal{V}_{\delta t}^{\alpha\alpha'} = \langle s | V_{\delta t} | s' \rangle \langle r' | V_{\delta t}^\dagger | r \rangle. \quad (6.74)$$

The influence tensors, $\mathcal{A}_\beta^{\alpha_i\alpha_j}$, describe the influence of the sample on the state of the qubit and contain all the temperature dependence of the probability. For linearly coupled models, the individual tensors depend only on the time separation $(i-j)\delta t/2$. The influence tensors are given by

$$\mathcal{A}_\beta^{\alpha_i\alpha_j} = e^{-(s_i - r_i)(\eta_{i-j} s_j - \eta_{i-j}^* r_j)}, \quad (6.75)$$

expressed in terms of the memory kernel elements

$$\eta_{i-j} = \begin{cases} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} dt' dt'' C(t' - t'') , & i \neq j \\ \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} dt' dt'' C(t' - t'') , & i = j \end{cases}, \quad (6.76)$$

which are themselves defined in terms of the bath auto-correlation function

$$C(t) = \frac{1}{\pi} \int_0^\infty d\omega \rho(\omega) \frac{\cosh[\omega(\beta - it)]}{\sinh(\beta\omega/2)}. \quad (6.77)$$

The bath auto-correlation function is given in terms of the spectral density $\rho(\omega)$ introduced in the main text.

The attainable temperature estimation precision depends not only on the ground state probability, but also on the derivative of this probability. Computing the derivative of the distribution with respect to the inverse temperature gives

$$\begin{aligned} \partial_\beta p_{0\beta}^{(k)} &= \sum_{i=1}^{2k} \sum_{j=1}^i \mu_{ij} \sum_{\{\alpha\}} [\Pi_{l=0}^k \delta_{\alpha_{2l+1}, \alpha_{2l}}] \\ &\times \hat{P}_0^{\alpha_{2k+1}} \mathcal{V}_{\delta t}^{\alpha_{2k} \alpha_{2k+1}} \dots \mathcal{V}_{\delta t}^{\alpha_2 \alpha_1} \hat{P}_0^{\alpha_0} \\ &\times [\Pi_{i=1}^{2k} \Pi_{j=1}^i \mathcal{A}_\beta^{\alpha_i \alpha_j}] \alpha_i^- \alpha_j^- \end{aligned} \quad (6.78)$$

where we have defined $\alpha^- = s - r$. It turns out that the same tensor network methods used to compute the probability can be used to compute the derivative of the probability. Furthermore we have defined $\mu_{ij} = -\partial_\beta \eta_{i-j}$, the square of which gives the Fisher information scaling at low-temperatures. At low temperatures, all the temperature dependence of the ground-state probability comes from these coefficients. We can approximate them by the series

$$\begin{aligned} \mu_{ij} &= \frac{\alpha \delta t^2}{4\beta^{\gamma+2}} \times \left[\Gamma(\gamma + 2) - \frac{\delta t^2}{8\beta^2} (i-j)^2 \Gamma(\gamma + 4) \right. \\ &\quad \left. + \frac{\delta t^4}{376\beta^4} (i-j)^4 \Gamma(\gamma + 6) - \dots \right] \end{aligned} \quad (6.79)$$

This shows that, to leading order, the exact expressions reproduce the low-temperature Fisher-information scaling obtained within the short-time approximation.

Chapter 7

Optimal Quantum Thermometry with Coarse-grained Measurements

This chapter is composed entirely of text and figures from Ref. [124] © 2021 American Physical Society. The appendix of the published paper [124] contains technical results and has not been included.

The chapter is self-contained. In particular the notation does not conform to previous chapters. This chapter takes a conventional approach to thermometry, i.e. one based on the absolute error and the associated noise-to-signal ratio. The technical results of the chapter are valid. However, the physical implications should be judged in light of the results of chapter 4.

7.1 Introduction

Thermometry is a basic metrological task, vital throughout science and technology. Estimating temperature is important on all scales, ranging from astronomical bodies with temperatures in the millions of Kelvin to atomic systems near absolute zero. In particular, applications of thermometry in nano- or micro-scaled devices are becoming increasingly relevant as technology advances [26, 156, 45, 157]. Examples include, for instance, accurate temperature estimation of ultra-cold gases [158, 159, 160, 19, 20], in electronic systems [25, 22, 161], or the use of atomic-size devices, such as colour centers in diamond or quantum dots, as probes to be employed in a variety of systems [162, 8, 130, 163]. At these scales, quantum effects have significant influence on the achievable precision. It is therefore important to understand what the fundamental limits for temperature estimation in quantum systems are.

Quantum features offer both advantages and challenges to thermometry [46, 45]. Advantages range from measurement enhancements due to strong coupling [106, 91, 19], correlated probes [110, 114], or nonequilibrium probes [164, 165, 103, 166, 123, 167, 168, 169]. The challenges are related to the inherent difficulty of accessing information in quantum systems, due, for instance, to measurement backaction or natural limitations in performing high-resolution measurements [51, 88].

When the measurement resolution is unlimited, the ultimate precision of temperature estimation allowed by quantum mechanics is obtained by performing projective measurements of energy [50, 83, 170]. However, for large (or even moderately-sized)

many-body systems, one seldom has access to measurements which distinguish individual energy levels. Instead, one usually measures only a local subsystem of the sample [104, 133, 91] or performs a global measurement with only a finite resolution [51, 88] [see Fig. 7.1(a)]; alternatively, one addresses the sample indirectly, by measuring a probe that has interacted with it [106, 91, 171, 167, 169] [see Fig. 7.1(b)]. All of these cases are examples of a coarse-grained measurement, which from an abstract point of view can be described by a d -outcome generalized quantum measurement of a D -dimensional system, with $d < D$. The fact that $d \ll D$ in most physically relevant cases may reduce the precision significantly. It is hence natural to ask what the optimal measurement strategy and the associated precision of temperature estimation are under such limitations. In this paper, we put forth a framework for addressing this question in detail. The framework is based on ideas from signal processing and parameter-estimation theory, and provides a simple, easy-to-use toolbox for studying coarse-grained thermometry of both few- and many-body systems. We illustrate the framework by applying it to paradigmatic many-body models of spin lattices, both close to and far from criticality.

In a second part of the article, these abstract ideas are applied to probe-based temperature measurements. Here, the temperature of a sample is estimated by letting it interact with a probe (and possibly some auxiliary system), and then measuring the probe, as illustrated in Fig. 7.1(b). These type of measurements are particularly appealing since they provide a natural way to overcome one of the main challenges in thermometry: the design of noninvasive measurements, , e.g., for ultra-cold atomic gases [17, 172, 173, 174, 175, 20]. In such probe-based measurements, a natural strategy is to let the probe reach thermal equilibrium with the sample [50]. Yet, it has been shown that the precision can be considerably enhanced by nonequilibrium strategies, where the probe either interacts with the sample for a finite time [164, 165, 103, 176, 166, 123, 177, 167, 168, 20], couples strongly to the sample [106, 91], or uses an ancillary catalytic system [169]. Here, we use our framework to obtain a fundamental bound on any such nonequilibrium strategy. We map the problem of probe-based thermometry to that of coarse-grained thermometry, and determine the maximal sensitivity that can be obtained with a probe of dimension d . We construct a specific fine-tuned sample-probe interaction that always saturates the bound, and notably show that it can also be reached in relevant experimental situations. In particular, when the sample and the probe are described by a harmonic oscillator and a qubit, respectively, the optimal nonequilibrium estimation scheme in the low-temperature regime can be obtained via the Jaynes-Cummings Hamiltonian. This is of direct relevance to temperature measurements of Bose-Einstein condensates [178] or micromechanical resonators [164, 165] via qubit probes.

7.2 Framework

To be specific, we consider a system S living in a D -dimensional Hilbert space, described by a Hamiltonian H , and initially in a canonical thermal state

$$\tau = \frac{1}{Z} e^{-\beta H}, \quad (7.1)$$

where $\beta = 1/T$ is the inverse temperature (we adopt units such that $k_B = 1$) and $Z = \text{Tr} e^{-\beta H}$ is the partition function. This is a family of states parameterized by

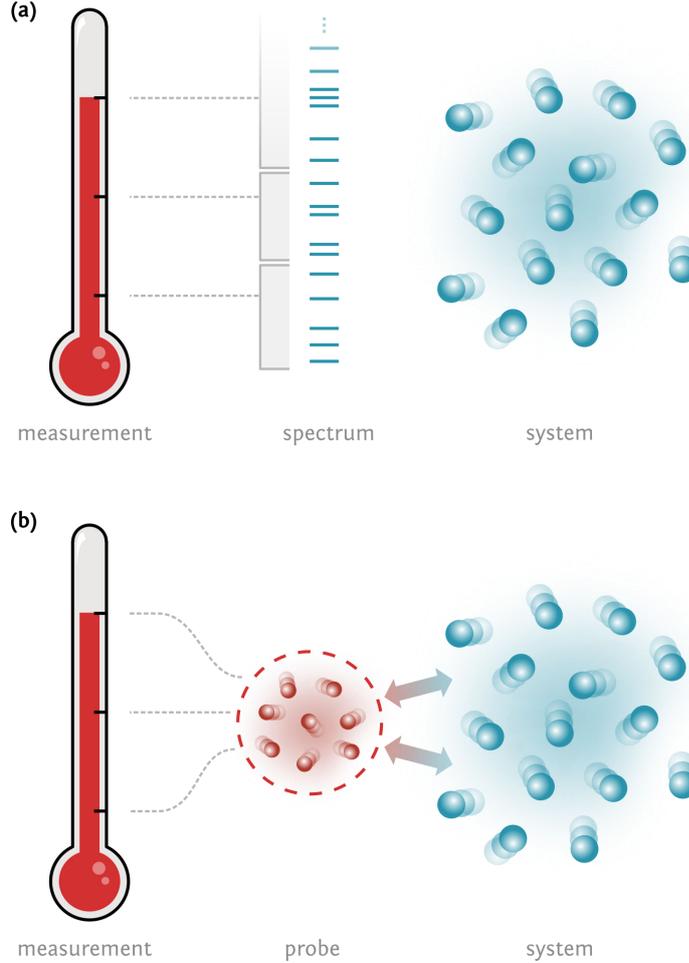


Figure 7.1: *This figure and caption is taken from Ref. [124] ©2021 American Physical Society.* **(a)** Thermometry with coarse-grained energy measurements. The measurement can be understood as resulting from post-processing of a fine-grained, projective energy measurement. Energies are grouped into bins and a single outcome is assigned to each bin. **(b)** Thermometry with probe-based measurements. A probe interacts unitarily with a target system. A measurement is then performed on the probe alone and used to infer the temperature of the target.

the temperature, and the smallest variance in estimating this parameter, based on any measurement, is hence lower bounded by the quantum Cramér-Rao bound [48] (see also [179, 180])

$$\Delta T^2 \geq \frac{1}{n\mathcal{F}(\tau)}, \quad (7.2)$$

where n is the number of repetitions of the measurement and \mathcal{F} is the quantum Fisher information with respect to T , which we refer to as the ‘thermal Fisher information’. It is given by

$$\mathcal{F}(\tau) = \beta^4 \delta(H) = \beta^2 C, \quad (7.3)$$

with $\delta(H) := \langle H^2 \rangle_\tau - \langle H \rangle_\tau^2$, where the angle brackets denote averaging over the quantum state: $\langle \hat{O} \rangle_\tau := \text{Tr}(\hat{O}\tau)$, and C is the heat capacity of the system ($C := d\langle H \rangle/dT$).

The optimal measurement, attaining the thermal Fisher information, is a projective measurement onto the eigenbasis of H , i.e., a projective measurement of the system energy. In this optimal scenario, the more the energy fluctuates, the more precise the measurement can be in principle. The optimal measurement saturates the inequality (7.2) for arbitrary n when the temperature estimator is unbiased. When an unbiased estimator is not available, a large class of generic biased estimators will still asymptotically saturate the inequality (7.2) in the limit of many repetitions ($n \rightarrow \infty$) [74]. Moreover, in the specific case of temperature estimation in many-body systems consisting of $N \gg 1$ particles, one can go even further and explicitly construct an estimator that, despite being biased for any finite N , will saturate the Cramér-Rao bound (7.2) as $N \rightarrow \infty$, even for $n = 1$ [84].

However, when fine-grained measurements of energy are not available, while remaining valid, the bound (7.2) will in general become too loose. We formalise the problem as follows. Suppose the resolution of the experiment is limited to $d < D$ measurement outcomes. *What is then the maximal precision for estimating the temperature of S and which is the optimal d -outcome measurement achieving it?* Below, we provide a systematic way to identify the optimal measurement.

Moreover, in Sec. 7.5, we prove that the highest Fisher information achievable by a $d < D$ dimensional probe undergoing an arbitrary interaction with the sample is equal to the optimal d -outcome coarse-grained Fisher information in the above sense. We thus provide a fundamental benchmark for any conceivable protocol of probe-based thermometry. In particular, this includes any standard thermometric technique in current experimental setups.

7.3 Optimal coarse graining

We consider coarse-grained thermometry on a D -dimensional system. We take coarse graining to mean that only generalised measurements, i.e., positive operator-valued measures (POVM) with at most $d < D$ outcomes are available. We then construct a framework for identifying the optimal POVM for thermometry in two steps.

First, we show that the optimal POVM is a projection onto energy subspaces of the system. This means that we can split the system spectrum into d subsets and study measurements which project onto the corresponding eigensubspaces.

Second, we show that the optimal choice of subsets consists of consecutive “bins,” i.e., the sets are not interspersed. We then lay down a method for constructing the optimal choice of bins, for any given system spectrum. This is done by casting the problem in the language of an analogous signal-processing problem, known as Lloyd-Max quantization [181].

7.3.1 Optimal POVM

We take a d -outcome POVM $\mathcal{M} = \{M_\alpha\}_{\alpha=1}^d$ and a system in the thermal state τ of Eq. (7.1). Each outcome α then occurs with probability

$$p_\alpha = \text{Tr}(M_\alpha \tau). \quad (7.4)$$

This distribution contains information about the temperature T , as quantified by the Fisher information [182]

$$\mathcal{C} = \sum_{\alpha=1}^d \frac{1}{p_{\alpha}} \left(\frac{\partial p_{\alpha}}{\partial T} \right)^2, \quad (7.5)$$

which, for a thermal state, becomes [51]

$$\mathcal{C} = \beta^4 \sum_{\alpha} \frac{1}{p_{\alpha}} \left[\text{Tr} (\tau M_{\alpha} (H - \langle H \rangle_{\tau})) \right]^2. \quad (7.6)$$

Note that, after coarse graining, p_{α} is no longer in the so-called exponential family with respect to T ¹, which means that no temperature estimator can satisfy the Cramér-Rao bound for any finite number of repetitions n [74]. However, the Fisher information (\mathcal{C}) is still a key precision quantifier as, for all unbiased and certain generic biased estimators, such as the maximum likelihood estimator, the Cramér-Rao bound is saturated asymptotically in the $n \rightarrow \infty$ limit. Moreover, in the same limit, the Fisher information retains its key role even in the Bayesian estimation approach².

We thus set as our goal to determine the optimal d -outcome POVM which maximises the Fisher information \mathcal{C} . We first observe that the POVM elements M_{α} can be taken to be diagonal in the energy eigenbasis. Because τ is diagonal in this basis, only diagonal elements of M_{α} contribute to the probability p_{α} in Eq. (7.4). Dropping all off-diagonal elements from each M_{α} results in a valid POVM since the operators remain positive and still sum to identity. Hence, for every POVM, there exists a diagonal POVM which achieves the same p_{α} and thus the same Fisher information. It is therefore sufficient to consider diagonal POVMs when looking for an optimal POVM maximising \mathcal{C} .

Next, let us note that \mathcal{C} is convex with respect to the POVM. That is, denoting by $\mathcal{C}_{\mathcal{M}}$ the Fisher information corresponding to a particular POVM and considering two POVMs \mathcal{M} and \mathcal{N} and a mixing parameter $0 \leq \lambda \leq 1$, we have

$$\mathcal{C}_{\lambda\mathcal{M} + \bar{\lambda}\mathcal{N}} \leq \lambda\mathcal{C}_{\mathcal{M}} + \bar{\lambda}\mathcal{C}_{\mathcal{N}}, \quad (7.7)$$

where $\bar{\lambda} = 1 - \lambda$. This can be seen by rewriting Eq. (7.6) as

$$\mathcal{C} = \beta^4 \sum_{\alpha} \frac{1}{p_{\alpha}} (W_{\alpha})^2, \quad (7.8)$$

where $W_{\alpha} = \text{Tr} (\tau M_{\alpha} [H - \langle H \rangle_{\tau}])$. Both p_{α} and W_{α} are linear in the POVM operators. Hence, when mixing POVMs, $W_{\alpha}^{(\lambda\mathcal{M} + \bar{\lambda}\mathcal{N})} = \lambda W_{\alpha}^{(\mathcal{M})} + \bar{\lambda} W_{\alpha}^{(\mathcal{N})}$ and $p_{\alpha}^{(\lambda\mathcal{M} + \bar{\lambda}\mathcal{N})} = \lambda p_{\alpha}^{(\mathcal{M})} + \bar{\lambda} p_{\alpha}^{(\mathcal{N})}$. Equation (7.7) is then an immediate consequence of the fact that $\frac{1}{p} W^2$ is a jointly convex function of p and W (see, e.g., Ref. [184]).

Finally, we show that the optimal POVM is necessarily a collection of non-overlapping projectors onto eigensubspaces of H . Indeed, take a POVM \mathcal{M} such that, for some

¹In that p_{α} cannot be written as $\exp[\mathfrak{R}(\alpha)f(T) + \mathfrak{L}(T) + \mathfrak{M}(\alpha)]$, where f , \mathfrak{R} , \mathfrak{L} , and \mathfrak{M} are some functions.

²Indeed, it follows from the van Trees inequality [183] that, for *any* estimator, in the $n \rightarrow \infty$ limit, $\Delta T \gtrsim 1/[n \int dT p(T) \mathcal{C}(T)]$, where $p(T)$ is the prior distribution of T .

eigenstate $|k\rangle$ of H , there are at least two POVM elements for which $\langle k| M_\alpha |k\rangle > 0$. Define

$$\varsigma_\alpha = \langle k| M_\alpha |k\rangle. \quad (7.9)$$

Now, construct d new POVMs $\mathcal{N}^{(\gamma)}$, with elements

$$N_\alpha^{(\gamma)} = M_\alpha + (\delta_{\alpha\gamma} - \varsigma_\alpha) |k\rangle \langle k|. \quad (7.10)$$

Each $\mathcal{N}^{(\gamma)}$ has the property that only $N_\gamma^{(\gamma)}$ has a nonzero k 'th diagonal, namely $\langle k| N_\gamma^{(\gamma)} |k\rangle = 1$ while $\langle k| N_{\alpha \neq \gamma}^{(\gamma)} |k\rangle = 0$. Furthermore, we note that

$$\sum_\gamma \varsigma_\gamma \mathcal{N}^{(\gamma)} = \mathcal{M}. \quad (7.11)$$

Since the M_α form a POVM, we have $\varsigma_\alpha \geq 0$ and $\sum_\alpha \varsigma_\alpha = 1$. By convexity (7.7) of the Fisher information,

$$\mathcal{C}_\mathcal{M} \leq \sum_\gamma \varsigma_\gamma \mathcal{C}_{\mathcal{N}^{(\gamma)}}, \quad (7.12)$$

and it follows that there must be at least one γ for which $\mathcal{C}_{\mathcal{N}^{(\gamma)}}(T) \geq \mathcal{C}_\mathcal{M}(T)$. This means that the optimal POVM will have to be one that consists of non-overlapping projectors on eigensubspaces of H . Given a (possibly degenerate) Hamiltonian $H = \sum_i E_i |i\rangle \langle i|$ the optimal POVM will thus be of the form

$$\Pi_\alpha = \sum_{E_i \in I_\alpha} |i\rangle \langle i|, \quad (7.13)$$

where the I_α define a partition of the set of all eigenenergies into non-overlapping subsets (“bins”). To summarize, optimal, coarse-grained thermometry can always be achieved by considering projective measurements onto non-overlapping eigenenergy subspaces.

7.3.2 Optimal binning

We now construct a method for determining the optimal eigenenergy subsets, defining the optimal POVM. For convenience, we choose the basis of H such that

$$E_1 \leq E_2 \leq \dots \leq E_D, \quad (7.14)$$

and write the probability of finding the system in bin I_α as [cf. Eq. (7.4)]

$$p_\alpha = \text{Tr}(\Pi_\alpha \tau) = \sum_{E_i \in I_\alpha} q_i, \quad (7.15)$$

where $q_i = \exp(-\beta E_i)/Z$. Next, we introduce the “bin energies” (normalized average energy within each bin),

$$\epsilon_\alpha = \frac{1}{p_\alpha} \sum_{E_i \in I_\alpha} q_i E_i, \quad (7.16)$$

and, with these definitions, reexpress the Fisher information in Eq. (7.6) for the corresponding measurement as

$$\mathcal{C} = \beta^4 \sum_{\alpha=1}^d p_{\alpha} (\epsilon_{\alpha} - \langle H \rangle)^2. \quad (7.17)$$

We shall henceforth refer to this as the *coarse-grained* Fisher information. Compared with Eq. (7.3), the expression for \mathcal{C} describes how each of the energies ϵ_{α} fluctuates away from the average.

As a first step towards finding the optimal sets I_{α} , in the appendix of Ref. [124], applying a result from Ref. [185], we prove that the best choice of I_{α} is given by a binning into consecutive intervals: $I_{\alpha} = \{E_{i_{\alpha-1}}, E_{i_{\alpha-1}+1}, \dots, E_{i_{\alpha-1}+|I_{\alpha}|-1}\}$, where $1 \leq i_{\alpha-1} \leq D$ and the iteration rule is $i_{\alpha} = i_{\alpha-1} + |I_{\alpha}|$. Introducing the ‘‘boundaries’’ $b_{\alpha} := E_{i_{\alpha}}$, we can conveniently write $I_{\alpha} = [b_{\alpha-1}, b_{\alpha})$, with the proviso that $b_0 = E_1$ and $b_d = E_{D+1} = E_D$ (the extra level E_{D+1} will not enter into any physical quantity and is introduced just so that E_D enters I_d despite $[b_{d-1}, b_d)$ having an open end). Note that, for discrete spectra, the boundaries b_{α} need not be exactly at energy eigenvalues. Positioning a boundary anywhere between neighbouring eigenvalues results in the same POVM in Eq. (7.13). The remaining task is then to find the optimal intervals $I_{\alpha} = [b_{\alpha-1}, b_{\alpha})$ which maximize \mathcal{C} . This will give the best strategy for temperature estimation using a d -outcome POVM.

Before carrying out this optimization, it is useful to recast the problem in terms of the density of states (DOS)

$$\Omega(E) = \sum_i \delta(E - E_i), \quad (7.18)$$

where δ denotes the Dirac’s delta function (note that this definition does not assume a continuous spectrum). Expectation values of any function $g(H)$ of the Hamiltonian may then be written as

$$\langle g(H) \rangle = \sum_i g(E_i) q_i = \int dE g(E) \Omega(E) \frac{e^{-\beta E}}{Z}. \quad (7.19)$$

This means that we can define the distribution of energy as

$$q(E) = \Omega(E) \frac{e^{-\beta E}}{Z}, \quad (7.20)$$

so that expectation values can be computed simply in terms of integrals over $q(E)$. This way, the probabilities p_{α} in Eq. (7.15) and the bin energies in Eq. (7.16) can be conveniently written as

$$p_{\alpha} = \int_{b_{\alpha-1}}^{b_{\alpha}} dE q(E) \quad \text{and} \quad \epsilon_{\alpha} = \frac{1}{p_{\alpha}} \int_{b_{\alpha-1}}^{b_{\alpha}} dE q(E) E. \quad (7.21)$$

These quantities are therefore all expressed explicitly as functions of the boundaries b_{α} .

The advantage of introducing the DOS is twofold. First, it allows for a unified treatment of Hamiltonians with discrete and quasi-continuous spectra (as one would expect in quantum many-body systems). Second, it allows us to frame the problem in

the language of signal processing. A common task in signal processing is to transmit a continuous function $q(E)$ through a channel. Often, in order to do so, the function must be discretized into a finite set of bins $I_\alpha = [b_{\alpha-1}, b_\alpha)$. The question is then which choice of bins leads to the optimal transmission. This problem is solved by using the scheme known as Lloyd-Max quantization (see Ref. [181], Chapter 3). If one uses the mean squared variations of energy as a figure of merit, one sees that the maximization of the Fisher information [see Eq. (7.17)] becomes entirely analogous to this signal processing problem.

To proceed, we introduce the quantity

$$\mathcal{D} = \beta^4 \sum_{\alpha=1}^d \int_{b_{\alpha-1}}^{b_\alpha} dE q(E) (E - \epsilon_\alpha)^2. \quad (7.22)$$

It can be directly verified that the thermal Fisher information given by Eq. (7.3) can be decomposed as

$$\mathcal{F} = \mathcal{C} + \mathcal{D}, \quad (7.23)$$

which means that the task of maximizing \mathcal{C} , for a fixed \mathcal{F} , is tantamount to that of minimizing \mathcal{D} .

The minimization can be carried out in the usual way, by equating to zero the derivatives of \mathcal{D} with respect to b_α . A straightforward calculation shows that the minima of \mathcal{D} occur when the intervals b_α satisfy the implicit (and generally nonlinear) equations,³

$$b_\alpha = \frac{\epsilon_{\alpha+1} + \epsilon_\alpha}{2}, \quad \alpha = 1, \dots, d-1. \quad (7.24)$$

These equations are implicit because ϵ_α itself is a function of the $\{b_\alpha\}$ [Eq. (7.16)]. This summarises the core of our framework. Equation (7.24) provides a recipe for how to optimize the energy bins in a d -outcome POVM in order to maximize the thermometric precision.

7.3.3 Illustrative examples

Let us now present two examples using our framework for optimal coarse-grained thermometry.

Noninteracting qubits

A simple, but illustrative example, is a system of N identical, noninteracting qubits. The system is in a thermal state, and take the ground- and excited-state energies to be 0 and 1, respectively. The energy levels of the system will thus range from 0 to N in integer steps. The probability to find the system in a state with energy j is then

$$q_j = \binom{N}{j} s^j r^{N-j}, \quad (7.25)$$

³The solutions to this equation are also guaranteed to be an actual minimum of \mathcal{D} , never a maximum. First, $\partial^2 \mathcal{D} / \partial b_\alpha \partial b_{\alpha'} = 0$ for $\alpha' \neq \alpha$. Second, at $b_\alpha = (\epsilon_{\alpha+1} + \epsilon_\alpha) / 2$, we have $\partial^2 \mathcal{D} / \partial b_\alpha^2 = 2\beta^4 (\epsilon_{\alpha+1} - \epsilon_\alpha) q(b_\alpha) \geq 0$, since $q(b_\alpha) \geq 0$ by construction and $\epsilon_{\alpha+1} \geq \epsilon_\alpha$ by hypothesis.

where we have defined the excited-state population $s = 1/(e^\beta + 1)$ and $r = 1 - s$. For a d -outcome measurement, the probabilities and bin energies take the form

$$\begin{aligned} p_\alpha &= \sum_{j=b_{\alpha-1}}^{b_\alpha-1} \binom{N}{j} s^j r^{N-j}, \\ \epsilon_\alpha &= \frac{1}{p_\alpha} \sum_{j=b_{\alpha-1}}^{b_\alpha-1} \binom{N}{j} j s^j r^{N-j}, \end{aligned} \quad (7.26)$$

where the bin positions b_α , which are integers in this case, are determined from Eq. (7.24) (with $b_0 = 0$). Note that, for this system, the average energy is simply $\langle H \rangle = Ns$, while the thermal Fisher information (7.3) is $\mathcal{F} = \beta^4 Nrs$. In Fig. 7.2, we show the ratio between \mathcal{C}/\mathcal{F} as a function of the bin positions b_α , for the cases $d = 2, 3$. The bins b_α which maximize \mathcal{C}/\mathcal{F} are precisely the solutions of Eq. (7.24).

According to the De Moivre–Laplace theorem [59], in the $N \gg 1$ limit, the distribution q_j becomes Gaussian. Using this, we show in the appendix of Ref. [124] that, for binary measurements ($d = 2$) and $N \gg 1$, optimal binning strategy leads to

$$\mathcal{C} = \frac{2}{\pi} \mathcal{F}, \quad (7.27)$$

and is achieved when the boundary is inserted at $b = \langle H \rangle = Ns$. This result is noteworthy, it shows that, irrespective of the number of qubits N , it is always possible to construct a dichotomic measurement strategy which captures $(2/\pi) \approx 0.63$ of the full thermal Fisher information.

Linear density of states

To contrast with the N qubits case, we now consider an example of a system with a continuous spectrum. Namely, we assume the system has a linear density of states: $\Omega(E) \propto E$. Such a DOS is met, for instance, in a noninteracting, ultra-relativistic gas in two dimensions. If the system is in a thermal state [Eq. (7.1)], the average energy is simply $\langle H \rangle = 2/\beta$ and the variance is $\delta(H) = 2/\beta^2$. Thus, the thermal Fisher information given in Eq. (7.3) is $\mathcal{F} = 2\beta^2$.

We first consider the case of binary measurements, $d = 2$, defined by a single boundary b . The probabilities p_1 and p_2 [Eq. (7.21)] are then given by

$$p_1 = 1 - p_2, \quad p_2 = \int_b^\infty dE q(E) = (1 + \beta b)e^{-\beta b},$$

and the corresponding bin energies become

$$\epsilon_1 = \frac{2}{\beta} + \frac{\beta b^2}{1 + \beta b - e^{\beta b}}, \quad \epsilon_2 = \frac{2}{\beta} + \frac{\beta b^2}{1 + \beta b}.$$

Thus, the Fisher information for the measurement, Eq. (7.17), is

$$\mathcal{C} = \frac{\beta^6 b^4}{(1 + \beta b)(e^{\beta b} - 1 - \beta b)}. \quad (7.28)$$

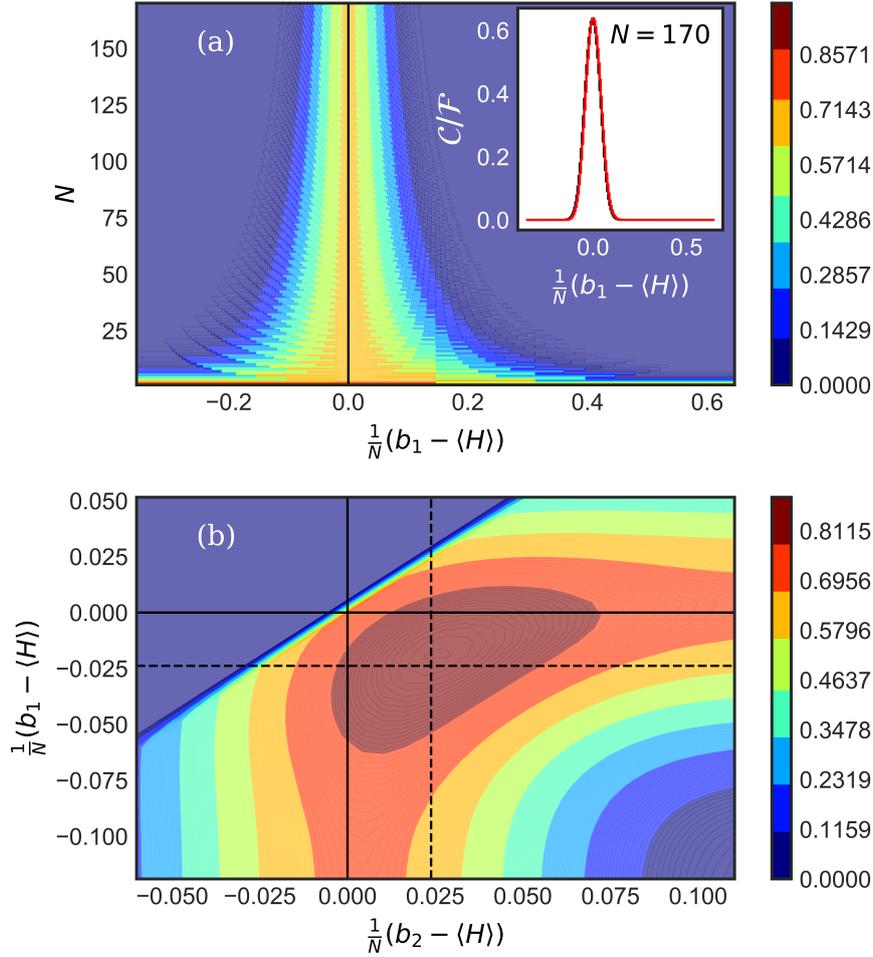


Figure 7.2: *This figure and caption is taken from Ref. [124] ©2021 American Physical Society.* Ratio between the coarse-grained and thermal Fisher informations \mathcal{C}/\mathcal{F} at $\beta = 0.6$. (a) The case of $d = 2$, plotted as a function of the partition $(b_1 - \langle H \rangle)/N$ and the number of qubits N . The ratio reaches a maximum value of approximately 0.64 in the large N limit. (Inset) Cross section at $N = 170$. (b) The case of $d = 3$ for $N = 170$, plotted as a function of the lower partition $(b_1 - \langle H \rangle)/N$ and the upper partition $(b_2 - \langle H \rangle)/N$. The ratio has a maximum value of approximately 0.82. The dashed lines gives the optimal partitions as predicted by a symmetric partition around the mean energy.

To find the optimal partition, we solve Eq. (7.24) for b , i.e., $b = (\epsilon_1 + \epsilon_2)/2$. This is a nonlinear equation which can be solved numerically. A plot of \mathcal{C}/\mathcal{F} is shown in the inset of Fig. 7.3(a). It attains a maximum at $\beta b \approx 2.589$. At this point $\mathcal{C} \approx 0.643\mathcal{F}$, i.e., the binary measurement reaches approximately 64% of maximal Fisher information for any possible measurement (this is slightly higher than in Eq. (7.27)).

The dependence of \mathcal{C} on the number of outcomes d , for optimal binnings, is shown in Fig. 7.3(a). Quite remarkably, even with as little as 8 bins, one can already reach a precision of $\approx 97\%$ of \mathcal{F} —the maximal possible precision. An illustration of the probabilities p_k and the corresponding bins b_k is given in Fig. 7.3(b) for $d = 8$.

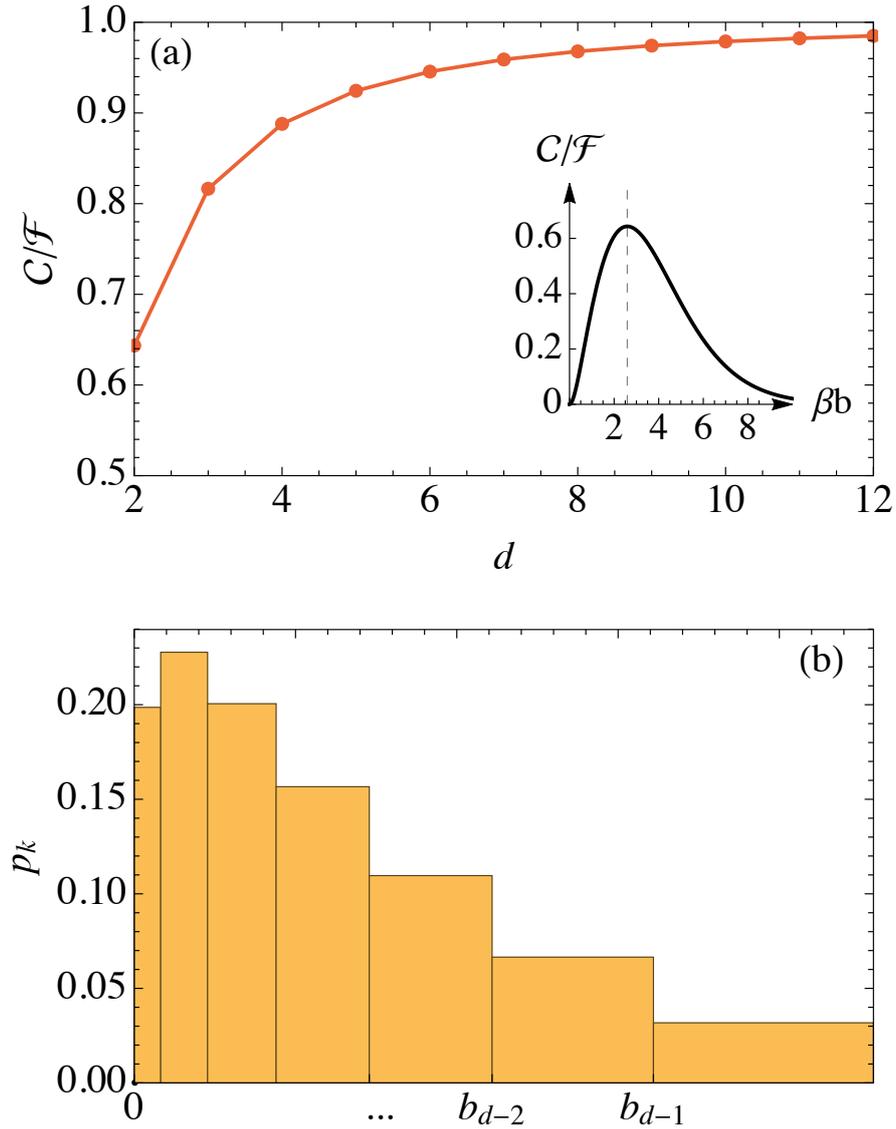


Figure 7.3: *This figure and caption is taken from Ref. [124] ©2021 American Physical Society.* Optimal thermometry binning for a system described by a linear density of states: $\Omega(E) \propto E$. (a) Optimal binned Fisher information, Eq. (7.17), as a function of d . The curves are normalized by the full Fisher information, which in this case reads $\mathcal{F} = 2\beta^2$. Inset: ratio C/\mathcal{F} as a function of the binning position b for $d = 2$. The optimal bin occurs at $\beta b = 2.58975$. The points in the main plot are already optimized over the binning positions. (b) Illustration of the corresponding probabilities p_k and average bins b_k for $d = 8$.

7.3.4 General remarks and extension to imperfect measurements

The two examples in Sec. 7.3.3 carry an important general message: even measurements so coarse-grained as to have only two outcomes can estimate the temperature of a generic system with precision (as quantified by the Fisher information) that is proportional to

the ultimate precision—the thermal Fisher information [Eq. (7.3)]. In the appendix of Ref. [124], we show that this is not a coincidence, by proving that any system for which the energy distribution is unimodal and has sufficiently fast decaying tails, displays a proportionality $\mathcal{C} \propto \mathcal{F}$. More specifically, we prove that there exists a finite number $\Xi \in [0, 1]$ such that $\mathcal{C} \geq \Xi \mathcal{F}$.

As will be discussed in Sec. 7.4, the relevance of these results lies in the fact that unimodal energy distributions with quickly decaying tails are a generic behaviour expected in finite-temperature many-body systems with short-range interactions, both at and away from criticality. In fact, we will show that $\Xi = 2/\pi$, as in Eq. (7.27), actually happens for a large variety of interacting lattice models. It is of course possible to conceive non-unimodal energy distributions for which the proportionality $\mathcal{C} \propto \mathcal{F}$ breaks down. This is illustrated in the appendix of Ref. [124], where we construct an example for which $\mathcal{C}/\mathcal{F} \rightarrow 0$ when $N \rightarrow \infty$.

Our framework can also be adapted to scenarios where the ideal energy binning cannot be implemented and imprecisions are present. In the simplest case, one could have some imprecision in determining the optimal bins through Eq. (7.24), which would lead to a decrease in \mathcal{C} . For the case $d = 2$, this is illustrated in the inset of Fig. 7.3(a), which shows how \mathcal{C}/\mathcal{F} decreases as the bin position deviates from its optimal value.

Another way in which imprecisions can enter our framework is when the POVMs (7.4) are noisy. For instance, experimental errors may cause energies close to a bin edge to sometimes result in outcomes corresponding to neighbouring bins. Such effects can be accounted for within our framework by modifying Eq. (7.21). To see that, we first rewrite Eq. (7.21) as

$$p_\alpha = \int_{-\infty}^{\infty} dE q(E) B_\alpha(E), \quad (7.29)$$

and similarly for ϵ_α . Here $B_\alpha(E)$ is a boxcar function, with value 1 when $E \in I_\alpha = [b_{\alpha-1}, b_\alpha)$, and 0 otherwise. A similar analysis can also be done at the discrete representation of Eq. (7.15). In this case, we would have $p_\alpha = \sum_i B_{\alpha,i} q_i$, where $B_{\alpha,i}$ is a matrix with entries 1 when $E_i \in I_\alpha$ and zero otherwise. However, it is more convenient to work with the continuous-energy representation (7.29).

It is now straightforward to generalize Eq. (7.29) to include the effects of noise by replacing $B_\alpha(E)$ by a different function. For instance, a smoothed boxcar as depicted in Fig. 7.4(a). Since $\sum_\alpha p_\alpha = 1$ for any initial distribution $q(E)$, it follows that $\sum_\alpha B_\alpha(E) = 1$ for all E . This can be viewed as a normalization condition for $B_\alpha(E)$. In fact, $B_\alpha(E)$ is actually a combination of a stochastic matrix (whose columns add up to one), plus an isometry, which reduces the dimension from a continuous energy E , to a discrete set of outcomes α . The precise form of $B_\alpha(E)$ will depend on the details of the experiment.

Measurement errors can cause not only a loss of precision, but also systematic shifts in energy by, e.g., falsely displacing the energies ϵ_α by a certain amount. For simplicity, we shall study these kinds of imprecision separately. We defer the discussion of robustness to energy shifts to Sections 7.4.1 and 7.4.2, while here we choose $B_\alpha(E)$ to be symmetric in the interval $[b_{\alpha-1}, b_\alpha)$, and centered at the midpoint $(b_{\alpha-1} + b_\alpha)/2$, so that the ϵ_α 's are not displaced.

The remaining feature to describe is errors associated with imperfect binning. This

can again be done using the smoothed boxcar of Fig. 7.4(a)

$$B_\alpha(E) = \frac{1}{2} \operatorname{erf} \left(\frac{b_\alpha - E}{\sigma\sqrt{2}} \right) - \frac{1}{2} \operatorname{erf} \left(\frac{b_{\alpha-1} - E}{\sigma\sqrt{2}} \right), \quad (7.30)$$

where $\operatorname{erf}(x)$ is the error function and σ is a parameter measuring the degree of imprecision (a sharp boxcar is recovered when $\sigma \rightarrow 0$). A function of this form defines a certain energy window 2σ , where measurements associated to one bin can be recorded in another. For this reason, the wider bins tend to be less affected than the thinner ones, which is physically reasonable.

We illustrate the above ideas with the linear DOS example of Sec. 7.3.3. In Fig. 7.4(b) we present \mathcal{C}/\mathcal{F} as a function of d for binning strategies computed using the smoothed boxcars (7.30), with different values of σ . This is contrasted with the ideal case, shown in red, which coincides with the red curve in Fig. 7.3(a). As can be seen, unsharp bin edges necessarily decrease the coarse-grained Fisher information. That said, \mathcal{C}/\mathcal{F} is surprisingly robust: even when the smearing occurs over a large part of the bin width (e.g., 30%), \mathcal{C}/\mathcal{F} does not decrease much (only about 7%).

7.4 Many-body lattice models

We now proceed to analyze quantum systems on a lattice, which is one of the most physically relevant settings where the coarse-grained measurements could be useful. We start with general considerations and a tight-binding chain as an illustrative example. Then we show a general result for all noncritical spin models and conclude with an analysis of a system undergoing a thermal phase transition.

7.4.1 Gaussian density of states

In many-body lattice models, the energy distribution (7.20) often displays an approximate Gaussian form in the thermodynamic limit [186, 187, 188] (see also the detailed discussion in Sec. 7.4.2 and in the appendix of Ref. [124]). As a simple, illustrative example of this principle, consider a fermionic one-dimensional tight-binding chain with N sites, under periodic boundary conditions:

$$H = \sum_{k=1}^N \varepsilon \hat{c}_k^\dagger \hat{c}_k - t \sum_{k=1}^N (\hat{c}_{k+1}^\dagger \hat{c}_k + \hat{c}_k^\dagger \hat{c}_{k+1}), \quad (7.31)$$

where \hat{c}_k is the fermionic annihilation operator at site k , ε is the on-site energy, t is hopping (tunnelling) strength, and $\hat{c}_{k+N} = \hat{c}_k$ ensures periodic boundary conditions. When diagonalized, the Hamiltonian of this model takes the form [189, 51]

$$H = \sum_{a=1}^N \varepsilon_a \hat{C}_a^\dagger \hat{C}_a, \quad (7.32)$$

with the (linearly) transformed \hat{C}_a 's satisfying standard fermionic anti-commutation relations, and with eigenenergies given by $\varepsilon_a = \varepsilon - t \cos(2\pi a/N)$.

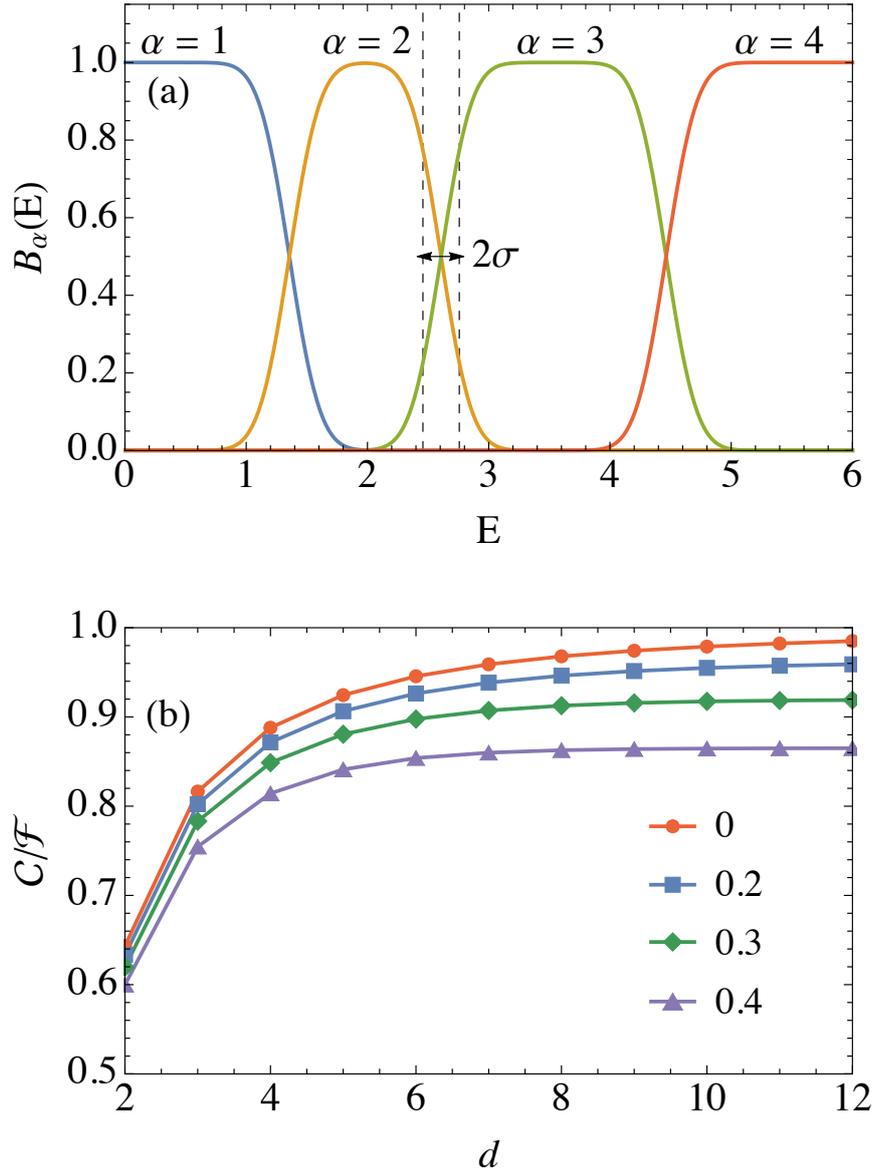


Figure 7.4: *This figure and caption is taken from Ref. [124] ©2021 American Physical Society.* The role of imperfect measurements in optimal thermometry. (a) Example of smoothed boxcars, Eq. (7.30), for $\sigma = 0.2$. (b) Illustration of the worsening of the coarse-grained Fisher information \mathcal{C} for the linear DOS example studied in Sec. 7.3.3. The red curve is the same as in Fig. 7.3(a), while the other curves were computed using the smoothed boxcars (7.30) with $\sigma = 0.2, 0.3$ and 0.4 (in units of $\beta = 1$).

In Fig. 7.5(a), we numerically compute the energy distribution (7.20) of this model, and compare the results to a continuous Gaussian distribution with average energy $\langle H \rangle$ and variance $\mu^2 \equiv \delta(H) = -\partial\langle H \rangle/\partial\beta$ (both of which depend implicitly on T); i.e.,

$$q(E) = \frac{1}{\sqrt{2\pi\mu^2}} e^{-\frac{(E-\langle H \rangle)^2}{2\mu^2}}. \quad (7.33)$$

We observe that the Gaussian distribution is a good approximation already with a modest number of sites and a modest hopping strength. The approximation, in fact, improves with the number of sites and becomes exact in the thermodynamic limit (see Refs. [187, 188]).

Let us now take a Gaussian distribution as a given and compute the Fisher information for different coarse-grainings of the continuous distribution (7.33). In this case, the probabilities and bin energies in (7.21) become

$$\begin{aligned} p_\alpha &= \frac{1}{2} [\operatorname{erf}(\tilde{b}_\alpha) - \operatorname{erf}(\tilde{b}_{\alpha-1})], \\ \epsilon_\alpha &= \langle H \rangle - \frac{\mu}{p_\alpha \sqrt{2\pi}} \left(e^{-\tilde{b}_\alpha^2} - e^{-\tilde{b}_{\alpha-1}^2} \right), \end{aligned} \quad (7.34)$$

where $\tilde{b}_\alpha := (b_\alpha - \langle H \rangle) / \sqrt{2}\mu$ (with $b_0 = -\infty$, $b_d = \infty$) are the shifted and rescaled bin positions. The full Fisher information is simply $\mathcal{F} = \beta^4 \mu^2$. For a d -outcome measurement, one can then numerically perform the optimization according to Eq. (7.24) to find the best such measurement and the corresponding coarse-grained Fisher information (7.17). The results for \mathcal{C}/\mathcal{F} for different numbers of bins d , are shown in Fig. 7.5(b). As in the linear density of states case, one sees a quick growth of \mathcal{C} with d towards the maximum value \mathcal{F} .

The particular case of $d = 2$ can be obtained by setting $b_0 = -\infty$, $b_2 = \infty$ and $b_1 = b$. We then get

$$p_{1,2} = \frac{1 \pm \operatorname{erf}(\tilde{b})}{2}, \quad \epsilon_{1,2} = \langle H \rangle \mp \frac{\mu}{\sqrt{2\pi} p_{1,2}} e^{-\tilde{b}^2},$$

so that the coarse-grained Fisher information [Eq. (7.17)] becomes

$$\mathcal{C} = \frac{2\mathcal{F}}{\pi} \frac{e^{-2\tilde{b}^2}}{1 - [\operatorname{erf}(\tilde{b})]^2}. \quad (7.35)$$

The result is expressed solely in terms of the shifted bin position \tilde{b} ; therefore, the minimization procedure is independent of $\langle H \rangle$ or μ . In fact, as one may readily verify, the function in Eq. (7.35) is maximized at $\tilde{b} = 0$. That is, the bin should be placed symmetrically, at $b = \langle H \rangle$. The corresponding maximum is

$$\mathcal{C} = \frac{2}{\pi} \mathcal{F}. \quad (7.36)$$

This relation is robust with respect to imprecise identification of the optimal boundary (which can be understood as a systematic error in the energy measurement, as mentioned in Sec. 7.3.4). Indeed, Taylor-expanding the right-hand side of Eq. (7.35) with respect to small \tilde{b} around its optimal value, $\tilde{b} = 0$, we find that $\frac{\mathcal{C}}{\mathcal{F}} = \frac{2}{\pi} [1 - 2\tilde{b}^2(1 - 2/\pi) + O(\tilde{b}^4)]$. Even for a significant deviation of $|b - \langle H \rangle| = 0.3\mu$, \mathcal{C}/\mathcal{F} degrades only by $\approx 3.3\%$.

Not coincidentally, the relation in Eq. (7.36) also appears in the case of non-interacting qubits in the limit of large N (Sec. 7.3.3). This is because the energy distribution in that case also becomes Gaussian in the $N \gg 1$ limit, due to the central limit theorem.

In Fig. 7.5(c), we illustrate the optimal bins and the corresponding probabilities for the distribution (7.33) in the case of $d = 8$. In this case, the optimal bins have to be located numerically. Unsurprisingly, it is found that the optimum is symmetric around the average energy.

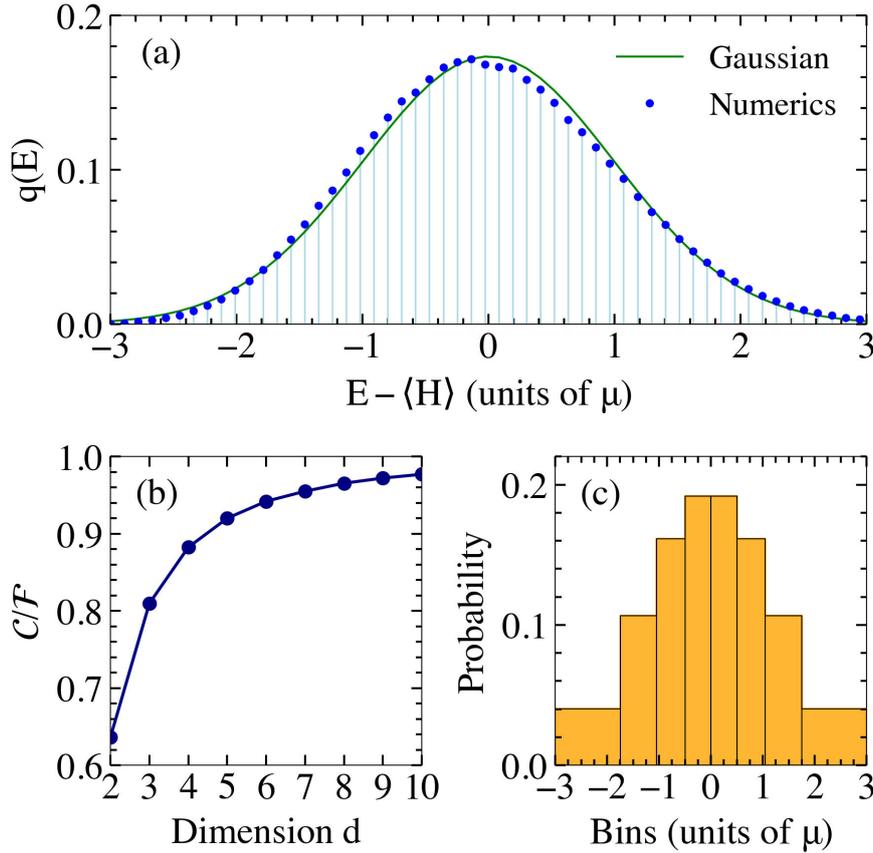


Figure 7.5: *This figure and caption is taken from Ref. [124] ©2021 American Physical Society.* (a) Energy distribution for a fermionic tight-binding chain with 20 sites. The hopping strength is $t = 0.3$, and the temperature is $T/\epsilon = 2$. The numerical results are compared to the Gaussian distribution (Eq. (7.33)). (b) Optimal binned Fisher information for the general Gaussian distribution as a function of d . The curves are normalized by the thermal Fisher information (Eq. (7.3)). (c) Illustration of optimal binning, and the corresponding probabilities, for the general Gaussian distribution with $d = 8$.

7.4.2 Noncritical, interacting systems on lattices

Now we will show how some of the conclusions of Sec. 7.4.1 actually hold universally in the thermodynamic limit. Intuitively speaking, the idea is that generic lattice models with finite-range interactions, when away from criticality, tend to have Gaussian energy distribution due to the many-body Berry-Esseen theorem [186, 187, 188]. Therefore, the same behaviour as in Fig. 7.5 is expected to occur when coarse-graining to different partitions in such lattice models.

In fact, in the appendix of Ref. [124], we prove that the maximal C/\mathcal{F} achievable by two-outcome measurements ($d = 2$) is $\frac{2}{\pi} + O(\ln^{-1} N)$, and the boundary of the optimal partition $I_1 = [E_0, b]$ and $I_2 = (b, E_D]$ is near the average energy:

$$b - \langle H \rangle = O(\ln^{-1/2} N) \sqrt{\delta(H)}. \quad (7.37)$$

Moreover, when $b = \langle H \rangle$ exactly, one still has $\frac{C}{\mathcal{F}} = \frac{2}{\pi} + O(\ln^{-1} N)$. In the thermodynamic limit, this coincides with the results for the exact Gaussian distribution

(Sec. 7.4.1) and for independent qubits (Sec. 7.3.3). We expect that, for $d \geq 3$ partitions, one should be able to prove results identical to those obtained for the exact Gaussian energy distribution in Sec. 7.4.1 by using arguments along the **lines of those given in the appendix of Ref. [124]**.

In order to prove Eq. (7.36), we need to assume that the thermal state of the lattice satisfies the following two generic conditions: (i) Exponential decay of correlations: For arbitrary regions \mathcal{X}, \mathcal{Y} separated by a distance l on the lattice, and some constant ξ ,

$$\max_{X \in \mathcal{X}, Y \in \mathcal{Y}} \left| \frac{\langle X \otimes Y \rangle - \langle X \rangle \langle Y \rangle}{\|X\| \|Y\|} \right| \leq e^{-l/\xi}, \quad (7.38)$$

and (ii) The variance in energy scales linearly with the number of lattice sites: $\delta(H) = \langle H^2 \rangle - \langle H \rangle^2 = s^2 N$. Assumption (i) is expected to hold for a very large class of systems away from criticality. Indeed, it has been rigorously proven for 1D translation-invariant thermal states [190], finite-range fermionic lattice systems of arbitrary spatial dimension at nonzero temperatures [191], and all finite-range lattice systems above a threshold lattice-dependent temperature [192]. Assumption (ii) is expected to hold for most systems at a high enough temperature. In fact, note that (i) already implies that $\delta(H) = O(N)$ [193].

The detailed proof of Eq. (7.36), provided in the appendix of Ref. [124], is based on the Berry-Esseen theorem for local Hamiltonians which relies on the two assumptions above and is proven in Refs. [187, 188]. This result can be seen as a strengthening of the central limit theorem, which gives a precise notion of how the energy distribution of lattice models converges to a Gaussian in the thermodynamic limit. It allows us to estimate functions of the form of Eq. (7.19), in this case, the bin energies ϵ_k .

Lastly, the Gaussian behaviour of noncritical many-body lattice systems also extends to the problem of how robust $\mathcal{C}/\mathcal{F} = 2/\pi$ is to imprecise identification of the optimal binning boundary. Indeed, as we show in the appendix of Ref. [124], for small $\tilde{b} = \frac{b - \langle H \rangle}{\sqrt{2\delta(H)}}$,

$$\frac{\mathcal{C}}{\mathcal{F}} = \frac{2}{\pi} [1 - 2\tilde{b}^2(1 - 2/\pi) + O(\ln^{-1} N) + O(\tilde{b}^3)],$$

so, as in Sec. 7.4.1, for e.g. $|b - \langle H \rangle| = 0.3\sqrt{\delta(H)}$ (and $N \gg 1$), \mathcal{C}/\mathcal{F} will degrade only by $\approx 3.3\%$.

7.4.3 Critical systems

The thermal Fisher information (7.3) is proportional to the heat capacity of the system, which scales as $C = Nc(\beta)$, where $c(\beta)$ is the specific heat. For noncritical systems, $c(\beta)$ is intensive. However, at a finite-temperature phase transition, it diverges as the temperature of the system approaches the critical temperature $T_c > 0$, according to $c(\beta) \propto |t|^{-\alpha}$, where $t := (\beta - \beta_c)/\beta_c$ and $\alpha \geq 0$ is the so-called critical exponent [194]. When $\alpha = 0$, the divergence is logarithmic: $c(\beta) \propto \ln |t|^{-1}$ [194]. In large but finite systems, there are of course no infinities and, at the phase transition point, $c_N(\beta_c) \propto N^{\frac{\alpha}{2-\alpha}}$ when $\alpha > 0$ [195] and $c_N(\beta_c) \propto \ln N$ when $\alpha = 0$ [195, 196]. Since $c_N = \beta^2 \delta(H)/N$ [Eq. (7.3)], the divergence of $c_N(\beta_c)$ with N implies that critical systems do not satisfy the condition (ii) of Sec. 7.4.2. In general, critical systems also feature diverging correlation lengths [194], thereby violating the condition (i) of Sec. 7.4.2 as

well. Therefore, the many-body Berry-Esseen theorem becomes inapplicable for critical systems.

In the appendix of Ref. [124], building on several rigorous results on translation-invariant lattices with finite-range interactions in Refs. [197, 189, 198], we develop an alternative approach towards determining the energy distribution of such lattices in arbitrary spatial dimensions. First of all, for noncritical lattices, we show that the energy distribution approximates a Gaussian in a way that complements the many-body Berry-Esseen theorem [187, 188]. Moreover, for this wide but specific class of lattices, our approach allows us to access the energy distribution even at criticality.

For critical lattices with $\alpha = 0$, we show in the appendix of Ref. [124] that the energy distribution still tends to a Gaussian in the $N \rightarrow \infty$ limit; however, the convergence does not include the tails of the distribution, which are $O(\sqrt{N})$ standard deviations away from $\langle H \rangle$. In a sense, for translation-invariant lattices, this result suggests that the Gaussianity of the distribution holds beyond assumptions (i) and (ii) above [187, 188]. Thus, Eq. (7.36) applies in the thermodynamic limit, both at criticality (with $\alpha = 0$) and away from it. We illustrate these ideas below in Sec. 7.4.3, with a detailed study on the classical 2D Ising model on a square-lattice, a paradigmatic model with $\alpha = 0$.

The case $1 > \alpha > 0$ is treated in the appendix of Ref. [124]. We show that the energy distribution is Gaussian only in a neighbourhood of the peak that is much smaller than the standard deviation. Hence, it is not Gaussian as a whole. Notwithstanding, we show that it is unimodal with exponentially decaying tails, which means that the considerations in the appendix of Ref. [124] are applicable; that is, a two-bin measurements with the boundary placed at $\langle H \rangle$ will yield a \mathcal{C} that scales proportionally to \mathcal{F} . In other words, since $\mathcal{F} = \beta^2 N c_N$, we will have $\mathcal{C} \propto \beta^2 N^{\frac{2}{2-\alpha}}$.

2D Ising model

The square-lattice 2D Ising model is defined on an $L \times L$ square lattice where each site i is characterized by a Pauli matrix σ_z^i , with $i = 1, \dots, N$ ($N = L^2$). The spins interact according to the Hamiltonian

$$H = -J \sum_{\langle i,j \rangle} \sigma_z^i \sigma_z^j, \quad (7.39)$$

where the sum is over all nearest neighbours. Since the interactions only involve σ_z operators, the Hamiltonian is already diagonal in the computational basis, with energy eigenvalues

$$E(\boldsymbol{\sigma}) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j, \quad (7.40)$$

where $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_N)$ and $\sigma_i = \pm 1$ are the eigenvalues of σ_z^i . Here, we will impose periodic boundary conditions. The model presents a phase transition at $T_c/J = 2/\ln(1 + \sqrt{2})$ [201, 202, 199]. This can be seen, for instance, in terms of the magnetization $m = \frac{1}{N} \sum_i \langle \sigma_i \rangle$, as plotted in Fig. 7.6(a).

For not very large N 's, the full energy distribution $q(E)$ can be computed exactly using a method developed in Ref. [200]. Results for $L = 8, 16, 32$, and 64 are shown in Fig. 7.6(b)-(e). Although irregular for small sizes, it can be seen that the distribution visually appears to approach a Gaussian as the lattice size increases.

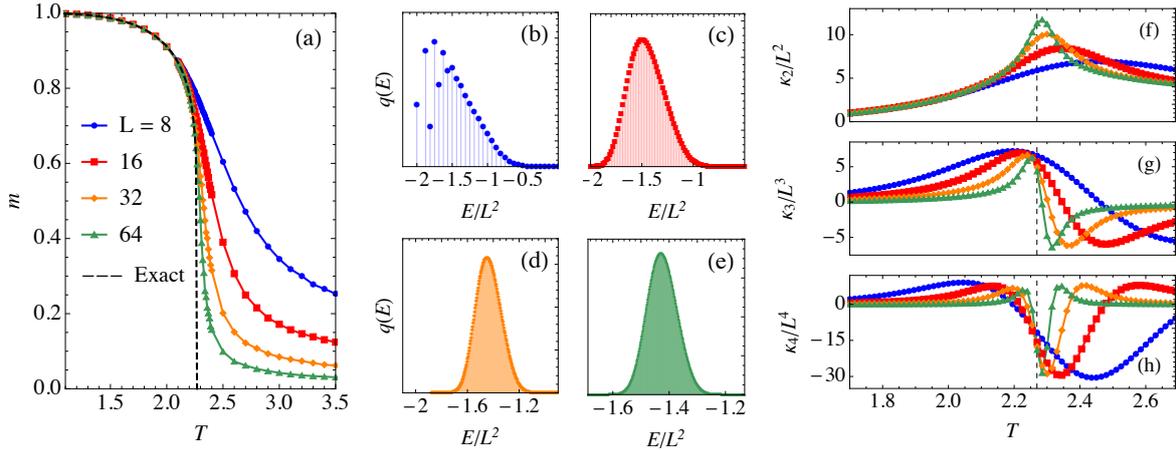


Figure 7.6: *This figure and caption is taken from Ref. [124] ©2021 American Physical Society.* Classical two-dimensional Ising model [Eq. (7.40)]. (a) Magnetization as a function of temperature for different lattice sizes, showing the phase transition at $T_c/J = 2.26919$. The dashed curve corresponds to the thermodynamic limit (where the exact solution is $m = [1 - \sinh^{-4}(2J/T)]^{1/8}$ [199]). (b)-(e) Exact energy distributions $q(E)$ for $T = T_c$, computed using the method of Ref. [200]. Each curve corresponds to a different lattice size, $L = 8, 16, 32,$ and 64 (same color code as image (a)). (f)-(h) Cumulants $\kappa_2, \kappa_3,$ and κ_4 of $q(E)$, as a function of temperature, for different lattice sizes. The third and fourth cumulants scale, respectively, as $N^{3/2}$ and N^2 , where $N = L^2$ is the number of sites. The second cumulant, on the other hand, scales as $N \ln N$.

In order to rigorously prove that this is indeed the case, one needs to show that the higher-order (≥ 3) cumulants of the energy distribution, κ_k , become irrelevant as N becomes large. As discussed in the appendix of Ref. [124], the cumulants of the energy distribution can be found from the free energy F_N through the simple relation

$$\kappa_k = (-1)^{k-1} \frac{\partial^k (\beta F_N)}{\partial \beta^k}. \quad (7.41)$$

Moreover, for the 2D Ising model, an exact expression for F_N , for finite N , is available [203]. Using these facts, we show in the appendix of Ref. [124] that $\kappa_2 \propto N \ln N$ [194, 196], while $\kappa_k \propto N^{k/2}$, for $k \geq 3$ (these results are also illustrated in Fig. 7.6(f)-(h)). As a consequence, we therefore have that

$$\frac{\kappa_k^{1/k}}{\kappa_2^{1/2}} \propto \ln^{-1/2} N, \quad (7.42)$$

showing that the higher-order cumulants do indeed become negligible as compared to $\kappa_2 = \delta(H)$; i.e., the distribution tends to a Gaussian as $N \rightarrow \infty$. Note that the previous discussion refers to the scaling at the vicinity of the critical point. Away from it, due to the extensivity and analyticity of F_N in the limit of $N \rightarrow \infty$, we simply have from Eq. (7.41) that $\kappa_k \propto N$ for $k \geq 1$. Hence, $\kappa_k^{1/k} / \kappa_2^{1/2} \propto N^{-\frac{k-2}{2k}}$, i.e., away from criticality, $q(E)$ approaches its Gaussian limit polynomially, as compared to the slow logarithmic convergence at criticality.

7.5 Probe-based measurements

Whereas the previous section allowed for arbitrary global measurements, in this section, we look at measurement schemes (both idealized and more realistic) which can be realised by the interaction of a probe with the system, possibly some auxiliary system of arbitrary size, and the subsequent measurement of the probe alone. We compare the performance of such probe-based measurements to the upper bounds obtained in Section 7.3.2.

First of all, we observe that the maximal thermometric precision achievable by measuring a d -dimensional probe P that has unitarily interacted with the system S in a thermal state τ , and an auxiliary system \mathcal{A} in some state $\rho_{\mathcal{A}}$, is the same as the maximal precision of a d -outcome measurement on S . Here we assume that $d < D$, because otherwise one can simply transfer all of the state of S —and the (Fisher) information on β , \mathcal{F} , along with it—to P ; however, when $d < D$, even the best possible strategy of encoding the state of S into that of P will bear losses. Indeed, if the initial state of P is some σ , then, whatever the optimal unitary U , standard quantum metrology tells us [4] that the optimal POVM on

$$\sigma' = \text{Tr}_{S\mathcal{A}} \{U(\sigma \otimes \tau \otimes \rho_{\mathcal{A}})U^\dagger\} \quad (7.43)$$

will have to have d outcomes. On the other hand, the probability distribution generated by any d -outcome POVM on σ' can also be generated by a d -outcome POVM on S . Thus, denoting the quantum Fisher information of σ' on β by $\mathcal{C}^{(P)}$, we have $\mathcal{C}^{(P)} \leq \mathcal{C}$.

To show that $\mathcal{C}^{(P)} = \mathcal{C}$, let us note that \mathcal{C} is delivered by a projective measurement on the system corresponding to some binning $I_1 \cup \dots \cup I_d$ which yields measurement statistics $p_\alpha = \sum_{q_j \in I_\alpha} q_j$. Now, let us choose $\sigma = |1\rangle\langle 1|$, so that, in the $\{|\alpha\rangle \otimes |E_j\rangle\}$ basis, $\sigma \otimes \tau = \text{diag}(\vec{\mathbf{q}}, \vec{\mathbf{0}}, \dots, \vec{\mathbf{0}})$. Here, $\vec{\mathbf{q}} = (q_1, \dots, q_D)$ and $\vec{\mathbf{0}}$ is made of D zeroes. Then, the permutation unitary acting on $\sigma \otimes \tau$ that permutes all the q_j 's in I_2 from $\vec{\mathbf{q}}$ with some of the zeroes from the $\vec{\mathbf{0}}$ next to $\vec{\mathbf{q}}$, all the q_j 's in I_3 with zeroes from the second $\vec{\mathbf{0}}$, etc., will render $\sigma' = \text{diag}(p_1, \dots, p_d)$. And this distribution will produce a $\mathcal{C}^{(P)}$ that is $= \mathcal{C}$. Note that, to show that $\mathcal{C}^{(P)}$ can be made equal to \mathcal{C} , there was no need to involve any auxiliary systems.

For the case of $d = 2$, where the optimal POVM on S is defined by the bins $I_1 = \{E_j : E_j < b\}$ and $I_2 = \{E_j : E_j \geq b\}$, with b being the boundary, we will now show that such a permutation can be generated by the quantum-optics-inspired Hamiltonian

$$H = \sum_{k=1}^D E_k |E_k\rangle\langle E_k| + b |\uparrow\rangle\langle\uparrow| + \lambda \sum_{E_k \leq E_D - b} (|\downarrow\rangle\langle\uparrow| \otimes |E_k + b\rangle\langle E_k| + \text{H.c.}), \quad (7.44)$$

where $|\downarrow\rangle$ and $|\uparrow\rangle$ are the eigenstates of the probe spin, with the corresponding eigenvalues $E_\downarrow = 0$ and $E_\uparrow = b$. This Hamiltonian may not be easily realizable in practice. However, the point is that, as we will show, it is guaranteed to provide the best possible precision using a two-level probe. This can then be used as a benchmark to compare against when using other interactions. Furthermore, we take the system's ground state to be at energy $E_1 = 0$, and since ultimately we are going to be interested in the case where the system's energy spectrum is effectively continuous, we will also assume that $|E_k + b\rangle$ is a valid eigenstate.

Let us initialize the system and the probe in the joint state $\rho(0) = |\Downarrow\rangle\langle\Downarrow| \otimes \tau$. The simplest way to characterize its evolution under H is to describe how the pure states comprising it, $|\Psi_j(0)\rangle = |\Downarrow\rangle \otimes |E_j\rangle$, evolve under H . It is easy to show that, in the interaction picture (labelled by the superscript I), $|\Psi_j^I(t)\rangle = |\Downarrow\rangle \otimes |E_j\rangle$ for $E_j < b$, and

$$|\Psi_j^I(t)\rangle = \cos(\lambda t) |\Downarrow\rangle \otimes |E_j\rangle - i \sin(\lambda t) |\Uparrow\rangle \otimes |E_j - b\rangle$$

for $E_j \geq b$. Thus, transitioning back to the Schrödinger picture, from

$$\rho(t) = \sum_j \frac{1}{Z} e^{-\beta E_j} |\Psi_j(t)\rangle\langle\Psi_j(t)| \quad (7.45)$$

we find the probability of finding the probe qubit in the spin-up state, \mathbb{P}_{\Uparrow} , when measuring it at the moment of time t , to be

$$\mathbb{P}_{\Uparrow}(t) = \sin^2(\lambda t) \sum_{E_j \geq b} \frac{e^{-\beta E_j}}{Z}. \quad (7.46)$$

Hence, for $t_{\text{meas}} = \pi/(2\lambda)$, the ideal projective measurement of the probe qubit's spin produces a probability distribution identical to that produced by the ideal binary measurement of the system corresponding to the binning $I_1 \cup I_2$.

Note that realizing this idealized scheme experimentally is far from being straightforward; we study a specific model realization in the next subsection, describing both its capabilities and limitations.

7.5.1 Jaynes–Cummings model

As a specific illustration of quantum probe-based thermometry, we consider an experimentally relevant system consisting of a superfluid Bose-Einstein condensate (BEC) reservoir in a shallow confining trap interacting with an atomic quantum dot [178]. Generally, the physics of this system is well-captured by a spin-boson model, in which the atomic quantum dot interacts with the phononic excitations of the BEC superfluid. Given suitably engineered boundary conditions, the spectral density will be such that a quantum dot with frequency ω_d will in effect couple only to the phonon modes which comes closest to be resonant with the quantum dot frequency (for simplicity we suppose all relevant phonon modes have the same frequency ω_a).

We can then ask the question of how well one can estimate the BEC temperature by measurements on the quantum dot probe. In Fig. 7.7(a) we plot the coarse-grained vs thermal Fisher information ratio for the optimal binary measurement within the effectively resonant subspace. The ratio is given as a function of the number of effectively resonant modes, which would be expected to increase in proportion to the width of the spectral density. From the figure we see that the ratio approaches a value of 0.64 as the number of modes increase. This provides an optimal value against which to compare specific binary measurement strategies. Interestingly we see that the obtained ratio agrees with the optimal ratio found for a binary measurement on a system described by a Gaussian density of states.

If we now consider the specific case where the effectively resonant subspace consist of a single phononic mode, and furthermore make a rotating wave approximation, the resulting system is modelled by the paradigmatic Jaynes-Cummings Hamiltonian [204]:

$$H = \omega_d \sigma^\dagger \sigma + \omega_a a^\dagger a + g (\sigma^\dagger a + \sigma a^\dagger), \quad (7.47)$$

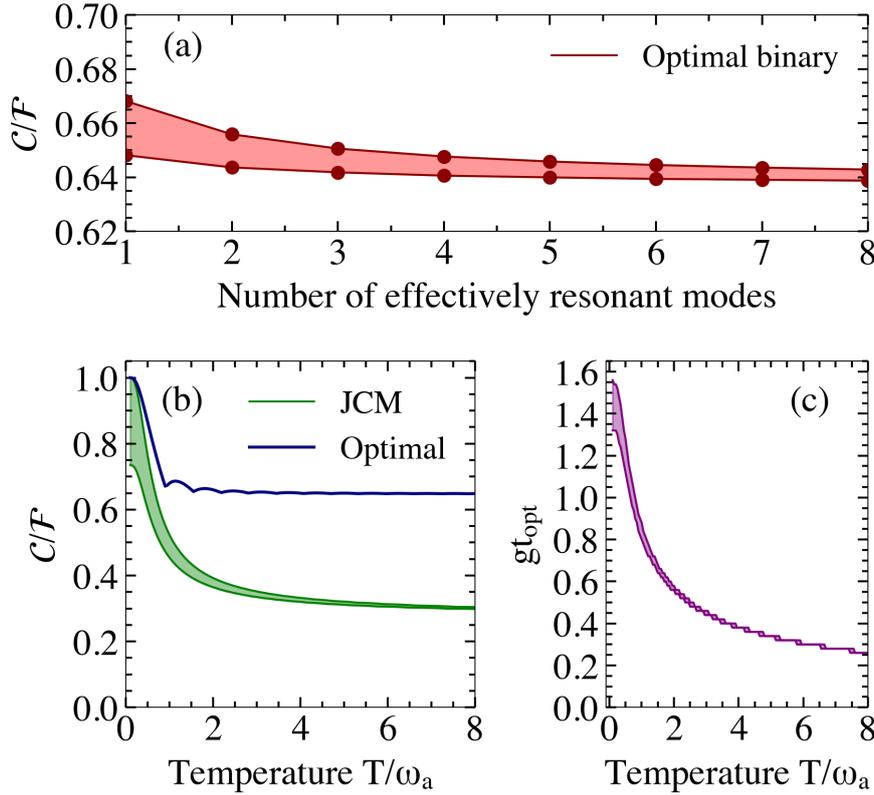


Figure 7.7: *This figure and caption is taken from Ref. [124] ©2021 American Physical Society.* (a) Optimal binary-outcome thermometry on a collection of bosonic modes of frequency ω_a as a function of the number of oscillators. The top line corresponds to $\beta\omega_a = 0.7$, the bottom line to $\beta\omega_a = 0.1$ and intermediate temperatures are contained within the shaded area. (b) Comparison between the optimal binary measurement strategy, and the two-level Jaynes-Cummings model probe, optimized over the measurement time gt , in the single oscillator case. The shaded area represents the range $\delta/g = [0, 1.2]$ of the detuning. (c) Optimal measurement time (gt_{opt}) as a function of temperature for the Jaynes-Cummings model.

where a, a^\dagger are the creation and annihilation operators of the bosonic cavity mode, g is the coupling strength, and $\sigma = |g\rangle\langle e|$ where the excited and ground states of the quantum dot are denoted by $|e\rangle$ and $|g\rangle$ respectively. Experimental work has shown how such models arise for specific thermometry protocols [159], and our results makes it possible to evaluate how close such a strategy is to being optimal.

Consider a measurement protocol in which the quantum dot is initialized in its ground state. The quantum dot then evolves jointly with the BEC for a time t , after which the probability of finding the quantum dot in its excited state is give by

$$\mathbb{P}_e(t) = \sum_{n=0}^{\infty} \frac{e^{-\beta\omega_a(n+1)}}{Z_\beta} \frac{g^2(n+1)}{\lambda_n^2} \sin^2(\lambda_n t), \quad (7.48)$$

where we have defined $\lambda_n = \sqrt{\delta^2/4 + g^2(n+1)}$ in terms of the detuning $\delta = \omega_d - \omega_a$. From this probability we can compute the coarse-grained Fisher information, see also [164]. In Fig. 7.7(b) we show the ratio of the coarse-grained Fisher information computed

from Eq. (7.48) and optimized over the measurement time at each temperature, with the thermal Fisher information of the phononic mode itself. Inspection of the results show that the probe-based measurement gives a Fisher information which never falls below 45% of the optimal binary measurement strategy. In Fig. 7.7(c) we plot the optimal measurement time as a function of temperature, and we observe an inverse relationship between the optimal measurement time and the temperature. Notice that as the temperature approach absolute zero, the optimal measurement time for zero detuning approach $\pi/2g$ in agreement with known results [164].

It is interesting to note that similar considerations for the Fisher information have been obtained for temperature measurements of micromechanical resonators via a qubit probe, whose interaction can also be described by the Jaynes-Cummings (7.47) [164]; and more general interaction Hamiltonians, either dropping the rotating wave approximation or considering interactions far off resonance, have also been considered [165, 162]. In all such cases, our considerations provide upper bounds on the maximal precision estimation with a qubit probe, as shown in Fig. 7.7. Indeed, the strength of our bounds is that they apply to arbitrary nonequilibrium strategies.

7.6 Conclusions

We considered the precision limits on temperature estimation when having access to coarse-grained measurements which have at most d outcomes. Using tools from signal processing, we derived the structure of the optimal POVM measurement. These abstract considerations have been applied to two physically relevant scenarios: temperature measurements of many-body systems and nonequilibrium thermometry.

For many-body systems, we considered spin lattices, both away from and at criticality, and found that the Fisher information \mathcal{C} can grow extensively with the system size even when d does not. In particular, for $d = 2$, we obtain that it is in principle possible that $\mathcal{C}/\mathcal{F} \approx 2/\pi$ in the thermodynamic limit $N \rightarrow \infty$ even for systems at criticality. While this will decrease for realistic strategies where the POVM are smoothed out (see the discussion in Sec. 7.3.4 and Fig. 7.4 specifically), we expect that the extensive scaling will be preserved as long as the binary measurement can distinguish system energies that are $O(\sqrt{\delta(H)})$ apart (see the discussion on displaced boundary in Sections 7.4.1 and 7.4.2).

Along the way, we also derived new results on the energy distribution of many-body systems in the regime of criticality, which might be of independent interest. These generic considerations were illustrated on the 2D Ising model, the energy distribution of which becomes well-approximated by a Gaussian distribution except in the tails of the distribution. We expect more pronounced non-Gaussian features in the energy distribution of other critical models, which we leave as interesting future research.

For nonequilibrium thermometry, we used our results to devise the optimal probe-system interaction and interrogation time, thus providing general guidelines on the design of optimal nonequilibrium thermometry strategies. This result also provides an upper bound on specific experimentally motivated protocols. This was illustrated for a temperature measurement of a Bose-Einstein condensate through a quantum dot via a Jaynes-Cummings interaction [164, 165]. It remains an exciting open question to find a realistic implementation of the optimal probe-sample interaction (7.44).

Lastly, in this work, we focused on asymptotic metrology, where one has access

to full measurement statistics and can possibly run the experiment many times. This may not always be feasible in practice, and the Fisher information may then no longer be an adequate precision quantifier. In such cases, alternative approaches are needed, such as global Bayesian estimation [78]. Analyzing the effect of coarse-graining in such situations is another interesting research direction.

Appendix A

Probability theory

A stochastic experiment is a repeatable physical process in which a random outcome is realized. In this section, following the treatment of U. Thygesen [36] and S. Gammelmark [35], we provide a brief introduction to the measure-theoretic probability theory used to model such experiments.

Probability measure spaces

We begin with the concept of a sample space M , which contains all possible outcomes m of an experiment, and note that this space could be discrete or continuous. In any realization of the experiment, exactly one of the potential outcomes is realized. Next we define events: an event is a subset of outcomes in the sample space, identified by an affirmative answer to a yes/no question. As an example we could consider the roll of a die, the sample space is $\{1, 2, 3, 4, 5, 6\}$, a question could be whether or not the die shows an even number and the corresponding event is the subset $\{2, 4, 6\}$ of the full sample space.

Our aim is to construct a framework in which we can assign probabilities to events. The first difficulty encountered, is that it is not generally possible to assign a probability to any conceivable event. The subset of measurable events, i.e. those to which probabilities can be assigned, forms the σ -algebra of events, denoted \mathcal{M} . A valid σ -algebra satisfies a number of requirements: it includes the certain event $M \in \mathcal{M}$, it is closed under set complements, and it must be closed under countable operations of union and intersection.

Given a sample space and a σ -algebra, the map which takes an event in the σ -algebra to a probability is called a probability measure. Generally, a measure is a map $\mathbb{P} : \mathcal{M} \rightarrow [0, \infty)$ which is countably additive i.e. given a countable family of events which are mutually exclusive we have $\mathbb{P}(\cup_i A_i) = \sum_i \mathbb{P}(A_i)$. A probability measure furthermore satisfies that $\mathbb{P}(M) = 1$, i.e. the probability that any outcome is obtained is one. The triple $(M, \mathcal{M}, \mathbb{P})$ is called a probability measure space, and it constitutes the basic mathematical model of any stochastic experiment.

Measurable functions & Random variables

We now introduce a function which maps an outcome from the measurable space (M, \mathcal{M}) into an element of another measurable space (N, \mathcal{N}) . Just as in the case of events, there are functions which are not measurable. To explain the meaning of measurable in this

context we define the pre-image: given a function $X : M \rightarrow N$ and a set $A \subset N$, the pre-image of the function is defined as the set of outcomes in M which maps to A , i.e.

$$X^{-1}(A) \equiv \{m \in M : X(m) \in A\}. \quad (\text{A.1})$$

Notice that the pre-image is not equivalent to the inverse of the function, this point will be important later. Now given the pre-image, we say that a function X is measurable if the pre-image of any event in \mathcal{N} is an event in \mathcal{M} , i.e.

$$X^{-1}(A) \in \mathcal{M} \text{ for all } A \in \mathcal{N}. \quad (\text{A.2})$$

Measurable functions are particularly interesting when dealing with a probability measure space $(M, \mathcal{M}, \mathbb{P})$. If we define an image measure on (N, \mathcal{N}) with respect to a measurable function X by $\mathbb{P}_X(A) \equiv \mathbb{P}(X^{-1}(A))$, then the image measure will itself be a probability measure. A special case of particular interest, is that of measurable functions mapping outcomes in the sample space into the space of real numbers \mathbf{R}^n . In this case, the measurable function is called a random variable. When considering the space of real numbers the standard σ -algebra is the Borel-algebra, denoted $\mathcal{B}(\mathbf{R}^n)$, which is defined as the smallest σ -algebra containing all open sets [36].

Integration & Expectation values

We now make use of the above ideas to construct the integral of a random variable. The integral of a one-dimensional random variable $X : M \rightarrow \mathbf{R}$ over a measure space $(M, \mathcal{M}, \mathbb{P})$, is constructed as follows: the measure assigns weights to the set of outcomes in M and weighs the random variable X accordingly, the integral is then the sum of these weighted contributions. If the integral is to be computed over a specific event $A \in \mathcal{M}$, then we write it as

$$\int_A d\mathbb{P}(m) X(m) \in \mathbf{R}. \quad (\text{A.3})$$

If we compute the integral over the full sample space, then the computed quantity is called the expectation value of X . Among other things, we can define the cumulative distribution function of a real-valued random variable in one-dimension: $F_X(x) \equiv \mathbb{P}_X((-\infty, x])$. The probability density is then the Radon-Nikodym derivative of the image measure with respect to the Lebesgue measure, i.e. $p_X(x) = d\mathbb{P}_X/d\mu_L(x)$ where μ_L is the Lebesgue measure on \mathbf{R} [35]. In these terms we can write the expectation value of X as

$$\begin{aligned} \mathbb{E}\{X\} &= \int_M d\mathbb{P}(m) X(m) \\ &= \int_{\mathbf{R}} d\mu_L(x) p_X(x) x, \end{aligned} \quad (\text{A.4})$$

which corresponds to the standard formula for an expectation value. The above ideas generalize naturally to functions of random variables, and random variables in higher dimensions. A more formal introduction to the concept of integration has been given by Thygesen [36].

Available & Generated information

We now consider how to model an observer, or multiple observers, with only partial knowledge about the outcome of a stochastic experiment, e.g. two dice are cast and the observer is only told the larger of the two outcomes. Recall that an event is a subset of outcomes in the sample space, identified by an affirmative answer to a question. We say that an observer can resolve an event, if her partial knowledge about the realized outcome allows her to know if the event has occurred or not. The set of events which can be resolved, forms a sub- σ -algebra to \mathcal{M} , and we say that the partial information available to an observer is represented by this sub- σ -algebra of events¹.

In almost all cases of physical interest, the knowledge available about the realized outcome is obtained by observing a random variable (e.g. measuring a current). If we let $X : M \rightarrow \mathbf{R}^n$ be a random variable on the probability space $(M, \mathcal{M}, \mathbb{P})$, then we define the information generated by X as the sub- σ -algebra:

$$\sigma(X) \equiv \{A \in \mathcal{M} : A = X^{-1}(B)\} \quad (\text{A.5})$$

for some event B in the Borel-algebra $\mathcal{B}(\mathbf{R}^n)$. The generated information represents the subset of questions which can be answered based on the observation of the realized value of the random variable.

One could also ask the related question: under what conditions will an observer with available information represented by the sub- σ -algebra \mathcal{H} , be able to infer the value of a random variable X which was realized. When this can be done we say that X is \mathcal{H} -measurable. If we suppose that one observer measures a random variable X and that another observer measures a random variable Y , such that the generated information is $\sigma(X)$ and $\sigma(Y)$ respectively. Then the answer to the above question is provided by the Doob-Dynkin lemma which states: If $X : M \rightarrow \mathbf{R}^m$ and $Y : M \rightarrow \mathbf{R}^n$ are both random variables on a probability measure space $(M, \mathcal{M}, \mathbb{P})$, then X is $\sigma(Y)$ -measurable if and only if there exist a Borel-measurable function $g : \mathbf{R}^n \rightarrow \mathbf{R}^m$ such that $X(m) = g \circ Y(m)$ for all $m \in M$.

On the other extreme, we could have complete independence of information. In contrast to the concepts above, independence depends on the probability measure. We say that two events $A, B \in \mathcal{M}$ are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Two σ -algebras are independent if all events in one are independent of events in the other. Lastly, we say that two random variable are independent if the information algebras they generate are independent.

Conditional expectations

We now ask what does an observer want to do with the available information? Basically, the available information is used in the construction of *conditional expectations* of random variables. Formally, conditioning is with respect to a sub- σ -algebra \mathcal{H} , and is written

$$\mathbb{E}\{X | \mathcal{H}\}. \quad (\text{A.6})$$

We define it as follows: given a random variable X on $(M, \mathcal{M}, \mathbb{P})$ and a sub- σ -algebra \mathcal{H} , the conditional expectation of X w.r.t. \mathcal{H} is the *almost surely* unique random

¹Notice that we use *knowledge* and *information* as distinct concepts here. Knowledge refers to the specific realization of the experiment, information refers to the set of questions which can be answered based on our knowledge of the realization.

variable $Z \equiv \mathbb{E}\{X | \mathcal{H}\}$ which is measurable w.r.t. \mathcal{H} and which satisfies the consistency requirement

$$\mathbb{E}\{Z \cdot \mathbf{1}_H\} = \mathbb{E}\{X \cdot \mathbf{1}_H\}, \quad (\text{A.7})$$

for any $H \in \mathcal{H}$. Here we have introduced the indicator function $\mathbf{1}_H$ of H defined such that $\mathbf{1}_H(m)$ equals one if $m \in H$ and zero otherwise. Informally, we can think of the conditional expectation as the effective random variable which is obtained after everything but \mathcal{H} has been averaged out.

Leaving the question of how to compute the conditional expectation aside for now, we can discuss some key properties. First of all, the conditional expectation satisfies a number of important mathematical relations [35]:

$$\mathbb{E}\{aX + bY | \mathcal{H}\} = a\mathbb{E}\{X | \mathcal{H}\} + b\mathbb{E}\{Y | \mathcal{H}\} \quad (\text{A.8})$$

$$\mathbb{E}\{\mathbb{E}\{X | \mathcal{G}\}\} = \mathbb{E}\{X\} \quad (\text{A.9})$$

$$\mathbb{E}\{\mathbb{E}\{X | \mathcal{G}\} | \mathcal{H}\} = \mathbb{E}\{X | \mathcal{H}\} \quad \text{for } \mathcal{G} \subset \mathcal{H} \quad (\text{A.10})$$

$$\mathbb{E}\{X | \mathcal{H}\} = X \quad \text{if } X \text{ } \mathcal{H}\text{-meas.} \quad (\text{A.11})$$

$$\mathbb{E}\{XY | \mathcal{H}\} = X\mathbb{E}\{Y | \mathcal{H}\} \quad \text{if } X \text{ } \mathcal{H}\text{-meas.} \quad (\text{A.12})$$

$$\mathbb{E}\{X | \mathcal{H}\} = \mathbb{E}\{X\} \quad \text{if } X \text{ } \mathcal{H}\text{-indep.} \quad (\text{A.13})$$

If we consider two measures \mathbb{P}_μ and \mathbb{P}_ν , and suppose that \mathbb{P}_μ is absolutely continuous with respect to \mathbb{P}_ν . The Radon-Nikodym derivative is written as $R_{\mu\nu} = d\mathbb{P}_\mu/d\mathbb{P}_\nu$. When the Radon-Nikodym derivative is well-defined, we have the important relation:

$$\mathbb{E}_\mu\{X | \mathcal{H}\} \mathbb{E}_\nu\{R_{\mu\nu} | \mathcal{H}\} = \mathbb{E}_\nu\{R_{\mu\nu}X | \mathcal{H}\}, \quad (\text{A.14})$$

where \mathbb{E}_μ labels the measure employed. This relation can be interpreted as the measure-theoretic version of Bayes rule. Lastly we mention that in the case where $\mathcal{H} = \sigma(Y)$, we know from the Doob-Dynkin lemma that there must exist a Borel-measurable function such that $\mathbb{E}\{X | \mathcal{H}\} = g \circ Y$, i.e. the conditional expectation is a function of Y .

The introduction of conditional expectations makes it possible to construct other conditional statistics as well. In particular we can define the conditional probability of an event:

$$\mathbb{P}(A | \mathcal{H}) \equiv \mathbb{E}\{\mathbf{1}_A | \mathcal{H}\}, \quad (\text{A.15})$$

which must be understood as a random variable, and we recall that $\mathbf{1}_A$ is the indicator function of A . From the conditional probability we could furthermore define conditional distribution functions and conditional probability densities. In addition we can consider the conditional variance of a random variable:

$$\text{Var}\{X | \mathcal{H}\} \equiv \mathbb{E}\{X^2 | \mathcal{H}\} - \mathbb{E}\{X | \mathcal{H}\}^2. \quad (\text{A.16})$$

There exist a useful decomposition formula for the variance known as the law of total variance [36]

$$\mathbb{V}\{X\} = \mathbb{E}\{\mathbb{V}\{X | \mathcal{H}\}\} + \mathbb{V}\{\mathbb{E}\{X | \mathcal{H}\}\}. \quad (\text{A.17})$$

This decomposition formula can be given a natural interpretation in terms of estimators of random variables and estimation error.

Appendix B

Exploiting the Causal Tensor Network Structure of Quantum Processes to Efficiently Simulate Non-Markovian Path Integrals

This chapter is composed entirely of text and figures from Ref. [127] © 2019 American Physical Society. The chapter is self-contained, in particular the notation does not conform to the main chapters of the thesis.

B.1 Introduction

All nanoscale quantum systems are open, meaning they inevitably interact with their environments, exchanging energy and generating correlations. If the system and its environment remain approximately uncorrelated, then the reduced system dynamics is well described by a Markovian model [37, 152, 38]. However, in physical systems such as photosynthetic complexes, nanoscale lasers and quantum thermal machines [205, 206, 207], the need to go beyond a Markovian description has long been recognized, and techniques accounting for non-Markovian physics have been developed, with greater or lesser breadth of applicability. Analytical methods involving time-local equations of motion exist, but tend to be highly restricted to specific parameter regimes [208, 209, 210]. Exact simulation often requires numerical methods, e.g. discrete path integrals [211, 212, 213, 214, 215], time non-local memory kernels [216, 217, 218, 219, 220, 221, 222], hierarchical equations of motion [223, 224] and others [225, 226, 227, 218]. Overall these methods tend to scale unfavourably with both the simulation time and the system size [228], making them inapplicable to important processes involving large complexes or when long time dynamics is important.

Recently, tensor network methods have been applied to the simulation [229, 230, 231, 232] and characterization [233, 234] of open quantum dynamics. Physically, these methods incorporate the fact that typical open quantum systems are only finitely correlated with their environments, massively reducing their description [235]. In particular, Strathearn et al. [154] reformulated the discrete path integral for open systems with Gaussian environments in terms of matrix product operators; the resulting time-evolving matrix product operator (TEMPO) algorithm is numerically exact and has an efficiency com-

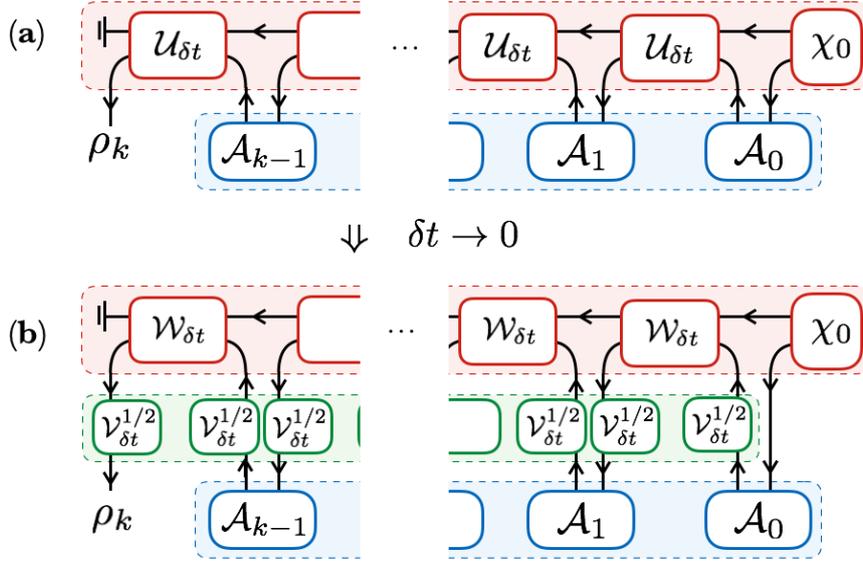


Figure B.1: (Color online) (a) An arbitrary process with interventions can be represented as a matrix product form tensor network, the process tensor (upper row), that contracts with a filter function consisting of a sequence of superoperators (lower row). This makes it possible to separate implemented control operations from the underlying uncontrolled process. (b) In the infinitesimal time step limit, the uncontrolled process can be further decomposed into free evolution of the system (middle row) and a generalized influence functional capturing the influence of the environment.

parable to other state of the art methods. By effectively only considering the most important non-Markovian contributions to the dynamics, the algorithm circumvents the exponential memory scaling of the bare path integral representation, in a similar spirit to earlier path-filtering techniques [236, 237, 238]. Motivated by this success, it is natural to ask if tensor network methods can be efficiently generalized from the simulation of reduced system density operators, to the simulation of general non-Markovian processes and multi-time correlations, which typically require many realizations of the dynamics to characterize.

In this Letter, we propose such a generalization, by making a formal connection between the path integral structure and the recently developed process tensor framework for characterizing general non-Markovian quantum processes [233]. We then use this to argue for an alternative formulation of the TEMPO algorithm, where we exploit the symmetry of the underlying tensor network to better account for the causal structure inherent in the dynamics. This not only allows us to efficiently compute multi-time correlation functions – the simulation need only run once to extract all multi-time observable properties – but also opens the door to the simulation of more general models. Our alternative formulation is demonstrated to significantly improve the efficiency of the method, which we use to straightforwardly compute non-Markovian emission spectra for the spin-boson model, beyond the point where the commonly used quantum regression theorem breaks down [239].

B.2 Process tensor framework

We consider stationary unitary dynamics of the system S we are interested in along with its environment E , and suppose the system is transformed by superoperators \mathcal{A}_j at the discrete times $\{t_{k-1}, \dots, t_0\}$, which we take to be at evenly spaced intervals of $\delta t = t_j - t_{j-1}$. In an experiment, these superoperators could correspond to actual interventions on the system as it evolves, i.e. unitary rotations, measurements etc., in which case they are completely-positive and, if the interventions are not conditional on a particular measurement outcome, trace preserving. Otherwise, the set $\{\mathcal{A}_j\}$ could represent more abstract transformations useful in the computation of physical quantities such as operator expectation values or emission spectra. The reduced, and potentially subnormalized, state of the system at time t_k is given by

$$\rho_k(\{\mathcal{A}_j\}) = \text{Tr}_E \{ \mathcal{U}_{\delta t} \mathcal{A}_{k-1} \dots \mathcal{U}_{\delta t} \mathcal{A}_0 [\chi_0] \} , \quad (\text{B.1})$$

where $\mathcal{U}_{\delta t}$ is a superoperator representation of the unitary evolution of duration δt , i.e. $\mathcal{U}_{\delta t}[\rho] = U_{\delta t} \rho U_{\delta t}^\dagger$, with $U_{\delta t}$ a unitary matrix, and χ_0 is the initial system-environment state. The inclusion of intermediate transformations makes it possible for us to consider a much broader class of physical properties than free evolution of the density operator (corresponding to $\mathcal{A}_j = \mathcal{I}$ the identity superoperator $\forall j$) would allow.

Since the state at time t_k in Eq. (B.1) is linearly related to each of the set of superoperators $\{\mathcal{A}_j\}$ it can be written as a linear function of the tensor product of their Choi state representations $\mathbf{A}_{k-1:0} = \mathbf{A}_{k-1} \otimes \dots \otimes \mathbf{A}_1 \otimes \mathbf{A}_0$, with $\mathbf{A}_j := \sum_{sr} \mathcal{A}_j[|s\rangle\langle r|] \otimes |s\rangle\langle r|$ obtained via the Choi-Jamiołkowski isomorphism [240, 34, 150]; here, $\{|s\rangle\}$ forms an orthonormal basis for S . Specifically, $\rho_k(\{\mathcal{A}_j\}) = \text{tr}_{k-1:0} \{ \Upsilon_{k:0} (\mathbb{1}_k \otimes \mathbf{A}_{k-1:0}^T) \}$, with the trace over all subsystems on which $\mathbf{A}_{k-1:0}$ acts. As we detail explicitly in Appendix B.7,

$$\begin{aligned} \Upsilon_{k:0} = & \sum_{\vec{s}', \vec{r}', \vec{s}, \vec{r}} \text{tr} \left\{ \mathcal{U}_{\delta t}^{(s'_k, r'_k, s_{k-1}, r_{k-1})} \dots \mathcal{U}_{\delta t}^{(s'_1, r'_1, s_0, r_0)} \left[\chi_0^{(r'_0, s'_0)} \right] \right\} \\ & \times |s'_k s_{k-1} \dots s'_1 s_0 s'_0\rangle \langle r'_k r_{k-1} \dots r'_1 r_0 r'_0| , \end{aligned} \quad (\text{B.2})$$

with environment superoperators $\mathcal{U}_{\delta t}^{(s', r', s, r)}[\rho^E] = \langle s' | U_{\delta t} (|s\rangle\langle r| \otimes \rho^E) U_{\delta t}^\dagger |r'\rangle$ and operators $\chi_0^{(r', s')} = \langle r' | \chi_0 |s'\rangle$, is the Choi representation of the *process tensor* [233], a many-body operator (on $2k+1$ copies of S) containing all information about the system's evolution that is independent of the transformations $\{\mathcal{A}_j\}$. Correlations between subsystems of $\Upsilon_{k:0}$ correspond to temporal correlations between observables, and a representation in terms of process tensors has been shown to consistently generalize stochastic processes, and related notions such as the Markov property and Markov order, to the quantum case [241, 242, 243, 244]. The process tensor is illustrated graphically in Fig. B.1, and can be thought of simply as a sequence of correlated maps on the system [245]. Unlike in a conventional open quantum systems picture, where density operators are mapped to density operators, this operational formulation stresses that the proper input to a quantum process is the set of interventions $\mathbf{A}_{k-1:0}$, and that the intermediate dynamics, and the initial state, are features of the process itself.

B.3 Gaussian influence functional

Here, we consider the specific structure of the process tensor for systems interacting with Gaussian environments, where the system-environment Hamiltonian and initial

state depend at most quadratically on environment creation and annihilation operators. For concreteness, we focus on spin-boson type models, but our results would extend to fermionic baths as well [30]. Working in natural units ($\hbar = k_B = 1$), we consider a spin system, with Hilbert space dimension d , interacting linearly with a bath of harmonic oscillators described by the Hamiltonian $H = H_0 + H_B$. Here, H_0 describes the free spin system and the full bath influence is collected in $H_B = \hat{s} \sum_n (g_n \hat{a}_n + g_n^* \hat{a}_n^\dagger) + \sum_n \omega_n \hat{a}_n^\dagger \hat{a}_n$. A bath mode n has energy ω_n , and is created (annihilated) by the bosonic operator \hat{a}_n^\dagger (\hat{a}_n). The system operator \hat{s} interacts with the bath with coupling constants g_n . Additional linear interaction terms to different system operators could be included, as long as all these system operators commute. For simplicity, we take the initial state to be product, such that $\chi_0 = \rho_0 \otimes \tau_\beta$, with the environment initially described by a thermal state $\tau_\beta = \exp[-\beta \sum_n \omega_n \hat{a}_n^\dagger \hat{a}_n] / \mathcal{Z}$ at inverse temperature β , where $\mathcal{Z} = \text{tr}\{\exp[-\beta \sum_n \omega_n \hat{a}_n^\dagger \hat{a}_n]\}$. This choice is not essential, and other, possibly correlated, Gaussian initial states of the environment could be considered.

In the limit that the time difference δt is small, the generated unitary dynamics can be approximately separated into contributions arising from H_0 and H_B as $\mathcal{U}_{\delta t} \simeq \mathcal{V}_{\delta t}^{1/2} \mathcal{W}_{\delta t} \mathcal{V}_{\delta t}^{1/2}$, where $\mathcal{V}_{\delta t}$ describes the free dynamics of S and $\mathcal{W}_{\delta t}$ describes the environment influence. The discrepancy between the approximate unitary maps and the actual ones vanishes as $\mathcal{O}(\delta t^3)$ for this symmetric decomposition [246]. Note that the approximate model does not break unitarity, and so corresponds to a valid physical process independently of step size. Moreover, since the Hamiltonian only contains a single interaction term, the interaction unitary preserves the eigenbasis of the corresponding system operator $\hat{s} = \sum_s \lambda_s |s\rangle\langle s|$: $\langle s' | \mathcal{W}_{\delta t} [|s\rangle\langle r |] |r'\rangle = \delta_{ss'} \delta_{rr'} \mathcal{W}_{\delta t}^{(s,r)}$. Together with the decomposition of unitary maps, this allows us to write the approximate process tensor Choi state as

$$\tilde{\Upsilon}_{k:0} = \left(\mathcal{V}_{\delta t}^{1/2} \otimes \mathcal{V}_{\delta t}^{*1/2} \right)^{\otimes k} [\mathcal{F}_{k:0}] \otimes \rho_0, \quad (\text{B.3})$$

where $\mathcal{F}_{k:0}$ is an operator representation of the discretized Feynman-Vernon influence functional [247] encoding environment induced correlations

$$\begin{aligned} \mathcal{F}_{k:0} = \sum_{\vec{s}, \vec{r}} \text{tr}_E \left\{ \mathcal{W}_{\delta t}^{(s_k, r_k)} \dots \mathcal{W}_{\delta t}^{(s_1, r_1)} [\tau_\beta] \right\} \\ \times |s_k s_k \dots s_1 s_1\rangle\langle r_k r_k \dots r_1 r_1|. \end{aligned} \quad (\text{B.4})$$

For Gaussian environments, the bath degrees of freedom can be traced over analytically using standard path integral techniques [211, 212, 248, 215]. In this case, introducing the d^2 compound indices $\alpha = (s, r)$, an element of the influence functional $\mathcal{F}_{k:0}^{\alpha_k \dots \alpha_1} := \langle s_k s_k \dots s_1 s_1 | \mathcal{F}_{k:0} | r_k r_k \dots r_1 r_1 \rangle$ can be decomposed as

$$\mathcal{F}_{k:0}^{\alpha_k \dots \alpha_1} = \prod_{i=1}^k \prod_{j=1}^i [b_{(i-j)}]^{\alpha_i \alpha_j}, \quad (\text{B.5})$$

where $b_{(i-j)}$ is called an influence tensor; the exact form, which can often be approximated by an analytic function [249], is given in Appendices B.7 and B.7 along with a full derivation of Eqs. (B.4) and (B.5). The influence tensors connect the dynamics around time step i with that around step j , quantifying the temporal correlations mediated by the environment between those two points; that is, they describe memory effects. Since the Hamiltonian is time-independent, the individual tensors $[b_i]$ depend only on the temporal separation $l\delta t$, simplifying the potential complexity considerably. However, since

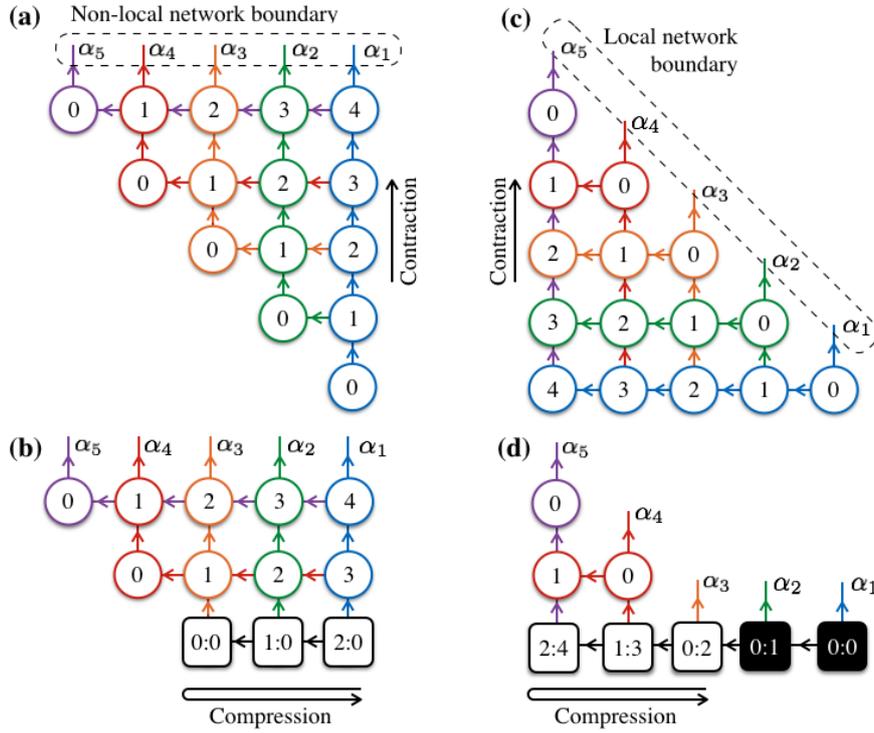


Figure B.2: (Color online) Tensor network representation of the influence functional on five time steps, with nodes representing influence tensors and labelled by time step separation. Before contraction, indices are constrained to be equal along rows and columns in the network; the open boundary can therefore be shifted to any tensor in the same column (panels (a) and (b)) or row (panels (c) and (d)). (a) With the non-local boundary choice of Ref. [154], the free indices are attached to influence tensors encoding memory effects over all different time-scales. (b) The full network is contracted iteratively from below, row by row, down to a boundary MPO, with the full influence functional changing at each step. (c) With the local boundary choice, the free indices are always attached to the time-local influence tensors. (d) Contraction proceeds as indicated, with the causal influence of each open leg sequentially incorporated into the wider network. The influence functional on an open leg is fixed once the corresponding layer has been contracted over.

the influence functional is a k index tensor, it is still potentially exponentially complex; we now show how viewing Eq. (B.5) as a tensor network can make its calculation more tractable.

B.4 Tensor network simulation

In many cases, the environment interaction produces only finite length correlations in $\mathcal{F}_{k:0}$, a fact used by the authors of Ref. [154] to circumvent the exponential complexity growth by representing it efficiently in terms of matrix product operators (MPOs) [250]. To introduce this representation we first extend our two-index influence tensors into three-index tensors as $[b_{(i-j)}]^{\gamma\alpha_i}_{\alpha_j} := \delta_{\alpha_j}^{\gamma} [b_{(i-j)}]^{\alpha_i}_{\alpha_j}$, where by convention an upper and a lower repeated index in a product of tensors is summed over (otherwise, tensor

elements differing only through raising or lowering are treated as equal). In terms of these, we then define the *non-local* time-evolving MPOs

$$\begin{aligned}
 \mathcal{F}_{k:0}^{\alpha_k \dots \alpha_1} &= \prod_{i=1}^k [b_0]_{\beta_1}^{\alpha_i} \prod_{j=1}^{i-2} [b_j]^{\beta_j \alpha_{i-j}} [b_{i-1}]^{\beta_{i-1} \alpha_1} \\
 &= \prod_{i=1}^k \text{Diagram}_i,
 \end{aligned} \tag{B.6}$$

where the outgoing (ingoing) arrows in the graphical representation indicate upper (lower) indices, and lines running through a given row or column are fixed to have the same index through Kronecker deltas. At the right boundary of the tensors shown in the second line of Eq. (B.6), we end up with a redundant lower index which we can trace over, while at the left boundary we impose that the two upper indices be equal. The full influence functional can then be constructed by iteratively multiplying such MPOs. If we label the individual MPOs in the product by $G^{\alpha_i \dots \alpha_1}$, then we can express the iterative multiplication as

$$\mathcal{F}_{k:0}^{\alpha_k \dots \alpha_1} = \mathcal{G}_{\beta_{k-1} \dots \beta_1}^{\alpha_k \alpha_{k-1} \dots \alpha_1} \mathcal{F}_{k-1:0}^{\beta_{k-1} \dots \beta_1}, \tag{B.7}$$

with $\mathcal{G}_{\beta_{k-1} \dots \beta_1}^{\alpha_k \alpha_{k-1} \dots \alpha_1} := G^{\alpha_k \dots \alpha_1} \delta_{\beta_{k-1}}^{\alpha_{k-1}} \dots \delta_{\beta_1}^{\alpha_1}$; this is represented graphically by the two-dimensional tensor network shown in Fig. B.2a. Conceptually the use of time-evolving MPOs, allows the state of the system to be propagated by updating indices to encode memory effects from the past process. This type of propagation is analogous to a description in terms of a time non-local memory kernel, since the open legs are connected to tensors describing the influence of the state at various points in its history ¹.

A key insight of this paper is that the decomposition of the influence functional into MPOs is not unique. Kronecker deltas implicit in Eqs. (B.6) and (B.7) mean that the open leg in a given row or column in Fig. B.2 could be shifted to any tensor in that same row or column. In particular, the causal structure of the process tensor motivates an alternative definition in terms of *local* time-evolving MPOs

$$\begin{aligned}
 \mathcal{F}_{k:0}^{\alpha_k \dots \alpha_1} &= \prod_{i=1}^k [b_{k-i}]_{\beta_{k-i}}^{\alpha_k} \prod_{j=1}^{k-i-1} [b_j]^{\beta_{j+1} \alpha_{j+1}} [b_0]^{\beta_1 \alpha_i} \\
 &= \prod_{i=1}^k \text{Diagram}_i,
 \end{aligned} \tag{B.8}$$

where now we end up with a redundant index at the left boundary, which we trace over, and at the right boundary we impose the condition that the lower index must equal α_i . As with the non-local propagators, the full influence functional is constructed iteratively by locally contracting MPOs. Labelling the individual MPOs in the product by $C^{\alpha_k \dots \alpha_i}$, the iterative multiplication can be expressed as (for $i \geq 1$)

$$\tilde{\mathcal{F}}_{(k:i+1)}^{\alpha_k \dots \alpha_1} := C_{\beta_k \dots \beta_{i+1}}^{\alpha_k \dots \alpha_{i+1}} \tilde{\mathcal{F}}_{(k:i)}^{\beta_k \dots \beta_{i+1} \alpha_i \dots \alpha_1}, \tag{B.9}$$

¹This correspondence is not precise, however, and our usage of ‘(non-)local’ should not be confused with that in the context of memory kernel convolution.

with $\mathcal{C}_{\gamma_k \dots \gamma_i}^{\alpha_k \dots \alpha_i} := C^{\alpha_k \dots \alpha_i} \delta_{\gamma_k}^{\alpha_k} \dots \delta_{\gamma_i}^{\alpha_i}$ and $\tilde{\mathcal{F}}_{(k:1)}^{\alpha_k \dots \alpha_1} := C^{\alpha_k \dots \alpha_1}$. The resulting network representation for the influence functional $\mathcal{F}_{k:0} = \tilde{\mathcal{F}}_{(k:k)}$ is shown in Fig. B.2c. Conceptually, the local time-evolving MPOs propagate the state by updating a set of effective memory space indices. These indices describe how the environment is conditioned by the process at a given time, and this information on the conditioning can be propagated locally. Since the process tensor, and hence the influence functional, has a well-defined causal structure, this conditioning only occurs from the past to the future. This means that, for a fixed evolution time, the size of the tensor to be updated decreases with each iteration.

In contracting the network, efficiency is achieved by incorporating a tensor compression procedure of the obtained boundary in each iteration. In this work we make use of the singular value compression procedure (see Appendix B.7) [250, 154]. Roughly speaking, the local tensors of the boundary are subjected to a singular value decomposition. The compression consists of discarding the eigenvalues below a specified singular value cutoff λ_c , quantifying the hardness of the compression. For the non-local algorithm (Fig. B.2a,b), the tensors contracted in each iteration encode information about the influence of multiple time-steps on each other. When correlations become smaller at longer time scales, as is typically the case, not all this information is relevant for describing the process as a whole. The local algorithm (Fig. B.2c,d) incorporates this insight, and separates out the most important contribution by including only the future influence of the environment at each timestep. Generally the most local contributions have the largest singular values, and therefore the separation means that the part of the boundary being propagated in the local case is less correlated, which translates into a more efficient algorithm.

B.5 Network complexity for a two level system

We now turn to the specific simulation of the dynamics of a two-level system, and compare the performance of the non-local and local algorithms. Consider the free Hamiltonian $H_0 = \Omega \sigma_x / 2$ and $\hat{s} = \sigma_z / 2$, where σ_x and σ_z are the usual Pauli operators. The environment is fully characterized by its spectral density defined as $J(\omega) = \sum_n |g_n|^2 \delta(\omega - \omega_n)$ [37]. Here we consider a continuum bath model with the spectral density

$$J(\omega) = (\alpha \omega_c / 2) (\omega / \omega_c)^\nu \exp(-\omega / \omega_c), \quad (\text{B.10})$$

with coupling strength α , cutoff frequency ω_c and Ohmicity ν , where for an Ohmic spectral density $\nu = 1$.

The computational complexity is quantified by the computation time required to contract a network of a certain size with a fixed singular value cutoff (see Appendix B.7 for details on implementation). In general, this will depend on the overall magnitude of the influence functional, as well as the characteristic memory time quantifying how the elements of the influence tensors $b_{(i-j)}$ decrease in magnitude at large $|i - j|$. In Appendix B.7, we show that, for a fixed evolution time, the memory time goes as $\alpha / (\beta \omega_c)$ when ω_c is large, and that the overall coupling goes as $\alpha \omega_c t_{\max}^2 / \beta$ when ω_c is small. In Fig. B.3a, we plot the computation time for the local and non-local algorithms as a function of coupling strength. We find that the local representation outperforms the non-local one by one-to-two orders of magnitude, even at weak coupling, and that

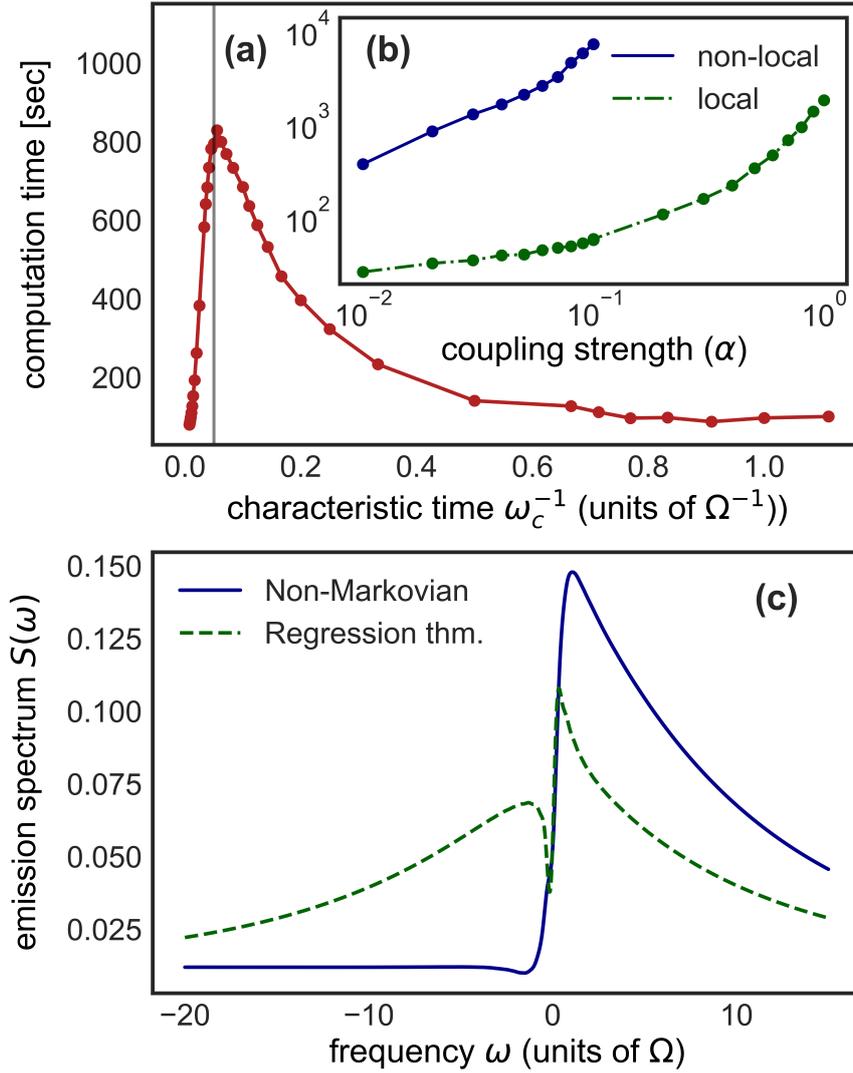


Figure B.3: (a) Variation of computation time with the inverse of the cutoff frequency for the local algorithm at a coupling strength of $\alpha = 0.7$, for an ohmic spectral density with $\omega_c = 10\Omega$, $T = 0.01\Omega$ and $\lambda_c = 10^{-6}$. (b) Comparison between the computation time of the non-local [Eq. (B.6)] and local [Eq. (B.8)] time-evolving MPO algorithms for an Ohmic spectral density with $\omega_c = 10\Omega$, $T = 0.01\Omega$ and $\lambda_c = 10^{-6}$ as a function of coupling strength. (c) Steady state phonon emission spectrum at $\alpha = 0.7$, the non-Markovian (numerically converged) spectrum is simulated using the local algorithm, and is compared with the spectrum obtained using the regression theorem (removing correlations across non-trivial superoperators).

the improvement increases at larger coupling strengths (in Appendix B.7 we show that there is an advantage everywhere across a wide range of parameters). Furthermore, in Fig. B.3b, we illustrate that this computational efficiency is maintained as the characteristic timescale of the bath (the inverse of the environment cutoff frequency) is varied, consistent with our predictions.

It should be kept in mind that, unlike in most other quantum simulation methods, including the original TEMPO algorithm, the object we are computing is the full process

tensor, which encodes all multi-time correlations, and from which a host of properties can be extracted efficiently. In particular, we can compute the steady state emission spectrum $S(\Delta\omega) = \text{Re} [\int_0^\infty d\tau (g^{(1)}(\tau) - g^{(1)}(\infty))e^{-i\Delta\omega\tau}]$, defined in terms of the two-point correlation function $g^{(1)}(\tau) = \lim_{t \rightarrow \infty} \langle \sigma^\dagger(t + \tau)\sigma(t) \rangle$. The two-point correlation function is defined in terms of the raising and lowering operators on the spin system. To compute it, we take all superoperators which the process tensor acts on to be the identity superoperator \mathcal{I} (with action $\mathcal{I}[\rho] = \rho$) except for two, which append a raising or lowering operator respectively. In Fig. B.3c, we study the physical effects of system-environment correlations by looking at the phonon emission spectrum. We compare this with the spectrum computed using the quantum regression theorem, which approximates intermediate dynamics with that from an initial product state and is valid in the weak-coupling limit [206]. The regression theorem correlations are obtained by breaking all correlations in the full process tensor Choi state across time steps at which the raising and lowering operators are evaluated. In addition, we compare the exact spectrum with a fully Markovian process in which correlations are broken after each time step. Fig. B.3c shows that non-Markovian effects produce a phonon sideband in the spectrum at positive frequencies, this is contrasted with the regression theorem result which gives a symmetric sideband structure. Furthermore the fully Markovian spectrum contains no phonon sidebands, but rather a resonant emission peak. These significant differences illustrates the importance of accounting for non-Markovian physics.

B.6 Conclusion

In this Letter, we have established a direct connection between recent frameworks for characterizing general non-Markovian quantum processes and the path integral formulation of open quantum dynamics. By relating the influence functional to a process tensor on an infinitesimal time grid, which has an explicit causal structure, we were able to build on recent progress in the classical simulation of Gaussian open quantum systems in terms of tensor networks. Specifically, we showed that the speed of the TEMPO algorithm, when computing the multi-time properties encapsulated in the process tensor, can be improved by orders-of-magnitude by shifting the boundary of the corresponding tensor network from a temporally non-local to a local one. Our contribution is immediately applicable to the efficient simulation of realistic complex open systems, and additionally illustrates the utility of thinking about a time-local propagation of a conditioned environment space, rather than a description resembling the use of a non-local memory kernel.

The utility of the non-local TEMPO algorithm has been amply illustrated by computing the Ohmic localization transition, and the dynamics of complex problems with multiple separated timescales [154]. The improved algorithm presented here is capable of exploring the same physics more efficiently, and, in addition, easily extends to the computation of multi-time observables, of the sort crucial to describing, for example, multi-dimensional spectroscopy experiments [251]. Moreover, relating path integral techniques to the more general process tensor formalism indicates how they might be generalized to more complex system-environment interactions, or even beyond the Gaussian regime. Even within the spin-boson model, the freedom of boundary choice in Fig. B.2 that we have identified could be further exploited in other contexts. While the local choice appears optimal here, it may be that for other, structured spectral densities,

where there are recurrent correlations, different boundary choices are more efficient, a point whose exploration we leave for future work.

B.7 Appendix

Process tensor formalism.

As described in the main text, we consider the scenario characterized by Eq. (B.1), where an open quantum system S is periodically interrogated as it evolves. The slightly more general case involves an arbitrary time separation between interventions and a possibly time-dependent SE Hamiltonian $H(t)$, such that the state at the k^{th} time is given by

$$\rho_k(\{\mathcal{A}_j\}_{j=0}^{k-1}) = \text{tr}_E \{ \mathcal{U}_{k:k-1} \mathcal{A}_{k-1} \dots \mathcal{U}_{1:0} \mathcal{A}_0 [\chi_0] \} \quad (\text{B.11})$$

where $\mathcal{U}_{j:j-1}$ is the time-evolution superoperator from time t_{j-1} to t_j with action $\mathcal{U}_{j:j-1}[\rho] = U_{j:j-1} \rho U_{j:j-1}^\dagger$ in terms of the time-ordered exponential

$$U_{j:j-1} = T_{\leftarrow} \exp \left\{ -i \int_{t_{j-1}}^{t_j} ds H(s) \right\}. \quad (\text{B.12})$$

As in the main text, χ_0 is the potentially correlated initial SE state and each \mathcal{A}_j can be any superoperator, though only those that are completely positive correspond to physically realizable transformations [240].

Any superoperator can be represented in terms of an operator sum as $\mathcal{A}[\rho] = \sum_n X_n \rho Y_n^\dagger$ (completely positive maps are characterized by $X_n = Y_n$, in which case the latter are called Kraus operators). As such, by taking the trace over the final state in Eq. (B.11), any multi-time correlation function $\langle B_0(t_0) \dots B_{k-1}(t_{k-1}) \mathcal{A}_{k-1}(t_{k-1}) \dots \mathcal{A}_0(t_0) \rangle_{\chi_0}$, where $\{A_j(t_j) = U_{j:0}^\dagger A_j U_{j:0}\}$ and $\{B_j = U_{j:0}^\dagger B_j U_{j:0}\}$ are Heisenberg picture operators on S , can be obtained by choosing $\mathcal{A}_j[\rho] = A_j \rho B_j$. In this way, with a sufficiently dense set of times $\{t_j\}$, any dynamically observable property of S can be represented in the form of Eq. (B.11) (correlation functions involving fewer observables can be obtained by choosing some of the $\{A_j\}$ and $\{B_j\}$ to be the identity operator). In addition, the freely evolved final state ρ_k can be obtained by choosing all $\mathcal{A}_j = \mathcal{I}$, where the latter is the identity superoperator with action $\mathcal{I}[\rho] = \rho$.

As we will now show, all this information can be encoded in a single object, the process tensor. While it is often introduced in terms of an abstract multi-linear map, we will here express it solely in terms of a concrete matrix representation via a version of the Choi-Jamiołkowski isomorphism. The Choi state (or Choi matrix) \mathbf{A} of a superoperator \mathcal{A} is defined in terms of its action on one half of the (unnormalized) maximally entangled state $\Psi = \sum_{sr} |s\rangle\langle sr|$, where $\{|s\rangle\}$ forms an orthonormal basis for the d -dimensional system, as

$$\mathbf{A} = \mathcal{A} \otimes \mathcal{I}[\Psi] = \sum_{sr s' r'} \mathcal{A}^{(s,r,s',r')} |s s'\rangle\langle r r'|, \quad (\text{B.13})$$

with $\mathcal{A}^{(s,r,s',r')} = \langle s | \mathcal{A}[|s'\rangle\langle r'|] |r\rangle$. Labelling the first and second copies of the original system's Hilbert space \mathfrak{o} (for output) and \mathfrak{i} (for input) respectively, the action of the superoperator on an initial state ρ can be written in terms of this representation as $\mathcal{A}[\rho] = \text{tr}_{\mathfrak{i}}[\mathbf{A} \mathbb{1}_{\mathfrak{o}} \otimes \rho^T]$, where the trace is over the input subsystem.

By expanding out the action of the superoperators in Eq. (B.11) and inserting a resolution of the identity on S to the left and right of every unitary matrix, one arrives at the equivalent expression

$$\rho_k(\{\mathcal{A}_j\}_{j=0}^{k-1}) = \text{tr}_{k-1:0} \left\{ \Upsilon_{k:0} (\mathbb{1}_k \otimes \mathbf{A}_{k-1:0}^T) \right\}, \quad (\text{B.14})$$

with $\mathbf{A}_{k-1:0} = \mathbf{A}_{k-1} \otimes \cdots \otimes \mathbf{A}_1 \otimes \mathbf{A}_0$ and the trace over all (input and output) Hilbert spaces on which $\mathbf{A}_{k-1:0}$ acts. The positive operator $\Upsilon_{k:0}$ is (the Choi state of) the process tensor, and it can be expressed in terms of the underlying SE dynamics as

$$\begin{aligned} \Upsilon_{k:0} = \sum_{\vec{s}', \vec{r}', \vec{s}, \vec{r}} \text{tr} \left\{ \mathcal{U}_{k:k-1}^{(s'_k, r'_k, s_{k-1}, r_{k-1})} \cdots \mathcal{U}_{1:0}^{(s'_1, r'_1, s_0, r_0)} \left[\chi_0^{(r'_0, s'_0)} \right] \right\} \\ \times |s'_k s_{k-1} \cdots s'_1 s_0 s'_0\rangle \langle r'_k r_{k-1} \cdots r'_1 r_0 r'_0|, \end{aligned} \quad (\text{B.15})$$

with $\mathcal{U}_{j:j-1}^{(s', r', s, r)}[\rho^E] = \langle s' | U_{j:j-1} (|s\rangle\langle r| \otimes \rho^E) U_{j:j-1}^\dagger |r'\rangle$ and $\chi_0^{(r', s')} = \langle r' | \chi_0 | s' \rangle$. This object can be directly constructed by swapping the system with one half of a maximally entangled state at each point where a superoperator \mathcal{A} is to be applied [233]. When the SE Hamiltonian is time independent and $t_j - t_{j-1} = \delta t$ for all j , Eq. (B.15) is equivalent to Eq. (B.2) of the main text.

A representation in terms of the process tensor separates the process and external interventions, as illustrated in Fig. B.1 of the main text. The process tensor's properties reflect the necessary features of any physical open dynamics: $\Upsilon_{k:0}$ is positive if and only if the process is completely positive, and causality is encoded in the hierarchy of trace conditions $\text{tr}_j \Upsilon_{j:0} = \Upsilon_{j-1:0} \otimes \mathbb{1}_{\mathfrak{o}_{j-1}}$. While we have expressed it in terms of SE quantities, it is an operator only on copies of the Hilbert space of S and it has a natural matrix product form, allowing for an efficient representation in many cases, a fact we exploit in this paper.

Connection with influence functional

In the case where the time spacing is small (and the Hamiltonian varies relatively slowly), the Trotter formula can be used to approximate the time evolution superoperators as $\mathcal{U}_{j:j-1} \simeq \mathcal{V}_{j:j-1}^{1/2} \mathcal{W}_{j:j-1} \mathcal{V}_{j:j-1}^{1/2}$, with $\mathcal{V}_{j:j-1}$ generated by the S part of the Hamiltonian and $\mathcal{W}_{j:j-1}$ by the remainder. Expanding out the superoperators appearing inside the trace in Eq. (B.15) and introducing further resolutions of the identity, one finds

$$\begin{aligned} \mathcal{U}_{j:j-1}^{(s', r', s, r)} &\simeq \langle s' | \mathcal{V}_{j:j-1}^{1/2} \mathcal{W}_{j:j-1} \mathcal{V}_{j:j-1}^{1/2} [|s\rangle\langle r|] |r'\rangle \\ &= \sum_{t', t, u', u} \langle s' | \mathcal{V}_{j:j-1}^{1/2} [|t'\rangle\langle u'|] |r'\rangle \\ &\quad \times \langle s | \mathcal{V}_{j:j-1}^{*1/2} [|t\rangle\langle u|] |r\rangle \mathcal{W}_{j:j-1}^{(t', u', t, u)}, \end{aligned} \quad (\text{B.16})$$

where $\mathcal{W}_{j:j-1}^{(s', r', s, r)} := \langle s' | \mathcal{W}_{j:j-1} [|s\rangle\langle r|] |r'\rangle$, and we have used that $\langle t | \mathcal{V}_{j:j-1}^{1/2} [|s\rangle\langle r|] |u\rangle = \langle s | \mathcal{V}_{j:j-1}^{*1/2} [|t\rangle\langle u|] |r\rangle$. Assuming a factorizing initial condition $\chi_0 = \rho_0 \otimes \tau$ and substituting Eq. (B.16) into Eq. (B.15) leads to the following slightly more general version of Eq. (B.5) for the approximate process tensor $\tilde{\Upsilon}_{k:0}$:

$$\tilde{\Upsilon}_{k:0} = \bigotimes_{j=1}^k \left(\mathcal{V}_{j:j-1}^{1/2} \otimes \mathcal{V}_{j:j-1}^{*1/2} \right) [\mathcal{F}_{k:0}] \otimes \rho_0, \quad (\text{B.17})$$

with

$$\begin{aligned} \mathcal{F}_{k:0} &= \sum_{\vec{s}', \vec{r}', \vec{s}, \vec{r}} \text{tr}_E \left\{ \mathcal{W}_{k:k-1}^{(s'_k, r'_k, s_{k-1}, r_{k-1})} \cdots \mathcal{W}_{1:0}^{(s'_1, r'_1, s_0, r_0)} [\mathcal{T}] \right\} \\ &\quad \times |s'_k s_{k-1} \cdots s_1 s'_1 s_0\rangle \langle r'_k r_{k-1} \cdots r_1 r'_1 r_0|. \end{aligned} \quad (\text{B.18})$$

In the special case that the bath coupling part Hamiltonian can be written in the form $H_B(t) = \sum_s |s\rangle\langle s| \otimes B_s(t)$ where $\sum_s |s\rangle\langle s| = \mathbb{1}_S$ (the Hamiltonian of the main text takes this form in the interaction picture with respect to the bath), $\mathcal{W}_{j:j-1}^{(s', r', s, r)} = \delta_{ss'} \delta_{rr'} \mathcal{W}_{j:j-1}^{(s, r)}$ with

$$\mathcal{W}_{j:j-1}^{(s, r)}[\rho^E] = W_{j:j-1}^{(s)} \rho^E W_{j:j-1}^{(r)\dagger} \quad (\text{B.19})$$

and

$$W_{j:j-1}^{(s)} = T_{\leftarrow} \exp \left[-i \int_{t_{j-1}}^{t_j} dx B_s(x) \right]. \quad (\text{B.20})$$

For a time-independent Hamiltonian with even time spacing δt , Eq. (B.18) then reduces to the operator representation of the discretized Feynman-Vernon influence functional in Eq. (B.4) of the main text.

Explicit form of influence tensors in the spin-boson model

Further decomposing the influence functional into a product of the form of Eq. (B.5) requires that the bath be composed of field modes coupled linearly to the system, and that the Hamiltonian and initial state are quadratic in the corresponding creation and annihilation operators $\{\hat{a}_n^\dagger\}$ and $\{\hat{a}_n\}$. In other words, the environment must be Gaussian and the operators coupling to the system must take the form $\hat{B}_s(t) = \sum_n (g_{s,n} \hat{a}_n e^{-i\omega_n t} + g_{s,n}^* \hat{a}_n^\dagger e^{i\omega_n t})$ in the interaction picture. In this case, Wick's theorem can be applied to express Eq. (B.18) as a product of exponentiated two point correlation functions. Specifically, we use the fact that for any linear functional of bath operators \hat{X} , $\text{tr}\{T \exp[\hat{X}] \rho\} = \exp\left[\frac{1}{2} \text{tr}\{T \hat{X}^2 \rho\}\right]$, with T any time ordering operator, when ρ is Gaussian [248]. Treating the left and right appended operators in Eq. (B.19) as a single contour ordered exponential under the trace, this results in the following expression:

$$\begin{aligned} \mathcal{F}_{k:0}^{\alpha_k \cdots \alpha_1} &= \text{tr}_E \left\{ \mathcal{W}_{k:k-1}^{(s_k, r_k)} \cdots \mathcal{W}_{1:0}^{(s_1, r_1)} [\mathcal{T}] \right\} \\ &= \exp \left[-\frac{1}{2} \sum_{i \geq j} \left(\zeta_{i,j}^{(s_i, s_j)} + \zeta_{i,j}^{(r_i, r_j)*} \right. \right. \\ &\quad \left. \left. - \zeta_{i,j}^{(r_i, s_j)} - \zeta_{i,j}^{(s_i, r_j)*} \right) \right], \end{aligned} \quad (\text{B.21})$$

where

$$\zeta_{i,j}^{(u,v)} = \int_{t_{i-1}}^{t_i} dx \int_{t_{j-1}}^{t_j} dy \text{tr} \left\{ \hat{B}_u(x) \hat{B}_v(y) \tau \right\} \quad (\text{B.22})$$

for $i \neq j$, else for $i = j$:

$$\zeta_{i,i}^{(u,v)} = \int_{t_{i-1}}^{t_i} dx \int_{t_{i-1}}^x dy \operatorname{tr} \left\{ \hat{B}_u(x) \hat{B}_v(y) \tau \right\}. \quad (\text{B.23})$$

Restricting to the spin-boson type Hamiltonian considered in the main text, with a single interaction term $\hat{s} \sum_n (g_n \hat{a}_n + g_n^* \hat{a}_n^\dagger)$ and with a thermal (and hence Gaussian) initial bath state τ_β , we can compute the influence tensors explicitly. Here, the interaction picture bath operators appearing in Eqs. (B.22) and (B.23) take the simple form $\hat{B}_u(t) = \lambda_u \sum_n (g_n \hat{a}_n e^{-i\omega_n t} + g_n^* \hat{a}_n^\dagger e^{i\omega_n t})$, written in terms of the eigenvalues of the system operator $\hat{s} = \sum_u \lambda_u |u\rangle\langle u|$.

In this case, $\zeta_{i,j}^{(u,v)} = 2\lambda_u \lambda_v \eta_{i-j}$, where the memory kernel elements

$$\eta_{i-j} = \begin{cases} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} dt' dt'' C(t' - t''), & i \neq j \\ \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t'} dt' dt'' C(t' - t''), & i = j \end{cases}, \quad (\text{B.24})$$

depend only on the difference between time steps and are expressed in terms of the environment auto-correlation function [154, 215]

$$C(t) = \frac{1}{\pi} \int_0^\infty d\omega J(\omega) \frac{\cosh[\omega(\beta/2 - it)]}{\sinh[\beta\omega/2]}. \quad (\text{B.25})$$

Here $J(\omega) = \sum_n |g_n|^2 \delta(\omega - \omega_n)$ is the spectral density, as defined in the main text. In terms of these quantities, the elements of the influence tensors $b_{(i-j)}$ that correspond to each of the terms in the exponentiated sum in Eq. (B.21) (such that $\mathcal{F}_{k:0}^{\alpha_k \dots \alpha_1} = \prod_{i \geq j} [b_{(i-j)}]^{\alpha_i \alpha_j}$), can be written

$$[b_{(i-j)}]^{\alpha_i \alpha_j} = e^{-(\lambda_{s_i} - \lambda_{r_i})(\eta_{i-j} \lambda_{s_j} - \eta_{i-j}^* \lambda_{r_j})}. \quad (\text{B.26})$$

These are the values that enter directly into our algorithm.

Tensor network compression

In contracting the network, efficiency is achieved by finding a minimal approximate representation for the boundary matrix product operator in each iteration [250]. This is obtained by replacing high rank tensors with small singular values by lower rank approximations. These are found by performing a singular value decomposition (SVD) of the local tensors in the matrix product state, and discarding the singular values below a cutoff λ_c . Indicating an index partition by raised and lowered indices, the SVD decomposed tensor takes the form

$$\mathcal{F}^{\alpha_k \dots \alpha_{j+1}}_{\alpha_j \dots \alpha_i} = U^{\alpha_k \dots \alpha_{j+1}}{}_\gamma \Lambda^\gamma{}_\delta (V^\dagger)^\delta_{\alpha_j \dots \alpha_i}, \quad (\text{B.27})$$

where the diagonal matrix $\Lambda^\gamma{}_\delta$ contains the singular values, and U, V are rectangular isometric matrices satisfying $U^\dagger U = \mathbb{1}$ and $V^\dagger V = \mathbb{1}$ (see Fig. B.4a). Truncating the singular values reduces the sizes of U and V (we will refer to the truncated versions as \bar{U} and \bar{V}), and introduces a corresponding truncation error, whose magnitude is determined by the cutoff. We truncate the singular values such that

$$\lambda_c \leq \sqrt{\frac{\Lambda^2 - \tilde{\Lambda}^2}{\Lambda^2}}, \quad (\text{B.28})$$

where $\tilde{\Lambda}$ denotes the truncated diagonal matrix. Having truncated the singular values, we contract the \tilde{U} and $\tilde{\Lambda}$ to give a new matrix Q with which to express the newly compressed local tensor:

$$\mathcal{F}^{\alpha_k \dots \alpha_{j+1}}_{\alpha_j \dots \alpha_i} \simeq Q^{\alpha_k \dots \alpha_{j+1}}_{\delta} \left(\tilde{V}^\dagger \right)_{\alpha_j \dots \alpha_i}^{\delta}. \quad (\text{B.29})$$

The singular value compression proceeds from one end of the boundary matrix product operator. Say we begin the compression at the right boundary (see Fig. B.2), then initially the first (farthest to the right) local tensor is singular value decomposed and compressed. The matrix \tilde{U} in the above decomposition is then contracted with the diagonal matrix $\tilde{\Lambda}$ to give Q , which is subsequently contracted with the second local tensor to the left. The compressed \tilde{V} tensor is stored as the new first local tensor (see Fig. B.4b) and this procedure is repeated for the second local tensor, and so on until the left boundary is reached. This constitutes a left sweep of the SVD compression procedure. After this left sweep, an equivalent right sweep is performed, where the left-most tensor is singular value decomposed and compressed, followed by the next left-most and so on. To produce the figures in this paper, we implemented one left sweep and one right sweep in each stage of the algorithm. Including more sweeps back and forth would in principle improve the quality of the compression; however, as we discuss below our implementation is sufficient to demonstrate an improvement of the local over the non-local algorithm.

Scaling of memory effects with physical parameters

We now proceed to estimate the complexity of contracting the network. This depends crucially on how quickly the memory decays, and hence the effective depth of the tensor network in Fig. B.2c (as we will see, the overall size of memory effects is also important). From Eq. (B.26), it is clear that non-trivial contributions of the $b_{(i-j)}$ tensors to the influence functional depend on the magnitude of the memory kernel elements η_{i-j} . Specifically, the effect of truncating the network in Fig. B.2c at a depth m , with resulting influence functional elements $[\mathcal{F}_{k:0}^{\alpha_k \dots \alpha_1}]_m$, is to introduce a relative error:

$$\begin{aligned} \varepsilon_m &:= \frac{\mathcal{F}_{k:0}^{\alpha_k \dots \alpha_1}}{[\mathcal{F}_{k:0}^{\alpha_k \dots \alpha_1}]_m} - 1 = \prod_{i=1}^{k-m} \prod_{j=1}^i [b_{(i-j+m)}]^{\alpha_{i+m} \alpha_j} - 1 \\ &\simeq \sum_{i=1}^{k-m} \sum_{j=1}^i (\lambda_{r_{i+m}} - \lambda_{s_{i+m}}) (\eta_{i-j+m} \lambda_{s_j} - \eta_{i-j+m}^* \lambda_{r_j}) \\ &\leq 2 \|\hat{s}\|_{\text{op}} \sum_{l=m}^k (k-l) |\eta_l|, \end{aligned} \quad (\text{B.30})$$

where $\|\hat{s}\|_{\text{op}} := \max \{|\lambda_r|\}$ (we will henceforth take $\|\hat{s}\|_{\text{op}} = 1$, effectively absorbing it into the coupling strength); in the second line we have assumed m is sufficiently large that the error is small. For a fixed error ε , the memory time $t_m = m\delta t$, and hence the complexity of our algorithm (we expect the error due to SVD compression to scale similarly), will therefore depend on how quickly $|\eta_l|$ decays with l . If it decays exponentially with rate c , then in the limit of large k , it is relatively straightforward to show that the memory time scales as $t_m \sim \delta t (\log k + \log \varepsilon^{-1})/c$. However, as we will

now see, for the spectral density we have chosen, the memory kernel decays as a power law.

When $\delta t = t_j - t_{j-1}$ is sufficiently small, we have

$$\eta_{i-j} \simeq \begin{cases} \delta t^2 C((i-j)\delta t) , & i \neq j \\ \frac{1}{2} \delta t^2 C(0) , & i = j \end{cases} . \quad (\text{B.31})$$

For the spectral density introduced in the main text $J(\omega) = (\alpha\omega_c/2)(\omega/\omega_c)^\nu \exp(-\omega/\omega_c)$ with $\nu = 1$ (i.e. the Ohmic case), the integral in Eq. (B.25) can be evaluated explicitly, giving

$$C(t) = \frac{\alpha\omega_c^2}{2\pi} \left(\frac{\omega_c^2 t^2 - 1}{(\omega_c^2 t^2 + 1)^2} + \frac{2}{\beta^2 \omega_c^2} \operatorname{Re} \psi^{(1)} \left[\frac{1 - i\omega_c t}{\beta\omega_c} \right] - 2i \frac{\omega_c t}{(\omega_c^2 t^2 + 1)^2} \right) ,$$

with $\psi^{(1)}[z] := \int_0^\infty dx x e^{-zx}/(1 - e^{-x})$ the order-1 polygamma function. For large $|z|$, the latter goes as $\psi^{(1)}[z] \sim 1/z + 1/(2z^2)$ [252]. Hence, in the limit that $t \gg \omega_c^{-1}$ and $t \gg \beta$, we can expand out Eq. (B.32) and combine with Eq. (B.31) to arrive at

$$|\eta_{i-j}| = \frac{\alpha}{\pi\beta\omega_c|i-j|^2} + \mathcal{O}(|i-j|^{-4}) , \quad (\text{B.32})$$

to leading order in $|i-j|^{-1}$. Therefore, for sufficiently large m and $k = t_{\max}/\delta t$, we can perform the sum in Eq. (B.30), finding $\varepsilon_m \lesssim \alpha k \psi^{(1)}[m]/(\pi\beta\omega_c) \simeq \alpha k/(\pi\beta\omega_c m)$. For fixed error ε , we therefore have that the bound on the memory time, and hence the complexity of the algorithm scales as

$$t_m \sim \frac{\alpha t_{\max}}{\pi\beta\omega_c \varepsilon} . \quad (\text{B.33})$$

This explains the behaviour of Fig. B.3a and the large ω_c behaviour of Fig. B.3b in the main text. However, the onset of this limit depends on the parameter combinations $\omega_c t$ and $\sqrt{(t^2 + \omega_c^{-2})/\beta^2}$ both being large (these elicit expansions for the first and second terms of Eq. (B.32) respectively). When $\omega_c \lesssim t_{\max}^{-1}$, the former limit is never reached, and only the term involving the polygamma function contributes significantly to the error. Specifically, the magnitude of the memory kernel is approximately constant, going as

$$|\eta_{i-j}| = \frac{\alpha\omega_c \delta t^2}{\pi\beta} + \mathcal{O}(\omega_c^2) . \quad (\text{B.34})$$

Therefore, for small δt , Eq. (B.30) leads to $\varepsilon_m \lesssim (\alpha\omega_c/\pi\beta)(t_{\max} - t_m)^2$ and one can see that, as long as t_{\max} is fixed, the error is bounded by a number that goes to zero as ω_c does, even for very small memory times. Hence, in this limit, the effective coupling to the bath is weak overall and even the local influence tensors do not contribute significantly to the dynamics, explaining the behaviour at small ω_c in Fig. B.3b.

Scaling comparison

In Fig. B.4c we compare the computation time per iteration of the non-local and the local network representations for a fixed network size and SVD cutoff. We see that the time per iteration of the non-local TEMPO algorithm increases approximately linearly.

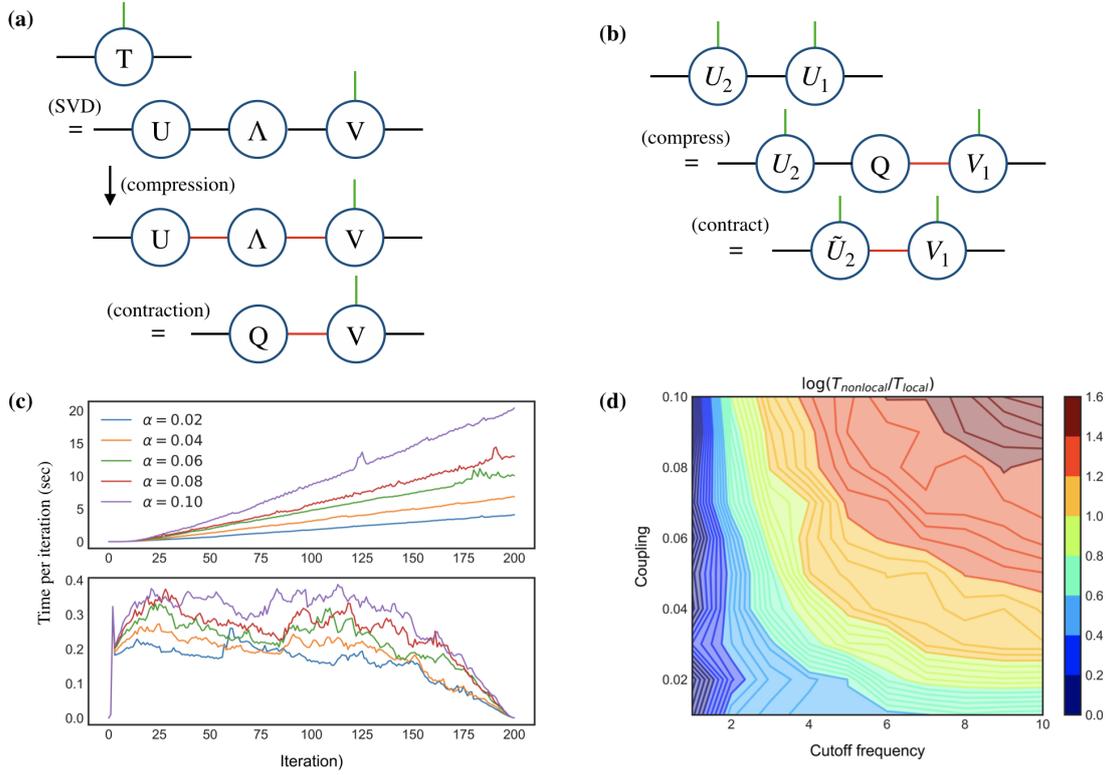


Figure B.4: (a) Singular value decomposition applied to input tensor. The singular values are truncated, giving a smaller bond dimension (indicated by red color). A new local tensor is defined and the contribution q_c is ready to be propagated along the matrix product state. (b) A local tensor of the boundary MPS is subjected to a SVD compression. The compression procedure is repeated iteratively across the full MPS. Here we illustrate a left compression sweep, the full procedure includes compression sweeps in both directions. (c) Comparison between the computation times per iteration for the local and non-local algorithms at weak coupling. The parameters used are $\omega_c = 5\Omega$, $\nu = 1$, $\delta t = 0.04/\Omega$ and a singular value cutoff of $\lambda_c = 10^{-6}$. We see that the non-local algorithm has a computation time increasing linearly with each iteration, this is contrasted with the local algorithm where an approximately constant computation time is observed. In addition we find a significant improvement in the actual value of the computation time, as argued in the main text this is due to a separation between relevant and irrelevant information in the compression. (d) Log ratio of total computation time for the non-local ($T_{nonlocal}$) and the local (T_{local}) algorithms for a range of couplings α and cutoff frequencies ω_c . These calculations were performed at zero temperature and otherwise for the same parameters as in Fig. B.3.

For a finite network, the local TEMPO algorithm has a time per iteration which rapidly goes to a non-increasing value. The growth of complexity in the non-local case is mainly due to the build up of irrelevant information, rather than a genuine build-up of temporal correlations. The observed decrease in the computation time per iteration for the local case, is a consequence of working with a finite network. In the original TEMPO proposal [154], it was argued that one could implement a truncation of the number of tensors

in the propagators and obtain a constant scaling at long times. The same method could be applied with the local TEMPO algorithm, only with a significantly reduced time per iteration. More rigorously, we could combine the tools developed here with the transfer tensor approach [218, 222], which infers long-time correlations from a short-time simulation. The problem would then become efficiently contracting the full network up to a sufficiently long-time. The advantage in contracting the network persists across a wide range of parameters, as depicted in Fig. B.4d. Even in the easier regime of weak coupling and small cutoff frequency, the non-local network takes longer to contract than the local one.

Bibliography

- [1] L Ballentine. *Quantum Mechanics: A Modern Development*. World scientific, 2014. ISBN: 978-0792336327.
- [2] Jonathan P. Dowling and Gerard J. Milburn. “Quantum technology: the second quantum revolution”. In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 361.1809 (2003), pp. 1655–1674. DOI: [10.1098/rsta.2003.1227](https://doi.org/10.1098/rsta.2003.1227).
- [3] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. “Quantum-Enhanced Measurements: Beating the Standard Quantum Limit”. In: *Science* 306.5700 (2004), pp. 1330–1336. ISSN: 0036-8075. DOI: [10.1126/science.1104149](https://doi.org/10.1126/science.1104149). eprint: <https://science.sciencemag.org/content/306/5700/1330.full.pdf>. URL: <https://science.sciencemag.org/content/306/5700/1330>.
- [4] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. “Advances in quantum metrology”. In: *Nature Photonics* 5.4 (Apr. 2011), pp. 222–229. ISSN: 1749-4893. DOI: [10.1038/nphoton.2011.35](https://doi.org/10.1038/nphoton.2011.35). URL: <https://doi.org/10.1038/nphoton.2011.35>.
- [5] B. P. Abbott et al. “Observation of Gravitational Waves from a Binary Black Hole Merger”. In: *Phys. Rev. Lett.* 116 (6 Feb. 2016), p. 061102. DOI: [10.1103/PhysRevLett.116.061102](https://doi.org/10.1103/PhysRevLett.116.061102). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.116.061102>.
- [6] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
- [7] L. D. Landau and E. M. Lifshitz. *Course of Theoretical Physics, Vol. 5: Statistical Physics, Part 1*. Pergamon Press, Oxford, 1959. ISBN: 978-0-08-057046-4. DOI: <https://doi.org/10.1016/C2009-0-24487-4>.
- [8] G. Kucsko et al. “Nanometre-scale thermometry in a living cell”. In: *Nature* 500.7460 (Aug. 2013), pp. 54–58. ISSN: 1476-4687. DOI: [10.1038/nature12373](https://doi.org/10.1038/nature12373). URL: <https://doi.org/10.1038/nature12373>.
- [9] Masazumi Fujiwara et al. “Real-time nanodiamond thermometry probing in vivo thermogenic responses”. In: *Science Advances* 6.37 (2020). DOI: [10.1126/sciadv.aba9636](https://doi.org/10.1126/sciadv.aba9636). URL: <https://advances.sciencemag.org/content/6/37/eaba9636>.
- [10] E. Moreva et al. “Practical Applications of Quantum Sensing: A Simple Method to Enhance the Sensitivity of Nitrogen-Vacancy-Based Temperature Sensors”. In: *Phys. Rev. Applied* 13 (5 May 2020), p. 054057. DOI: [10.1103/PhysRevApplied.13.054057](https://doi.org/10.1103/PhysRevApplied.13.054057). URL: <https://link.aps.org/doi/10.1103/PhysRevApplied.13.054057>.

- [11] Cécile Carcy et al. “Certifying the Adiabatic Preparation of Ultracold Lattice Bosons in the Vicinity of the Mott Transition”. In: *Phys. Rev. Lett.* 126 (4 Jan. 2021), p. 045301. DOI: [10.1103/PhysRevLett.126.045301](https://doi.org/10.1103/PhysRevLett.126.045301). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.126.045301>.
- [12] Clément De Daniloff et al. “In Situ Thermometry of Fermionic Cold-Atom Quantum Wires”. In: *Phys. Rev. Lett.* 127 (11 Sept. 2021), p. 113602. DOI: [10.1103/PhysRevLett.127.113602](https://doi.org/10.1103/PhysRevLett.127.113602). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.127.113602>.
- [13] Jean-Philippe Brantut et al. “A Thermoelectric Heat Engine with Ultracold Atoms”. In: *Science* 342.6159 (2013), pp. 713–715. ISSN: 0036-8075. DOI: [10.1126/science.1242308](https://doi.org/10.1126/science.1242308). URL: <https://science.sciencemag.org/content/342/6159/713>.
- [14] Thomas Hartke et al. “Doublon-Hole Correlations and Fluctuation Thermometry in a Fermi-Hubbard Gas”. In: *Phys. Rev. Lett.* 125 (11 Sept. 2020), p. 113601. DOI: [10.1103/PhysRevLett.125.113601](https://doi.org/10.1103/PhysRevLett.125.113601). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.125.113601>.
- [15] S. Nascimbène et al. “Exploring the thermodynamics of a universal Fermi gas”. In: *Nature* 463.7284 (Feb. 2010), pp. 1057–1060. ISSN: 1476-4687. DOI: [10.1038/nature08814](https://doi.org/10.1038/nature08814). URL: <https://doi.org/10.1038/nature08814>.
- [16] D C McKay and B DeMarco. “Cooling in strongly correlated optical lattices: prospects and challenges”. In: *Reports on Progress in Physics* 74.5 (Apr. 2011), p. 054401. DOI: [10.1088/0034-4885/74/5/054401](https://doi.org/10.1088/0034-4885/74/5/054401). URL: <https://doi.org/10.1088/0034-4885/74/5/054401>.
- [17] F. M. Spiegelhalder et al. “Collisional Stability of ^{40}K Immersed in a Strongly Interacting Fermi Gas of ^6Li ”. In: *Phys. Rev. Lett.* 103 (22 Nov. 2009), p. 223203. DOI: [10.1103/PhysRevLett.103.223203](https://doi.org/10.1103/PhysRevLett.103.223203). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.103.223203>.
- [18] Mark T. Mitchison et al. “In Situ Thermometry of a Cold Fermi Gas via Dephasing Impurities”. In: *Phys. Rev. Lett.* 125 (8 Aug. 2020), p. 080402. DOI: [10.1103/PhysRevLett.125.080402](https://doi.org/10.1103/PhysRevLett.125.080402). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.125.080402>.
- [19] Mohammad Mehboudi et al. “Using Polarons for sub-nK Quantum Nondemolition Thermometry in a Bose-Einstein Condensate”. In: *Phys. Rev. Lett.* 122 (3 Jan. 2019), p. 030403. DOI: [10.1103/PhysRevLett.122.030403](https://doi.org/10.1103/PhysRevLett.122.030403). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.122.030403>.
- [20] Quentin Bouton et al. “Single-Atom Quantum Probes for Ultracold Gases Boosted by Nonequilibrium Spin Dynamics”. In: *Phys. Rev. X* 10 (1 Jan. 2020), p. 011018. DOI: [10.1103/PhysRevX.10.011018](https://doi.org/10.1103/PhysRevX.10.011018). URL: <https://link.aps.org/doi/10.1103/PhysRevX.10.011018>.
- [21] Matthew Mecklenburg et al. “Nanoscale temperature mapping in operating microelectronic devices”. In: *Science* 347.6222 (2015), pp. 629–632. ISSN: 0036-8075. DOI: [10.1126/science.aaa2433](https://doi.org/10.1126/science.aaa2433). URL: <https://science.sciencemag.org/content/347/6222/629>.

- [22] D. Halbertal et al. “Nanoscale thermal imaging of dissipation in quantum systems”. In: *Nature* 539.7629 (Nov. 2016), pp. 407–410. ISSN: 1476-4687. DOI: [10.1038/nature19843](https://doi.org/10.1038/nature19843). URL: <https://doi.org/10.1038/nature19843>.
- [23] Bayan Karimi et al. “Reaching the ultimate energy resolution of a quantum detector”. In: *Nature Communications* 11.1 (Jan. 2020), p. 367. ISSN: 2041-1723. DOI: [10.1038/s41467-019-14247-2](https://doi.org/10.1038/s41467-019-14247-2). URL: <https://doi.org/10.1038/s41467-019-14247-2>.
- [24] Jukka P. Pekola. “Towards quantum thermodynamics in electronic circuits”. In: *Nature physics* 11 (Feb. 2015), p. 118. DOI: [10.1038/NPHYS3169](https://doi.org/10.1038/NPHYS3169).
- [25] S. Gasparinetti et al. “Fast Electron Thermometry for Ultrasensitive Calorimetric Detection”. In: *Phys. Rev. Applied* 3 (1 Jan. 2015), p. 014007. DOI: [10.1103/PhysRevApplied.3.014007](https://link.aps.org/doi/10.1103/PhysRevApplied.3.014007). URL: <https://link.aps.org/doi/10.1103/PhysRevApplied.3.014007>.
- [26] Francesco Giazotto et al. “Opportunities for mesoscopes in thermometry and refrigeration: Physics and applications”. In: *Rev. Mod. Phys.* 78 (1 Mar. 2006), pp. 217–274. DOI: [10.1103/RevModPhys.78.217](https://link.aps.org/doi/10.1103/RevModPhys.78.217). URL: <https://link.aps.org/doi/10.1103/RevModPhys.78.217>.
- [27] Dustin Kleckner and Dirk Bouwmeester. “Sub-kelvin optical cooling of a micromechanical resonator”. In: *Nature* 444.7115 (Nov. 2006), pp. 75–78. ISSN: 1476-4687. DOI: [10.1038/nature05231](https://doi.org/10.1038/nature05231). URL: <https://doi.org/10.1038/nature05231>.
- [28] A. Schliesser et al. “Resolved-sideband cooling of a micromechanical oscillator”. In: *Nature Physics* 4.5 (May 2008), pp. 415–419. ISSN: 1745-2481. DOI: [10.1038/nphys939](https://doi.org/10.1038/nphys939). URL: <https://doi.org/10.1038/nphys939>.
- [29] T. Rocheleau et al. “Preparation and detection of a mechanical resonator near the ground state of motion”. In: *Nature* 463.7277 (Jan. 2010), pp. 72–75. ISSN: 1476-4687. DOI: [10.1038/nature08681](https://doi.org/10.1038/nature08681). URL: <https://doi.org/10.1038/nature08681>.
- [30] A. Atland and B. Simon. *Condensed matter field theory*. Cambridge University Press, 2010. ISBN: 978-0-521-76975-4.
- [31] Subir Sachdev. *Quantum Phase Transitions*. Cambridge University Press, 2009. DOI: [10.1017/cbo9780511973765](https://doi.org/10.1017/cbo9780511973765). URL: <https://doi.org/10.1017/cbo9780511973765>.
- [32] Paolo Zanardi, Matteo G. A. Paris, and Lorenzo Campos Venuti. “Quantum criticality as a resource for quantum estimation”. In: *Phys. Rev. A* 78 (4 Oct. 2008), p. 042105. DOI: [10.1103/PhysRevA.78.042105](https://link.aps.org/doi/10.1103/PhysRevA.78.042105). URL: <https://link.aps.org/doi/10.1103/PhysRevA.78.042105>.
- [33] N. D. Mermin and H. Wagner. “Absence of Ferromagnetism or Antiferromagnetism in One- or Two-Dimensional Isotropic Heisenberg Models”. In: *Phys. Rev. Lett.* 17 (22 Nov. 1966), pp. 1133–1136. DOI: [10.1103/PhysRevLett.17.1133](https://link.aps.org/doi/10.1103/PhysRevLett.17.1133). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.17.1133>.
- [34] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2011. ISBN: 978-1107002173.
- [35] Søren Gammelmark. “Efficient parametric inference, estimation and simulation of open quantum system”. PhD thesis. Aarhus University, 2013.

- [36] U. H. Thygesen. *Lecture notes on Diffusion and Stochastic Differential Equations*. Polyteknisk boghandel, 2020.
- [37] H.-P. Breuer and F. Petruccione. *The Theory of Open Quantum Systems*. Oxford University Press, 2002. ISBN: 978-0-19-921390-0.
- [38] H. J. Carmichael. *Statistical Methods in Quantum Optics 1: Master Equations and Fokker-Planck Equations*. Springer, 2003. ISBN: 978-3540548829.
- [39] John Goold et al. “The role of quantum information in thermodynamics—a topical review”. In: *Journal of Physics A: Mathematical and Theoretical* 49.14 (Feb. 2016), p. 143001. DOI: [10.1088/1751-8113/49/14/143001](https://doi.org/10.1088/1751-8113/49/14/143001). URL: <https://doi.org/10.1088/1751-8113/49/14/143001>.
- [40] Sheldon Goldstein et al. “Canonical Typicality”. In: *Phys. Rev. Lett.* 96 (5 2006), p. 050403. DOI: [10.1103/PhysRevLett.96.050403](https://doi.org/10.1103/PhysRevLett.96.050403).
- [41] E. T. Jaynes. “Information Theory and Statistical Mechanics”. In: *Phys. Rev.* 106 (4 May 1957), pp. 620–630. DOI: [10.1103/PhysRev.106.620](https://link.aps.org/doi/10.1103/PhysRev.106.620). URL: <https://link.aps.org/doi/10.1103/PhysRev.106.620>.
- [42] E. T. Jaynes. “Information Theory and Statistical Mechanics. II”. In: *Phys. Rev.* 108 (2 Oct. 1957), pp. 171–190. DOI: [10.1103/PhysRev.108.171](https://link.aps.org/doi/10.1103/PhysRev.108.171). URL: <https://link.aps.org/doi/10.1103/PhysRev.108.171>.
- [43] Yuto Ashida, Keiji Saito, and Masahito Ueda. “Thermalization and Heating Dynamics in Open Generic Many-Body Systems”. In: *Phys. Rev. Lett.* 121 (17 Oct. 2018), p. 170402. DOI: [10.1103/PhysRevLett.121.170402](https://link.aps.org/doi/10.1103/PhysRevLett.121.170402). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.121.170402>.
- [44] Christian Gogolin and Jens Eisert. “Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems”. In: *Reports on Progress in Physics* 79.5 (Apr. 2016), p. 056001. DOI: [10.1088/0034-4885/79/5/056001](https://doi.org/10.1088/0034-4885/79/5/056001). URL: <https://doi.org/10.1088/0034-4885/79/5/056001>.
- [45] Mohammad Mehboudi, Anna Sanpera, and Luis A Correa. “Thermometry in the quantum regime: recent theoretical progress”. In: *Journal of Physics A: Mathematical and Theoretical* 52.30 (July 2019), p. 303001. DOI: [10.1088/1751-8121/ab2828](https://doi.org/10.1088/1751-8121/ab2828). URL: <https://doi.org/10.1088/1751-8121/ab2828>.
- [46] Antonella De Pasquale and Thomas M. Stace. “Quantum Thermometry”. In: *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*. Ed. by Felix Binder et al. Vol. 195. 2018, p. 503. DOI: [10.1007/978-3-319-99046-0_21](https://doi.org/10.1007/978-3-319-99046-0_21).
- [47] Ian Jermyn. “Invariant Bayesian estimation on manifolds”. In: *Ann. Statist.* 33 (0 Nov. 2005), pp. 583–605. DOI: [10.1214/009053604000001273](https://projecteuclid.org/euclid.aos/1117114330). URL: <https://projecteuclid.org/euclid.aos/1117114330>.
- [48] Samuel L. Braunstein and Carlton M. Caves. “Statistical distance and the geometry of quantum states”. In: *Phys. Rev. Lett.* 72 (22 May 1994), pp. 3439–3443. DOI: [10.1103/PhysRevLett.72.3439](https://link.aps.org/doi/10.1103/PhysRevLett.72.3439). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.72.3439>.
- [49] Gavin E. Crooks. “Measuring Thermodynamic Length”. In: *Phys. Rev. Lett.* 99 (10 Sept. 2007), p. 100602. DOI: [10.1103/PhysRevLett.99.100602](https://link.aps.org/doi/10.1103/PhysRevLett.99.100602). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.99.100602>.

- [50] Luis A. Correa et al. “Individual Quantum Probes for Optimal Thermometry”. In: *Phys. Rev. Lett.* 114 (22 June 2015), p. 220405. DOI: [10.1103/PhysRevLett.114.220405](https://doi.org/10.1103/PhysRevLett.114.220405). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.114.220405>.
- [51] Patrick P. Potts, Jonatan Bohr Brask, and Nicolas Brunner. “Fundamental limits on low-temperature quantum thermometry with finite resolution”. In: *Quantum* 3 (July 2019), p. 161. ISSN: 2521-327X. DOI: [10.22331/q-2019-07-09-161](https://doi.org/10.22331/q-2019-07-09-161). URL: <https://doi.org/10.22331/q-2019-07-09-161>.
- [52] Mathias R. Jørgensen et al. “Bayesian quantum thermometry based on thermodynamic length”. In: *arXiv e-prints*, arXiv:2108.05901 (Aug. 2021), arXiv:2108.05901. arXiv: [2108.05901](https://arxiv.org/abs/2108.05901) [quant-ph].
- [53] E. T. Jaynes. *Probability Theory: The Logic of Science*. Ed. by G. Larry Editor Bretthorst. Cambridge University Press, 2003. DOI: [10.1017/CB09780511790423](https://doi.org/10.1017/CB09780511790423).
- [54] Harry L Van Trees. *Detection, estimation, and modulation theory, part I: detection, estimation, and linear modulation theory*. John Wiley & Sons, 2004.
- [55] H. L. Van Trees and K. L. Bell. *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Wiley, 2007. ISBN: 9780470120958.
- [56] Erich Leo Lehmann and George Casella. *Theory of Point Estimation*. Vol. 31. Springer, 1998.
- [57] S. Amari. *Differential-Geometrical Methods in Statistics*. Springer-Verlag, 1985.
- [58] Shun-Ichi Amari and Hiroshi Nagaoka. *Methods of Information Geometry*. Vol. 191. American Mathematical Society, 2007.
- [59] G. Casella and R. L. Berger. *Statistical inference*. Australia: Thomson Learning, 2002.
- [60] R. Barlow. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*. Wiley, 2013. ISBN: 978-0-471-92295-7.
- [61] Udo von Toussaint. “Bayesian inference in physics”. In: *Rev. Mod. Phys.* 83 (3 Sept. 2011), pp. 943–999. DOI: [10.1103/RevModPhys.83.943](https://doi.org/10.1103/RevModPhys.83.943). URL: <https://link.aps.org/doi/10.1103/RevModPhys.83.943>.
- [62] Matteo G. A. Paris. “Quantum estimation for quantum technology”. In: *International Journal of Quantum Information* 07.supp01 (2009), pp. 125–137. DOI: [10.1142/S0219749909004839](https://doi.org/10.1142/S0219749909004839). eprint: <https://doi.org/10.1142/S0219749909004839>. URL: <https://doi.org/10.1142/S0219749909004839>.
- [63] Marcin Jarzyna and Jan Kołodyński. “Geometric Approach to Quantum Statistical Inference”. In: *IEEE Journal on Selected Areas in Information Theory* 1.2 (2020), pp. 367–386. DOI: [10.1109/JSAIT.2020.3017469](https://doi.org/10.1109/JSAIT.2020.3017469).
- [64] N. N. Chentsov. “Algebraic foundation of mathematical statistics”. In: *Math. Operationsforsch. statist.* 9 (1978), pp. 267–276.
- [65] F. Weinhold. “Metric geometry of equilibrium thermodynamics”. In: *The Journal of Chemical Physics* 63.6 (Sept. 1975), pp. 2479–2483. DOI: [10.1063/1.431689](https://doi.org/10.1063/1.431689). URL: <https://doi.org/10.1063/1.431689>.

- [66] Peter Salamon and R. Stephen Berry. “Thermodynamic Length and Dissipated Availability”. In: *Physical Review Letters* 51.13 (Sept. 1983), pp. 1127–1130. DOI: [10.1103/physrevlett.51.1127](https://doi.org/10.1103/physrevlett.51.1127). URL: <https://doi.org/10.1103/physrevlett.51.1127>.
- [67] Matteo Scandi and Martí Perarnau-Llobet. “Thermodynamic length in open quantum systems”. In: *Quantum* 3 (Oct. 2019), p. 197. ISSN: 2521-327X. DOI: [10.22331/q-2019-10-24-197](https://doi.org/10.22331/q-2019-10-24-197). URL: <https://doi.org/10.22331/q-2019-10-24-197>.
- [68] Ariel Caticha. “Lectures on Probability, Entropy, and Statistical Physics”. In: *arXiv e-prints*, arXiv:0808.0012 (July 2008), arXiv:0808.0012. arXiv: [0808.0012](https://arxiv.org/abs/0808.0012) [[physics.data-an](https://arxiv.org/archive/physics)].
- [69] Yan Li et al. “Frequentist and Bayesian Quantum Phase Estimation”. In: *Entropy* 20.9 (2018). ISSN: 1099-4300. DOI: [10.3390/e20090628](https://doi.org/10.3390/e20090628). URL: <https://www.mdpi.com/1099-4300/20/9/628>.
- [70] Lucien Bacharach et al. “Some Results on Tighter Bayesian Lower Bounds on the Mean-Square Error”. In: *arXiv e-prints*, arXiv:1907.09509 (July 2019), arXiv:1907.09509. arXiv: [1907.09509](https://arxiv.org/abs/1907.09509) [[cs.IT](https://arxiv.org/archive/cs)].
- [71] Katarzyna Macieszczak, Martin Fraas, and Rafał Demkowicz-Dobrzański. “Bayesian quantum frequency estimation in presence of collective dephasing”. In: *New Journal of Physics* 16.11 (Oct. 2014), p. 113002. DOI: [10.1088/1367-2630/16/11/113002](https://doi.org/10.1088/1367-2630/16/11/113002). URL: <https://doi.org/10.1088/1367-2630/16/11/113002>.
- [72] Jesús Rubio and Jacob Dunningham. “Bayesian multiparameter quantum metrology with limited data”. In: *Phys. Rev. A* 101 (3 Mar. 2020), p. 032114. DOI: [10.1103/PhysRevA.101.032114](https://doi.org/10.1103/PhysRevA.101.032114). URL: <https://link.aps.org/doi/10.1103/PhysRevA.101.032114>.
- [73] Richard Nickl. *Statistical theory*. Statistical Laboratory, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 2013. URL: http://www.statslab.cam.ac.uk/~nickl/Site/_files/stat2013.pdf.
- [74] A. W. van der Vaart. *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998. DOI: [10.1017/CB09780511802256](https://doi.org/10.1017/CB09780511802256).
- [75] Natalia A. Bochkina and Peter J. Green. “The Bernstein–von Mises theorem and nonregular models”. In: *The Annals of Statistics* 42.5 (2014), pp. 1850–1878. DOI: [10.1214/14-AOS1239](https://doi.org/10.1214/14-AOS1239). URL: <https://doi.org/10.1214/14-AOS1239>.
- [76] Harald Cramer. *Mathematical methods of statistics*. Princeton University Press, 1999. ISBN: 978-0691005478.
- [77] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. “Quantum Metrology”. In: *Phys. Rev. Lett.* 96 (1 Jan. 2006), p. 010401. DOI: [10.1103/PhysRevLett.96.010401](https://doi.org/10.1103/PhysRevLett.96.010401). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.96.010401>.
- [78] Jesús Rubio, Janet Anders, and Luis A. Correa. “Global Quantum Thermometry”. In: *arXiv e-prints*, arXiv:2011.13018 (Nov. 2020), arXiv:2011.13018. arXiv: [2011.13018](https://arxiv.org/abs/2011.13018) [[quant-ph](https://arxiv.org/archive/quant)].

- [79] Gabriel O. Alves and Gabriel T. Landi. “Bayesian estimation for collisional thermometry”. In: *arXiv e-prints*, arXiv:2106.12072 (June 2021), arXiv:2106.12072. arXiv: [2106.12072](https://arxiv.org/abs/2106.12072) [quant-ph].
- [80] Wai-Keong Mok et al. “Optimal Probes for Global Quantum Thermometry”. In: *arXiv e-prints*, arXiv:2010.14200 (Oct. 2020), arXiv:2010.14200. arXiv: [2010.14200](https://arxiv.org/abs/2010.14200) [quant-ph].
- [81] Jesús Rubio, Paul Knott, and Jacob Dunningham. “Non-asymptotic analysis of quantum metrology protocols beyond the Cramér–Rao bound”. In: *Journal of Physics Communications* 2.1 (Jan. 2018), p. 015027. DOI: [10.1088/2399-6528/aaa234](https://doi.org/10.1088/2399-6528/aaa234). URL: <https://doi.org/10.1088/2399-6528/aaa234>.
- [82] Giacomo Guarnieri et al. “Thermodynamics of precision in quantum nonequilibrium steady states”. In: *Phys. Rev. Research* 1 (3 Oct. 2019), p. 033021. DOI: [10.1103/PhysRevResearch.1.033021](https://doi.org/10.1103/PhysRevResearch.1.033021). URL: <https://link.aps.org/doi/10.1103/PhysRevResearch.1.033021>.
- [83] Matteo G A Paris. “Achieving the Landau bound to precision of quantum thermometry in systems with vanishing gap”. In: *Journal of Physics A: Mathematical and Theoretical* 49.3 (Dec. 2015), 03LT02. DOI: [10.1088/1751-8113/49/3/03lt02](https://doi.org/10.1088/1751-8113/49/3/03lt02). URL: <https://doi.org/10.1088/1751-8113/49/3/03lt02>.
- [84] T. Jahnke, S. Lanéry, and G. Mahler. “Operational approach to fluctuations of thermodynamic variables in finite quantum systems”. In: *Phys. Rev. E* 83 (1 Jan. 2011), p. 011109. DOI: [10.1103/PhysRevE.83.011109](https://doi.org/10.1103/PhysRevE.83.011109). URL: <https://link.aps.org/doi/10.1103/PhysRevE.83.011109>.
- [85] Felix A. Pollock et al. “Operational Markov Condition for Quantum Processes”. In: *Phys. Rev. Lett.* 120 (4 Jan. 2018), p. 040405. DOI: [10.1103/PhysRevLett.120.040405](https://doi.org/10.1103/PhysRevLett.120.040405). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.120.040405>.
- [86] V. Vedral. “The role of relative entropy in quantum information theory”. In: *Rev. Mod. Phys.* 74 (1 Mar. 2002), pp. 197–234. DOI: [10.1103/RevModPhys.74.197](https://doi.org/10.1103/RevModPhys.74.197). URL: <https://link.aps.org/doi/10.1103/RevModPhys.74.197>.
- [87] Marlon Brenes et al. “Multipartite Entanglement Structure in the Eigenstate Thermalization Hypothesis”. In: *Phys. Rev. Lett.* 124 (4 Jan. 2020), p. 040605. DOI: [10.1103/PhysRevLett.124.040605](https://doi.org/10.1103/PhysRevLett.124.040605). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.124.040605>.
- [88] Mathias R. Jørgensen et al. “Tight bound on finite-resolution quantum thermometry at low temperatures”. In: *Phys. Rev. Research* 2 (3 Sept. 2020), p. 033394. DOI: [10.1103/PhysRevResearch.2.033394](https://doi.org/10.1103/PhysRevResearch.2.033394). URL: <https://link.aps.org/doi/10.1103/PhysRevResearch.2.033394>.
- [89] Ryan Olf et al. “Thermometry and cooling of a Bose gas to 0.02 times the condensation temperature”. In: *Nature Physics* 11 (9 Sept. 2015), pp. 720–723. DOI: [10.1038/nphys3408](https://doi.org/10.1038/nphys3408). URL: <https://doi.org/10.1038/nphys3408>.
- [90] David M. Weld et al. “Spin Gradient Thermometry for Ultracold Atoms in Optical Lattices”. In: *Phys. Rev. Lett.* 103 (24 Dec. 2009), p. 245301. DOI: [10.1103/PhysRevLett.103.245301](https://doi.org/10.1103/PhysRevLett.103.245301). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.103.245301>.

- [91] Karen V. Hovhannisyanyan and Luis A. Correa. “Measuring the temperature of cold many-body quantum systems”. In: *Phys. Rev. B* 98 (4 July 2018), p. 045101. DOI: [10.1103/PhysRevB.98.045101](https://doi.org/10.1103/PhysRevB.98.045101). URL: <https://link.aps.org/doi/10.1103/PhysRevB.98.045101>.
- [92] Zvika Ben-Haim and Yonina C. Eldar. “A Lower Bound on the Bayesian MSE Based on the Optimal Bias Function”. In: *arXiv e-prints*, arXiv:0804.4391 (Apr. 2008), arXiv:0804.4391. arXiv: [0804.4391](https://arxiv.org/abs/0804.4391) [cs.IT].
- [93] Mohammad Mehboudi et al. “Fundamental limits in Bayesian thermometry and attainability via adaptive strategies”. In: *arXiv e-prints*, arXiv:2108.05932 (Aug. 2021), arXiv:2108.05932. arXiv: [2108.05932](https://arxiv.org/abs/2108.05932) [quant-ph].
- [94] Alessio Celi et al. “Quantum optics and frontiers of physics: the third quantum revolution”. In: *Physica Scripta* 92.1 (Dec. 2016), p. 013003. DOI: [10.1088/1402-4896/92/1/013003](https://doi.org/10.1088/1402-4896/92/1/013003). URL: <https://doi.org/10.1088/1402-4896/92/1/013003>.
- [95] Immanuel Bloch, Jean Dalibard, and Wilhelm Zwerger. “Many-body physics with ultracold gases”. In: *Rev. Mod. Phys.* 80 (3 July 2008), pp. 885–964. DOI: [10.1103/RevModPhys.80.885](https://doi.org/10.1103/RevModPhys.80.885). URL: <https://link.aps.org/doi/10.1103/RevModPhys.80.885>.
- [96] Immanuel Bloch, Jean Dalibard, and Sylvain Nascimbène. “Quantum simulations with ultracold quantum gases”. In: *Nature Physics* 8.4 (Apr. 2012), pp. 267–276. ISSN: 1745-2481. DOI: [10.1038/nphys2259](https://doi.org/10.1038/nphys2259). URL: <https://doi.org/10.1038/nphys2259>.
- [97] A. E. Leanhardt et al. “Cooling Bose-Einstein Condensates Below 500 Picokelvin”. In: *Science* 301.5639 (2003), pp. 1513–1515. ISSN: 0036-8075. DOI: [10.1126/science.1088827](https://doi.org/10.1126/science.1088827). eprint: <https://science.sciencemag.org/content/301/5639/1513.full.pdf>. URL: <https://science.sciencemag.org/content/301/5639/1513>.
- [98] Marco Scigliuzzo et al. “Primary Thermometry of Propagating Microwaves in the Quantum Regime”. In: *Phys. Rev. X* 10 (4 Dec. 2020), p. 041054. DOI: [10.1103/PhysRevX.10.041054](https://doi.org/10.1103/PhysRevX.10.041054). URL: <https://link.aps.org/doi/10.1103/PhysRevX.10.041054>.
- [99] Rudolf Gati et al. “Noise Thermometry with Two Weakly Coupled Bose-Einstein Condensates”. In: *Phys. Rev. Lett.* 96 (13 Apr. 2006), p. 130404. DOI: [10.1103/PhysRevLett.96.130404](https://doi.org/10.1103/PhysRevLett.96.130404). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.96.130404>.
- [100] Alberto Ronzani et al. “Tunable photonic heat transport in a quantum heat valve”. In: *Nature Physics* 14.10 (Oct. 2018), pp. 991–995. ISSN: 1745-2481. DOI: [10.1038/s41567-018-0199-4](https://doi.org/10.1038/s41567-018-0199-4). URL: <https://doi.org/10.1038/s41567-018-0199-4>.
- [101] Kuan Yen Tan et al. “Quantum-circuit refrigerator”. In: *Nature Communications* 8.1 (May 2017), p. 15189. ISSN: 2041-1723. DOI: [10.1038/ncomms15189](https://doi.org/10.1038/ncomms15189). URL: <https://doi.org/10.1038/ncomms15189>.

- [102] Daniel Adam et al. “Coherent and dephasing spectroscopy for single-impurity probing of an ultracold bath”. In: *arXiv e-prints*, arXiv:2105.03331 (May 2021), arXiv:2105.03331. arXiv: [2105.03331](https://arxiv.org/abs/2105.03331) [quant-ph].
- [103] Sania Jevtic et al. “Exchange fluctuation theorem for correlated quantum systems”. In: *Phys. Rev. E* 92 (4 Oct. 2015), p. 042113. DOI: [10.1103/PhysRevE.92.042113](https://doi.org/10.1103/PhysRevE.92.042113). URL: <https://link.aps.org/doi/10.1103/PhysRevE.92.042113>.
- [104] Antonella De Pasquale et al. “Local quantum thermal susceptibility”. In: *Nature Communications* 7.1 (Sept. 2016), p. 12782. ISSN: 2041-1723. DOI: [10.1038/ncomms12782](https://doi.org/10.1038/ncomms12782). URL: <https://doi.org/10.1038/ncomms12782>.
- [105] Sholeh Razavian et al. “Quantum thermometry by single-qubit dephasing”. In: *The European Physical Journal Plus* 134.6 (June 2019), p. 284. ISSN: 2190-5444. DOI: [10.1140/epjp/i2019-12708-9](https://doi.org/10.1140/epjp/i2019-12708-9). URL: <https://doi.org/10.1140/epjp/i2019-12708-9>.
- [106] Luis A. Correa et al. “Enhancement of low-temperature thermometry by strong coupling”. In: *Phys. Rev. A* 96 (6 Dec. 2017), p. 062103. DOI: [10.1103/PhysRevA.96.062103](https://doi.org/10.1103/PhysRevA.96.062103). URL: <https://link.aps.org/doi/10.1103/PhysRevA.96.062103>.
- [107] M Mehboudi et al. “Thermometry precision in strongly correlated ultracold lattice gases”. In: *New Journal of Physics* 17.5 (May 2015), p. 055020. DOI: [10.1088/1367-2630/17/5/055020](https://doi.org/10.1088/1367-2630/17/5/055020). URL: <https://doi.org/10.1088/1367-2630/17/5/055020>.
- [108] Safoura S. Mirkhalaf et al. “Criticality-enhanced quantum sensing in ferromagnetic Bose-Einstein condensates: Role of readout measurement and detection noise”. In: *Phys. Rev. A* 103 (2 Feb. 2021), p. 023317. DOI: [10.1103/PhysRevA.103.023317](https://doi.org/10.1103/PhysRevA.103.023317). URL: <https://link.aps.org/doi/10.1103/PhysRevA.103.023317>.
- [109] Patrick P. Hofer et al. “Quantum Thermal Machine as a Thermometer”. In: *Phys. Rev. Lett.* 119 (9 Sept. 2017), p. 090603. DOI: [10.1103/PhysRevLett.119.090603](https://doi.org/10.1103/PhysRevLett.119.090603). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.119.090603>.
- [110] Stella Seah et al. “Collisional Quantum Thermometry”. In: *Phys. Rev. Lett.* 123 (18 Oct. 2019), p. 180602. DOI: [10.1103/PhysRevLett.123.180602](https://doi.org/10.1103/PhysRevLett.123.180602). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.123.180602>.
- [111] Gerardo A. Paz-Silva, Leigh M. Norris, and Lorenza Viola. “Multiqubit spectroscopy of Gaussian quantum noise”. In: *Phys. Rev. A* 95 (2 Feb. 2017), p. 022121. DOI: [10.1103/PhysRevA.95.022121](https://doi.org/10.1103/PhysRevA.95.022121). URL: <https://link.aps.org/doi/10.1103/PhysRevA.95.022121>.
- [112] Steve Campbell et al. “Global and local thermometry schemes in coupled quantum systems”. In: *New Journal of Physics* 19.10 (Oct. 2017), p. 103003. DOI: [10.1088/1367-2630/aa7fac](https://doi.org/10.1088/1367-2630/aa7fac). URL: <https://doi.org/10.1088/1367-2630/aa7fac>.
- [113] C L Latune, I Sinayskiy, and F Petruccione. “Collective heat capacity for quantum thermometry and quantum engine enhancements”. In: *New Journal of Physics* 22.8 (Aug. 2020), p. 083049. DOI: [10.1088/1367-2630/aba463](https://doi.org/10.1088/1367-2630/aba463). URL: <https://doi.org/10.1088/1367-2630/aba463>.

- [114] Guim Planella et al. “Bath-induced correlations lead to sub-shot-noise thermometry precision”. In: *arXiv e-prints*, arXiv:2001.11812 (Jan. 2020), arXiv:2001.11812. arXiv: [2001.11812](https://arxiv.org/abs/2001.11812) [quant-ph].
- [115] Marcin Płodzień, Rafał Demkowicz-Dobrzański, and Tomasz Sowiński. “Few-fermion thermometry”. In: *Phys. Rev. A* 97 (6 June 2018), p. 063619. DOI: [10.1103/PhysRevA.97.063619](https://doi.org/10.1103/PhysRevA.97.063619). URL: <https://link.aps.org/doi/10.1103/PhysRevA.97.063619>.
- [116] Steve Campbell, Marco G Genoni, and Sebastian Deffner. “Precision thermometry and the quantum speed limit”. In: *Quantum Science and Technology* 3.2 (Feb. 2018), p. 025002. DOI: [10.1088/2058-9565/aaa641](https://doi.org/10.1088/2058-9565/aaa641). URL: <https://doi.org/10.1088/2058-9565/aaa641>.
- [117] Ricardo Román-Ancheyta, Barış Çakmak, and Özgür E Müstecaplıoğlu. “Spectral signatures of non-thermal baths in quantum thermalization”. In: *Quantum Science and Technology* 5.1 (Dec. 2019), p. 015003. DOI: [10.1088/2058-9565/ab5e4f](https://doi.org/10.1088/2058-9565/ab5e4f). URL: <https://doi.org/10.1088/2058-9565/ab5e4f>.
- [118] Kerson Huang. *Introduction to statistical physics*. Chapman and Hall/CRC, 2009.
- [119] Mathias R. Jørgensen et al. “Bayesian quantum thermometry based on thermodynamic length”. In: (2021).
- [120] D. W. Berry and H. M. Wiseman. “Optimal States and Almost Optimal Adaptive Measurements for Quantum Interferometry”. In: *Phys. Rev. Lett.* 85 (24 Dec. 2000), pp. 5098–5101. DOI: [10.1103/PhysRevLett.85.5098](https://doi.org/10.1103/PhysRevLett.85.5098). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.85.5098>.
- [121] Wojciech Górecki et al. “ π -Corrected Heisenberg Limit”. In: *Phys. Rev. Lett.* 124 (3 Jan. 2020), p. 030501. DOI: [10.1103/PhysRevLett.124.030501](https://doi.org/10.1103/PhysRevLett.124.030501). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.124.030501>.
- [122] H. M. Wiseman et al. “Adaptive Measurements in the Optical Quantum Information Laboratory”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 15.6 (Nov. 2009), pp. 1661–1672. ISSN: 1077-260X. DOI: [10.1109/JSTQE.2009.2020810](https://doi.org/10.1109/JSTQE.2009.2020810).
- [123] Vasco Cavina et al. “Bridging thermodynamics and metrology in nonequilibrium quantum thermometry”. In: *Phys. Rev. A* 98 (5 Nov. 2018), p. 050101. DOI: [10.1103/PhysRevA.98.050101](https://doi.org/10.1103/PhysRevA.98.050101). URL: <https://link.aps.org/doi/10.1103/PhysRevA.98.050101>.
- [124] Karen V. Hovhannisyán et al. “Optimal Quantum Thermometry with Coarse-Grained Measurements”. In: *PRX Quantum* 2 (2 May 2021), p. 020322. DOI: [10.1103/PRXQuantum.2.020322](https://doi.org/10.1103/PRXQuantum.2.020322). URL: <https://link.aps.org/doi/10.1103/PRXQuantum.2.020322>.
- [125] Irénée Frérot and Tommaso Roscilde. “Quantum Critical Metrology”. In: *Phys. Rev. Lett.* 121 (2 July 2018), p. 020402. DOI: [10.1103/PhysRevLett.121.020402](https://doi.org/10.1103/PhysRevLett.121.020402). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.121.020402>.

- [126] Marek M. Rams et al. “At the Limits of Criticality-Based Quantum Metrology: Apparent Super-Heisenberg Scaling Revisited”. In: *Phys. Rev. X* 8 (2 Apr. 2018), p. 021022. DOI: [10.1103/PhysRevX.8.021022](https://doi.org/10.1103/PhysRevX.8.021022). URL: <https://link.aps.org/doi/10.1103/PhysRevX.8.021022>.
- [127] Mathias R. Jørgensen and Felix A. Pollock. “Exploiting the Causal Tensor Network Structure of Quantum Processes to Efficiently Simulate Non-Markovian Path Integrals”. In: *Phys. Rev. Lett.* 123 (24 Dec. 2019), p. 240602. DOI: [10.1103/PhysRevLett.123.240602](https://doi.org/10.1103/PhysRevLett.123.240602). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.123.240602>.
- [128] Luís Dias Carlos and Fernando Palacio, eds. *Thermometry at the Nanoscale*. The Royal Society of Chemistry, 2016, P007–522. DOI: [10.1039/9781782622031](https://doi.org/10.1039/9781782622031).
- [129] S. Gasparinetti et al. “Probing the local temperature of a two-dimensional electron gas microdomain with a quantum dot: Measurement of electron-phonon interaction”. In: *Phys. Rev. B* 83 (2011), 201306(R). DOI: [10.1103/PhysRevB.83.201306](https://doi.org/10.1103/PhysRevB.83.201306).
- [130] P. Neumann et al. “High-Precision Nanoscale Temperature Sensing Using Single Defects in Diamond”. In: *Nano Letters* 13 (6 May 2013), pp. 2738–2742. DOI: [10.1021/nl401216y](https://doi.org/10.1021/nl401216y). URL: <https://doi.org/10.1021/nl401216y>.
- [131] Carlos Sabin et al. “Impurities as a quantum thermometer for a Bose-Einstein condensate”. In: *Scientific Reports* 4 (Sept. 2014), p. 6436. DOI: [10.1038/srep06436](https://doi.org/10.1038/srep06436).
- [132] Michael R. Moldover, Weston L. Tew, and Howard W. Yoon. “Advances in thermometry”. In: *Nature Physics* 12 (2016), p. 7. DOI: [10.1038/nphys3618](https://doi.org/10.1038/nphys3618).
- [133] M. Palma et al. “On-and-off chip cooling of a Coulomb blockade thermometer down to 2.8 mK”. In: *Applied Physics Letters* 111.25 (2017), p. 253105. DOI: [10.1063/1.5002565](https://doi.org/10.1063/1.5002565). URL: <https://doi.org/10.1063/1.5002565>.
- [134] M. Zgirski et al. “Nanosecond Thermometry with Josephson Junctions”. In: *Phys. Rev. Applied* 10 (4 Oct. 2018), p. 044068. DOI: [10.1103/PhysRevApplied.10.044068](https://doi.org/10.1103/PhysRevApplied.10.044068). URL: <https://link.aps.org/doi/10.1103/PhysRevApplied.10.044068>.
- [135] Luca Mancino et al. “Quantum Simulation of Single-Qubit Thermometry Using Linear Optics”. In: *Phys. Rev. Lett.* 118 (13 Mar. 2017), p. 130502. DOI: [10.1103/PhysRevLett.118.130502](https://doi.org/10.1103/PhysRevLett.118.130502). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.118.130502>.
- [136] T. H. Johnson et al. “Thermometry of ultracold atoms via nonequilibrium work distributions”. In: *Phys. Rev. A* 93 (5 May 2016), p. 053619. DOI: [10.1103/PhysRevA.93.053619](https://doi.org/10.1103/PhysRevA.93.053619). URL: <https://link.aps.org/doi/10.1103/PhysRevA.93.053619>.
- [137] Alexander H. Küllerich, Antonella De Pasquale, and Vittorio Giovannetti. “Dynamical approach to ancilla-assisted quantum thermometry”. In: *Phys. Rev. A* 98 (4 Oct. 2018), p. 042124. DOI: [10.1103/PhysRevA.98.042124](https://doi.org/10.1103/PhysRevA.98.042124).
- [138] Michele M. Feyles et al. “Dynamical role of quantum signatures in quantum thermometry”. In: *Phys. Rev. A* 99 (6 June 2019), p. 062114. DOI: [10.1103/PhysRevA.99.062114](https://doi.org/10.1103/PhysRevA.99.062114).

- [139] Victor Mukherjee et al. “Enhanced precision bound of low-temperature quantum thermometry via dynamical control”. In: *Communications Physics* 2.1 (2019), pp. 1–8. DOI: <https://doi.org/10.1038/s42005-019-0265-y>. URL: <https://doi.org/10.1038/s42005-019-0265-y>.
- [140] Marcin Jarzyna and Marcin Zwierz. “Quantum interferometric measurements of temperature”. In: *Phys. Rev. A* 92 (3 Sept. 2015), p. 032112. DOI: [10.1103/PhysRevA.92.032112](https://doi.org/10.1103/PhysRevA.92.032112). URL: <https://link.aps.org/doi/10.1103/PhysRevA.92.032112>.
- [141] Thomas M. Stace. “Quantum limits of thermometry”. In: *Phys. Rev. A* 82 (1 July 2010), 011611(R). DOI: [10.1103/PhysRevA.82.011611](https://doi.org/10.1103/PhysRevA.82.011611). URL: <https://link.aps.org/doi/10.1103/PhysRevA.82.011611>.
- [142] Antonella De Pasquale, Kazuya Yuasa, and Vittorio Giovannetti. “Estimating temperature via sequential measurements”. In: *Phys. Rev. A* 96 (1 July 2017), p. 012316. DOI: [10.1103/PhysRevA.96.012316](https://doi.org/10.1103/PhysRevA.96.012316). URL: <https://link.aps.org/doi/10.1103/PhysRevA.96.012316>.
- [143] Florian Fröwis, Pavel Sekatski, and Wolfgang Dür. “Detecting Large Quantum Fisher Information with Finite Measurement Precision”. In: *Phys. Rev. Lett.* 116 (9 Mar. 2016), p. 090801. DOI: [10.1103/PhysRevLett.116.090801](https://doi.org/10.1103/PhysRevLett.116.090801). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.116.090801>.
- [144] Yelena Guryanova, Nicolai Friis, and Marcus Huber. “Ideal Projective Measurements Have Infinite Resource Costs”. In: *Quantum* 4 (Jan. 2020), p. 222. ISSN: 2521-327X. DOI: [10.22331/q-2020-01-13-222](https://doi.org/10.22331/q-2020-01-13-222). URL: <https://doi.org/10.22331/q-2020-01-13-222>.
- [145] Giacomo De Palma, Antonella De Pasquale, and Vittorio Giovannetti. “Universal locality of quantum thermal susceptibility”. In: *Phys. Rev. A* 95 (5 May 2017), p. 052115. DOI: [10.1103/PhysRevA.95.052115](https://doi.org/10.1103/PhysRevA.95.052115). URL: <https://link.aps.org/doi/10.1103/PhysRevA.95.052115>.
- [146] Harold Jeffreys. *Methods of mathematical physics*. Cambridge university press; 3rd edition, 2000. ISBN: 978-0521664028.
- [147] L. Ballentine. *Quantum mechanics: a modern development*. World Scientific, 2014. ISBN: 978-981-4578-58-5.
- [148] George B. Arfken, Hans J. Weber, and Frank E. Harris. *Mathematical methods for physicists: A comprehensive guide; 7th edition*. Academic press, 2013. ISBN: 978-0123846549.
- [149] Kevin Cahill. *Physical mathematics*. Cambridge University Press, 2013. ISBN: 978-1107005211.
- [150] Mark Wilde. *Quantum information theory; 2nd edition*. Cambridge University Press, 2017. ISBN: 978-1107176164.
- [151] Ines de Vega and Daniel Alonso. “Dynamics of non-Markovian open quantum systems”. In: *Rev. Mod. Phys.* 89 (1 Jan. 2017), p. 015001. DOI: [10.1103/RevModPhys.89.015001](https://doi.org/10.1103/RevModPhys.89.015001).
- [152] U. Weiss. *Quantum Dissipative Systems*. World Scientific, 2012. ISBN: 978-981-4374-91-0.

- [153] Ning-Hua Tong and Matthias Vojta. “Signatures of a Noise-Induced Quantum Phase Transition in a Mesoscopic Metal Ring”. In: *Phys. Rev. Lett.* 97 (1 July 2006), p. 016802. DOI: [10.1103/PhysRevLett.97.016802](https://doi.org/10.1103/PhysRevLett.97.016802). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.97.016802>.
- [154] A. Strathearn et al. “Efficient non-Markovian quantum dynamics using time-evolving matrix product operators”. In: *Nat. Commun.* 9 (2018), p. 3322. DOI: [10.1038/s41467-018-05617-3](https://doi.org/10.1038/s41467-018-05617-3).
- [155] Maximilian Buser et al. “Initial system-environment correlations via the transfer-tensor method”. In: *Phys. Rev. A* 96 (6 Dec. 2017), p. 062122. DOI: [10.1103/PhysRevA.96.062122](https://doi.org/10.1103/PhysRevA.96.062122). URL: <https://link.aps.org/doi/10.1103/PhysRevA.96.062122>.
- [156] Yanan Yue and Xinwei Wang. “Nanoscale thermal probing”. In: *Nano Rev.* 3.1 (2012), p. 11586. DOI: [10.3402/nano.v3i0.11586](https://doi.org/10.3402/nano.v3i0.11586).
- [157] Luís Dias Carlos and Fernando Palacio, eds. *Thermometry at the Nanoscale*. Nanoscience & Nanotechnology Series. The Royal Society of Chemistry, 2016, P007–522. ISBN: 978-1-84973-904-7. DOI: [10.1039/9781782622031](https://doi.org/10.1039/9781782622031).
- [158] Ugo Marzolino and Daniel Braun. “Precision measurements of temperature and chemical potential of quantum gases”. In: *Phys. Rev. A* 88 (6 2013), p. 063609. DOI: [10.1103/PhysRevA.88.063609](https://doi.org/10.1103/PhysRevA.88.063609).
- [159] Carlos Sabin et al. “Impurities as a quantum thermometer for a Bose-Einstein condensate”. In: *Sci. Rep.* 4 (2014), p. 6436. DOI: [10.1038/srep06436](https://doi.org/10.1038/srep06436).
- [160] Roberto Onofrio. “Cooling and thermometry of atomic Fermi gases”. In: *Phys. Usp.* 59.11 (2016), pp. 1129–1153. DOI: [10.3367/ufne.2016.07.037873](https://doi.org/10.3367/ufne.2016.07.037873).
- [161] Bayan Karimi and Jukka P. Pekola. “Quantum Trajectory Analysis of Single Microwave Photon Detection by Nanocalorimetry”. In: *Phys. Rev. Lett.* 124 (17 Apr. 2020), p. 170601. DOI: [10.1103/PhysRevLett.124.170601](https://doi.org/10.1103/PhysRevLett.124.170601). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.124.170601>.
- [162] Kieran D. B. Higgins, Brendon W. Lovett, and Erik M. Gauger. “Quantum thermometry using the ac Stark shift within the Rabi model”. In: *Phys. Rev. B* 88 (15 2013), p. 155409. DOI: [10.1103/PhysRevB.88.155409](https://doi.org/10.1103/PhysRevB.88.155409).
- [163] Florian Haupt, Atac Imamoglu, and Martin Kroner. “Single Quantum Dot as an Optical Thermometer for Millikelvin Temperatures”. In: *Phys. Rev. Applied* 2 (2 2014), p. 024001. DOI: [10.1103/PhysRevApplied.2.024001](https://doi.org/10.1103/PhysRevApplied.2.024001).
- [164] Matteo Brunelli, Stefano Olivares, and Matteo G. A. Paris. “Qubit thermometry for micromechanical resonators”. In: *Phys. Rev. A* 84 (3 2011), p. 032105. DOI: [10.1103/PhysRevA.84.032105](https://doi.org/10.1103/PhysRevA.84.032105).
- [165] Matteo Brunelli et al. “Qubit-assisted thermometry of a quantum harmonic oscillator”. In: *Phys. Rev. A* 86 (1 2012), p. 012125. DOI: [10.1103/PhysRevA.86.012125](https://doi.org/10.1103/PhysRevA.86.012125).
- [166] Antonella De Pasquale, Kazuya Yuasa, and Vittorio Giovannetti. “Estimating temperature via sequential measurements”. In: *Phys. Rev. A* 96 (1 2017), p. 012316. DOI: [10.1103/PhysRevA.96.012316](https://doi.org/10.1103/PhysRevA.96.012316).

- [167] Mark T. Mitchison et al. “*In Situ* Thermometry of a Cold Fermi Gas via Dephasing Impurities”. In: *Phys. Rev. Lett.* 125 (8 2020), p. 080402. DOI: [10.1103/PhysRevLett.125.080402](https://doi.org/10.1103/PhysRevLett.125.080402).
- [168] Luca Mancino et al. “Nonequilibrium readiness and precision of Gaussian quantum thermometers”. In: *Phys. Rev. Research* 2 (3 2020), p. 033498. DOI: [10.1103/PhysRevResearch.2.033498](https://doi.org/10.1103/PhysRevResearch.2.033498).
- [169] Ivan Henao and Raam Uzdin. “Catalytic transformations with finite-size environments: applications to cooling and thermometry”. In: (). arXiv: [2010.09070](https://arxiv.org/abs/2010.09070) [quant-ph].
- [170] Steve Campbell, Marco G. Genoni, and Sebastian Deffner. “Precision thermometry and the quantum speed limit”. In: *Quantum Sci. Technol.* 3.2 (2018), p. 025002. DOI: [10.1088/2058-9565/aaa641](https://doi.org/10.1088/2058-9565/aaa641).
- [171] H. J. D. Miller and J. Anders. “Energy-temperature uncertainty relation in quantum thermodynamics”. In: *Nature Communications* 9.1 (June 2018), p. 2203. ISSN: 2041-1723. DOI: [10.1038/s41467-018-04536-7](https://doi.org/10.1038/s41467-018-04536-7). URL: <https://doi.org/10.1038/s41467-018-04536-7>.
- [172] D. McKay and B. DeMarco. “Thermometry with spin-dependent lattices”. In: *New J. Phys.* 12.5 (2010), p. 055013. DOI: [10.1088/1367-2630/12/5/055013](https://doi.org/10.1088/1367-2630/12/5/055013).
- [173] Ryan Olf et al. “Thermometry and cooling of a Bose gas to 0.02 times the condensation temperature”. In: *Nat. Phys.* 11.9 (2015), pp. 720–723. DOI: [10.1038/nphys3408](https://doi.org/10.1038/nphys3408).
- [174] Michael Hohmann et al. “Single-atom thermometer for ultracold gases”. In: *Phys. Rev. A* 93 (4 2016), p. 043607. DOI: [10.1103/PhysRevA.93.043607](https://doi.org/10.1103/PhysRevA.93.043607).
- [175] Rianne S. Lous et al. “Thermometry of a deeply degenerate Fermi gas with a Bose-Einstein condensate”. In: *Phys. Rev. A* 95 (5 2017), p. 053627. DOI: [10.1103/PhysRevA.95.053627](https://doi.org/10.1103/PhysRevA.95.053627).
- [176] Li-Sha Guo et al. “Improved thermometry of low-temperature quantum systems by a ring-structure probe”. In: *Phys. Rev. A* 92 (5 2015), p. 052112. DOI: [10.1103/PhysRevA.92.052112](https://doi.org/10.1103/PhysRevA.92.052112).
- [177] Victor Mukherjee et al. “Enhanced precision bound of low-temperature quantum thermometry via dynamical control”. In: *Commun. Phys.* 2 (2019), p. 162. DOI: [10.1038/s42005-019-0265-y](https://doi.org/10.1038/s42005-019-0265-y).
- [178] Alessio Recati et al. “Atomic Quantum Dots Coupled to a Reservoir of a Superfluid Bose-Einstein Condensate”. In: *Phys. Rev. Lett.* 94 (4 2005), p. 040404. DOI: [10.1103/PhysRevLett.94.040404](https://doi.org/10.1103/PhysRevLett.94.040404).
- [179] Benoit B. Mandelbrot. “An outline of a purely phenomenological theory of statistical thermodynamics: I. Canonical ensembles”. In: *IRE Trans. Inf. Theory* 2.3 (1956), pp. 190–203. DOI: [10.1109/TIT.1956.1056804](https://doi.org/10.1109/TIT.1956.1056804).
- [180] Jos Uffink and Janneke van Lith. “Thermodynamic Uncertainty Relations”. In: *Found. Phys.* 29 (1999), pp. 655–692. DOI: [10.1023/A:1018811305766](https://doi.org/10.1023/A:1018811305766).
- [181] Robert G. Gallager. *Principles of Digital Communication*. Cambridge University Press, New York, 2008.
- [182] Solomon Kullback. *Information Theory and Statistics*. Dover, New York, 1968.

- [183] Richard D. Gill and Boris Y. Levit. “Applications of the van Trees Inequality: A Bayesian Cramér-Rao Bound”. In: *Bernoulli* 1 (1995), pp. 59–79. DOI: [10.2307/3318681](https://doi.org/10.2307/3318681).
- [184] Edward G. Effros. “A matrix convexity approach to some celebrated quantum inequalities”. In: *Proc. Natl. Acad. Sci. U.S.A.* 106.4 (2009), pp. 1006–1008. DOI: [10.1073/pnas.0807965106](https://doi.org/10.1073/pnas.0807965106).
- [185] Amiya K. Chakravarty, James B. Orlin, and Uriel G. Rothblum. “Consecutive Optimizers for a Partitioning Problem with Applications to Optimal Inventory Groupings for Joint Replenishment”. In: *Oper. Res.* 33.4 (1985), pp. 820–834. DOI: [10.1287/opre.33.4.820](https://doi.org/10.1287/opre.33.4.820).
- [186] Jon P. Keating, Noah Linden, and Huw J. Wells. “Spectra and eigenstates of spin chain Hamiltonians”. In: *Commun. Math. Phys.* 338.1 (2015), pp. 81–102. DOI: [10.1007/s00220-015-2366-0](https://doi.org/10.1007/s00220-015-2366-0).
- [187] Fernando G. S. L. Brandão and Marcus Cramer. “Equivalence of Statistical Mechanical Ensembles for Non-Critical Quantum Systems”. In: (). arXiv: [1502.03263](https://arxiv.org/abs/1502.03263) [quant-ph].
- [188] Fernando G. S. L. Brandão, Marcus Cramer, and Mădălin Guță. “Berry-Esseen Theorem for quantum lattice systems and the equivalence of statistical mechanical ensembles”. In: *QIP2015 Talk* (2015). URL: <http://quantum-lab.org/qip2015/talks/125-Brandao.pdf>.
- [189] Barry Simon. *The Statistical Mechanics of Lattice Gases*. Vol. 1. Princeton University Press, Princeton, 1993.
- [190] Huzihiro Araki. “On the equivalence of the KMS condition and the variational principle for quantum lattice systems”. In: *Commun. Math. Phys.* 38 (1974), pp. 1–10. DOI: [10.1007/BF01651545](https://doi.org/10.1007/BF01651545).
- [191] Matthew B. Hastings. “Decay of Correlations in Fermi Systems at Nonzero Temperature”. In: *Phys. Rev. Lett.* 93 (12 2004), p. 126402. DOI: [10.1103/PhysRevLett.93.126402](https://doi.org/10.1103/PhysRevLett.93.126402).
- [192] Martin Kliesch et al. “Locality of Temperature”. In: *Phys. Rev. X* 4 (3 2014), p. 031019. DOI: [10.1103/PhysRevX.4.031019](https://doi.org/10.1103/PhysRevX.4.031019).
- [193] Terry Farrelly, Fernando G. S. L. Brandão, and Marcus Cramer. “Thermalization and return to equilibrium on finite quantum lattice systems”. In: *Phys. Rev. Lett.* 118 (14 2017), p. 140601. DOI: [10.1103/PhysRevLett.118.140601](https://doi.org/10.1103/PhysRevLett.118.140601).
- [194] M. E. Fisher. “The theory of equilibrium critical phenomena”. In: *Rep. Progr. Phys.* 30.2 (1967), pp. 615–730. DOI: [10.1088/0034-4885/30/2/306](https://doi.org/10.1088/0034-4885/30/2/306).
- [195] Édouard Brézin. “An investigation of finite size scaling”. In: *J. Phys. France* 43.1 (1982), pp. 15–22. DOI: [10.1051/jphys:0198200430101500](https://doi.org/10.1051/jphys:0198200430101500).
- [196] Nikolay Sh. Izmailian and Chin-Kun Hu. “Exact amplitude ratio and finite-size corrections for the $M \times N$ square lattice Ising model”. In: *Phys. Rev. E* 65 (3 2002), p. 036103. DOI: [10.1103/PhysRevE.65.036103](https://doi.org/10.1103/PhysRevE.65.036103).
- [197] Robert B. Griffiths. “A proof that the free energy of a spin system is extensive”. In: *J. Math. Phys.* 5.9 (1964), pp. 1215–1222. DOI: [10.1063/1.1704228](https://doi.org/10.1063/1.1704228).

- [198] Markus P. Müller et al. “Thermalization and Canonical Typicality in Translation-Invariant Quantum Lattice Systems”. In: *Commun. Math. Phys.* 340.2 (2015), pp. 499–561. DOI: [10.1007/s00220-015-2473-y](https://doi.org/10.1007/s00220-015-2473-y).
- [199] Theodore David Schultz, Daniel Charles Mattis, and Elliott Hershel Lieb. “Two-Dimensional Ising Model as a Soluble Problem of Many Fermions”. In: *Rev. Mod. Phys.* 36 (1964), pp. 856–871. DOI: [10.1103/RevModPhys.36.856](https://doi.org/10.1103/RevModPhys.36.856).
- [200] Paul D. Beale. “Exact Distribution of Energies in the Two-Dimensional Ising Model”. In: *Phys. Rev. Lett.* 76 (1 1996), pp. 78–81. DOI: [10.1103/PhysRevLett.76.78](https://doi.org/10.1103/PhysRevLett.76.78).
- [201] H. A. Kramers and G. H. Wannier. “Statistics of the Two-Dimensional ferromagnet. Part I”. In: *Phys. Rev.* 60 (1941), p. 252. DOI: [10.1103/PhysRev.60.252](https://doi.org/10.1103/PhysRev.60.252).
- [202] Lars Onsager. “Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition”. In: *Phys. Rev.* 65 (3-4 1944), pp. 117–149. DOI: [10.1103/PhysRev.65.117](https://doi.org/10.1103/PhysRev.65.117).
- [203] Rodney J. Baxter. *Exactly solved models in statistical mechanics*. Academic Press, London, 1982.
- [204] N. K. Langford et al. “Experimentally simulating the dynamics of quantum light and matter at deep-strong coupling”. In: *Nat. Commun.* 8 (2017), p. 1715. DOI: [10.1038/s41467-017-01061-x](https://doi.org/10.1038/s41467-017-01061-x).
- [205] C. A. Mujica-Martinez, P. Nalbach, and M. Thorwart. “Quantification of non-Markovian effects in the Fenna-Matthews-Olson complex”. In: *Phys. Rev. E* 88 (6 Dec. 2013), p. 062719. DOI: [10.1103/PhysRevE.88.062719](https://doi.org/10.1103/PhysRevE.88.062719). URL: <https://link.aps.org/doi/10.1103/PhysRevE.88.062719>.
- [206] Dara P. S. McCutcheon. “Optical signatures of non-Markovian behavior in open quantum systems”. In: *Phys. Rev. A* 93 (2 Feb. 2016), p. 022119. DOI: [10.1103/PhysRevA.93.022119](https://doi.org/10.1103/PhysRevA.93.022119). URL: <https://link.aps.org/doi/10.1103/PhysRevA.93.022119>.
- [207] David Newman, Florian Mintert, and Ahsan Nazir. “Performance of a quantum heat engine at strong reservoir coupling”. In: *Phys. Rev. E* 95 (1 Mar. 2017), p. 032139. DOI: [10.1103/PhysRevE.95.032139](https://doi.org/10.1103/PhysRevE.95.032139).
- [208] Bassano Vacchini and Heinz-Peter Breuer. “Exact master equations for the non-Markovian decay of a qubit”. In: *Phys. Rev. A* 81 (4 Apr. 2010), p. 042103. DOI: [10.1103/PhysRevA.81.042103](https://doi.org/10.1103/PhysRevA.81.042103). URL: <https://link.aps.org/doi/10.1103/PhysRevA.81.042103>.
- [209] Chee Kong Lee, Jeremy Moix, and Jianshu Cao. “Accuracy of second order perturbation theory in the polaron and variational polaron frames”. In: *J. Chem. Phys.* 136.20 (2012), p. 204120. DOI: [10.1063/1.4722336](https://doi.org/10.1063/1.4722336). URL: <https://doi.org/10.1063/1.4722336>.
- [210] Amir Fruchtmann, Neill Lambert, and Erik M. Gauger. “When do perturbative approaches accurately capture the dynamics of complex quantum systems?” In: *Sci. Rep.* 6 (2016), p. 28204. DOI: [10.1038/srep28204](https://doi.org/10.1038/srep28204). URL: <http://dx.doi.org/10.1038/srep28204>.

- [211] Nancy Makri and Dmitrii E. Makarov. “Tensor propagator for iterative quantum time evolution of reduced density matrices. I. Theory”. In: *J. Chem. Phys.* 102.11 (1995), p. 4600. DOI: [10.1063/1.469508](https://doi.org/10.1063/1.469508). URL: <https://doi.org/10.1063/1.469508>.
- [212] Nancy Makri and Dmitrii E. Makarov. “Tensor propagator for iterative quantum time evolution of reduced density matrices. II. Numerical methodology”. In: *J. Chem. Phys.* 102.11 (1995), p. 4611. DOI: [10.1063/1.469509](https://doi.org/10.1063/1.469509). URL: <https://doi.org/10.1063/1.469509>.
- [213] Peter Nalbach et al. “Iterative path-integral algorithm versus cumulant time-nonlocal master equation approach for dissipative biomolecular exciton transport”. In: *New J. Phys.* 13.6 (2011), p. 063040. DOI: <https://doi.org/10.1088/1367-2630/13/6/063040>. URL: <http://stacks.iop.org/1367-2630/13/i=6/a=063040>.
- [214] Nikesh S. Dattani. “Numerical Feynman integrals with physically inspired interpolation: Faster convergence and significant reduction of computational cost”. In: *AIP Adv.* 2.1 (2012), p. 012121. DOI: [10.1063/1.3680607](https://doi.org/10.1063/1.3680607).
- [215] A Strathearn, B W Lovett, and P Kirton. “Efficient real-time path integrals for non-Markovian spin-boson models”. In: *New J. Phys.* 19.9 (2017), p. 093009. DOI: [10.1088/1367-2630/aa8744](https://doi.org/10.1088/1367-2630/aa8744).
- [216] Qiang Shi and Eitan Geva. “A new approach to calculating the memory kernel of the generalized quantum master equation for an arbitrary system-bath coupling”. In: *J. Chem. Phys.* 119.23 (2003), pp. 12063–12076. DOI: [10.1063/1.1624830](https://doi.org/10.1063/1.1624830). URL: <https://doi.org/10.1063/1.1624830>.
- [217] Guy Cohen and Eran Rabani. “Memory effects in nonequilibrium quantum impurity models”. In: *Phys. Rev. B* 84 (7 Aug. 2011), p. 075150. DOI: [10.1103/PhysRevB.84.075150](https://link.aps.org/doi/10.1103/PhysRevB.84.075150). URL: <https://link.aps.org/doi/10.1103/PhysRevB.84.075150>.
- [218] Javier Cerrillo and Jianshu Cao. “Non-Markovian Dynamical Maps: Numerical Processing of Open Quantum Trajectories”. In: *Phys. Rev. Lett.* 112 (11 Mar. 2014), p. 110401. DOI: [10.1103/PhysRevLett.112.110401](https://doi.org/10.1103/PhysRevLett.112.110401). URL: <http://link.aps.org/doi/10.1103/PhysRevLett.112.110401>.
- [219] Robert Rosenbach et al. “Efficient simulation of non-Markovian system-environment interaction”. In: *New J. Phys.* 18.2 (Feb. 2016), p. 023035. DOI: [10.1088/1367-2630/18/2/023035](https://doi.org/10.1088/1367-2630/18/2/023035). URL: <https://doi.org/10.1088/1367-2630/18/2/023035>.
- [220] Maximilian Buser et al. “Initial system-environment correlations via the transfer-tensor method”. In: *Phys. Rev. A* 96 (6 Dec. 2017), p. 062122. DOI: [10.1103/PhysRevA.96.062122](https://doi.org/10.1103/PhysRevA.96.062122). URL: <https://link.aps.org/doi/10.1103/PhysRevA.96.062122>.
- [221] Andrius Gelzinis, Edvardas Rybakovas, and Leonas Valkunas. “Applicability of transfer tensor method for open quantum system dynamics”. In: *J. Chem. Phys.* 147.23 (2017), p. 234108. DOI: [10.1063/1.5009086](https://doi.org/10.1063/1.5009086). URL: <https://doi.org/10.1063/1.5009086>.

- [222] Felix A. Pollock and Kavan Modi. “Tomographically reconstructed master equations for any open quantum dynamics”. In: *Quantum* 2 (July 2018), p. 76. ISSN: 2521-327X. DOI: [10.22331/q-2018-07-11-76](https://doi.org/10.22331/q-2018-07-11-76). URL: <https://doi.org/10.22331/q-2018-07-11-76>.
- [223] Yoshitaka Tanimura. “Stochastic Liouville, Langevin, Fokker-Planck, and Master Equation Approaches to Quantum Dissipative Systems”. In: *J. Phys. Soc. Jpn.* 75.8 (2006), p. 082001. DOI: [10.1143/JPSJ.75.082001](https://doi.org/10.1143/JPSJ.75.082001). URL: <http://dx.doi.org/10.1143/JPSJ.75.082001>.
- [224] Johan Strümpfer and Klaus Schulten. “Open Quantum Dynamics Calculations with the Hierarchy Equations of Motion on Parallel Computers”. In: *J. Chem. Theory Comput.* 8.8 (2012), p. 2808. DOI: [10.1021/ct3003833](https://doi.org/10.1021/ct3003833). URL: <http://dx.doi.org/10.1021/ct3003833>.
- [225] Ralf Bulla, Theo A. Costi, and Thomas Pruschke. “Numerical renormalization group method for quantum impurity systems”. In: *Rev. Mod. Phys.* 80 (2 Apr. 2008), p. 395. DOI: [10.1103/RevModPhys.80.395](https://link.aps.org/doi/10.1103/RevModPhys.80.395). URL: <https://link.aps.org/doi/10.1103/RevModPhys.80.395>.
- [226] Hsing-Ta Chen, Guy Cohen, and David R. Reichman. “Inchworm Monte Carlo for exact non-adiabatic dynamics. I. Theory and algorithms”. In: *J. Chem. Phys.* 146.5 (2017), p. 054105. DOI: [10.1063/1.4974328](https://doi.org/10.1063/1.4974328). URL: <https://doi.org/10.1063/1.4974328>.
- [227] Hsing-Ta Chen, Guy Cohen, and David R. Reichman. “Inchworm Monte Carlo for exact non-adiabatic dynamics. II. Benchmarks and comparison with established methods”. In: *J. Chem. Phys.* 146.5 (2017), p. 054106. DOI: [10.1063/1.4974329](https://doi.org/10.1063/1.4974329). URL: <https://doi.org/10.1063/1.4974329>.
- [228] Inés de Vega and Daniel Alonso. “Dynamics of non-Markovian open quantum systems”. In: *Rev. Mod. Phys.* 89 (1 Jan. 2017), p. 015001. DOI: [10.1103/RevModPhys.89.015001](https://link.aps.org/doi/10.1103/RevModPhys.89.015001). URL: <https://link.aps.org/doi/10.1103/RevModPhys.89.015001>.
- [229] Alex W. Chin et al. “Exact mapping between system-reservoir quantum models and semi-infinite discrete chains using orthogonal polynomials”. In: *J. Math. Phys.* 51.9 (2010), p. 092109. DOI: [10.1063/1.3490188](https://doi.org/10.1063/1.3490188). URL: <https://doi.org/10.1063/1.3490188>.
- [230] Javier Prior et al. “Efficient Simulation of Strong System-Environment Interactions”. In: *Phys. Rev. Lett.* 105 (5 July 2010), p. 050404. DOI: [10.1103/PhysRevLett.105.050404](https://link.aps.org/doi/10.1103/PhysRevLett.105.050404). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.105.050404>.
- [231] Florian A. Y. N. Schröder and Alex W. Chin. “Simulating open quantum dynamics with time-dependent variational matrix product states: Towards microscopic correlation of environment dynamics and reduced system evolution”. In: *Phys. Rev. B* 93 (7 Feb. 2016), p. 075105. DOI: [10.1103/PhysRevB.93.075105](https://link.aps.org/doi/10.1103/PhysRevB.93.075105). URL: <https://link.aps.org/doi/10.1103/PhysRevB.93.075105>.
- [232] Michael L. Wall, Arghavan Safavi-Naini, and Ana Maria Rey. “Simulating generic spin-boson models with matrix product states”. In: *Phys. Rev. A* 94 (5 Nov. 2016), p. 053637. DOI: [10.1103/PhysRevA.94.053637](https://link.aps.org/doi/10.1103/PhysRevA.94.053637). URL: <https://link.aps.org/doi/10.1103/PhysRevA.94.053637>.

- [233] Felix A. Pollock et al. “Non-Markovian quantum processes: Complete framework and efficient characterization”. In: *Phys. Rev. A* 97 (1 Jan. 2018), p. 012127. DOI: [10.1103/PhysRevA.97.012127](https://doi.org/10.1103/PhysRevA.97.012127). URL: <https://link.aps.org/doi/10.1103/PhysRevA.97.012127>.
- [234] I. A. Luchnikov et al. “Simulation Complexity of Open Quantum Dynamics: Connection with Tensor Networks”. In: *Phys. Rev. Lett.* 122 (16 Apr. 2019), p. 160401. DOI: [10.1103/PhysRevLett.122.160401](https://doi.org/10.1103/PhysRevLett.122.160401). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.122.160401>.
- [235] Román Orús. “A practical introduction to tensor networks: Matrix product states and projected entangled pair states”. In: *Ann. Phys.* 349 (2014), p. 117. ISSN: 0003-4916. DOI: <https://doi.org/10.1016/j.aop.2014.06.013>.
- [236] Eunji Sim and Nancy Makri. “Filtered propagator functional for iterative dynamics of quantum dissipative systems”. In: *Comput. Phys. Commun.* 99.2 (1997), p. 335. ISSN: 0010-4655. DOI: [https://doi.org/10.1016/S0010-4655\(96\)00130-0](https://doi.org/10.1016/S0010-4655(96)00130-0).
- [237] Eunji Sim. “Quantum dynamics for a system coupled to slow baths: On-the-fly filtered propagator method”. In: *J. Chem. Phys.* 115.10 (2001), p. 4450. DOI: [10.1063/1.1394208](https://doi.org/10.1063/1.1394208).
- [238] Nikesh S. Dattani. “FeynDyn: A MATLAB program for fast numerical Feynman integral calculations for open quantum system dynamics on GPUs”. In: *Comput. Phys. Commun.* 184.12 (2013), p. 2828. ISSN: 0010-4655. DOI: <https://doi.org/10.1016/j.cpc.2013.07.001>.
- [239] Li Li, Michael J. W. Hall, and Howard M. Wiseman. “Concepts of quantum non-Markovianity: A hierarchy”. In: *Phys. Rep.* 759 (2018). Concepts of quantum non-Markovianity: A hierarchy, p. 1. ISSN: 0370-1573. DOI: <https://doi.org/10.1016/j.physrep.2018.07.001>.
- [240] Simon Milz, Felix A. Pollock, and Kavan Modi. “An introduction to operational quantum dynamics”. In: *Open Sys. Info. Dyn.* (Dec. 2017), p. 1740016. DOI: [10.1142/S1230161217400169](https://doi.org/10.1142/S1230161217400169).
- [241] Simon Milz et al. “Kolmogorov extension theorem for (quantum) causal modelling and general probabilistic theories”. In: *arXiv:1712.02589* (2017). URL: <https://www.arxiv.org/abs/1712.02589>.
- [242] Felix A. Pollock et al. “Operational Markov Condition for Quantum Processes”. In: *Phys. Rev. Lett.* 120 (4 Jan. 2018), p. 040405. DOI: [10.1103/PhysRevLett.120.040405](https://doi.org/10.1103/PhysRevLett.120.040405). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.120.040405>.
- [243] Philip Taranto et al. “Quantum Markov Order”. In: *Phys. Rev. Lett.* 122 (14 Apr. 2019), p. 140401. DOI: [10.1103/PhysRevLett.122.140401](https://doi.org/10.1103/PhysRevLett.122.140401). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.122.140401>.
- [244] Philip Taranto et al. “Structure of quantum stochastic processes with finite Markov order”. In: *Phys. Rev. A* 99 (4 Apr. 2019), p. 042108. DOI: [10.1103/PhysRevA.99.042108](https://doi.org/10.1103/PhysRevA.99.042108). URL: <https://link.aps.org/doi/10.1103/PhysRevA.99.042108>.

- [245] Fattah Sakuldee et al. “Non-Markovian quantum control as coherent stochastic trajectories”. In: *J. Phys. A: Math. Theor.* 51.41 (Sept. 2018), p. 414014. DOI: [10.1088/1751-8121/aabb1e](https://doi.org/10.1088/1751-8121/aabb1e). URL: <https://doi.org/10.1088%2F1751-8121%2Faabb1e>.
- [246] H. F. Trotter. “On the product of semi-groups of operators”. In: *Proc. Amer. Math. Soc.* 10 (1959), p. 545. URL: <https://doi.org/10.1090/S0002-9939-1959-0108732-6>.
- [247] R. P. Feynman and F. L. Vernon. “The theory of a general quantum system interacting with a linear dissipative system”. In: *Ann. Phys.* 24 (1963), p. 118. ISSN: 0003-4916. DOI: [https://doi.org/10.1016/0003-4916\(63\)90068-X](https://doi.org/10.1016/0003-4916(63)90068-X).
- [248] L. Diósi and L. Ferialdi. “General Non-Markovian Structure of Gaussian Master and Stochastic Schrödinger Equations”. In: *Phys. Rev. Lett.* 113 (20 Nov. 2014), p. 200403. DOI: [10.1103/PhysRevLett.113.200403](https://doi.org/10.1103/PhysRevLett.113.200403). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.113.200403>.
- [249] Nikesh S. Dattani, Felix A. Pollock, and David M. Wilkins. “Analytic influence functionals for numerical Feynman integrals in most open quantum systems”. In: *Quant. Phys. Lett.* 1 (2012), p. 35. DOI: <http://dx.doi.org/10.18576/qpl>.
- [250] Ulrich Schollwöck. “The density-matrix renormalization group in the age of matrix product states”. In: *Ann. Phys.* 326.1 (2011), p. 96. ISSN: 0003-4916. DOI: <https://doi.org/10.1016/j.aop.2010.09.012>. URL: <http://www.sciencedirect.com/science/article/pii/S0003491610001752>.
- [251] Joel Yuen-Zhou et al. *Ultrafast Spectroscopy*. IOP, 2014. ISBN: 978-0-750-31062-8. DOI: [10.1088/978-0-750-31062-8](https://doi.org/10.1088/978-0-750-31062-8).
- [252] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions*. New York: Dover Publications, 1964. ISBN: 978-0-486-61272-0.