Bayesian parameter estimation and thermodynamics in equilibrating quantum systems

Dissertation

zur Erlangung des Grades eines Doktors
der Naturwissenschaften

vorgelegt von

Julia Boeyens

eingereicht bei der Naturwissenschaftlich-Technischen Fakultät
der Universität Siegen,
Siegen 2025

Betreuer und erster Gutachter

Jun. Prof. Stefan Nimmrichter
Universität Siegen

Zweiter Gutachter

Dr. Luis Alberto Correa Marichal
Universidad de La Laguna

Tag der mündlichen prüfung
31.01.2025

Abstract

The trend towards miniaturised devices has resulted in systems small enough to be governed by quantum mechanics. The goal of these systems is to perform thermodynamic tasks like refrigeration or work extraction, which requires a description of how they interact with their environment. This has spurred the development of quantum thermodynamics. However, achieving the desired experimental control requires precise knowledge of the system and highly accurate measurements. This thesis explores these two aspects of open quantum systems from a theoretical perspective. First, we explore how Bayesian techniques can be applied to the sensing of environmental parameters of a quantum system to find better estimation protocols, particularly in situations with little data, and where adaptive strategies are allowed. The advantages and drawbacks of various Bayesian estimation approaches are explored. These methods are specifically examined with respect to the role of prior distributions, estimators, and cost functions in achieving accurate estimates. The primary example involves qubit thermometry, where a two-level probe interacts with an environment. This example highlights the effectiveness of Bayesian estimation for small data sets and quantifies the scaling of the accuracy with number of measurements using Bayesian bounds. The sensitivity of probes based on environmental interactions is also analysed. The estimation of rate parameters is studied with particular emphasis on how the specification of the prior information can influence the entire estimation strategy. Next, another example of thermometry is considered using continuously monitored probes. An adaptive strategy is proposed which showcases the benefits of Bayesian estimation. We also study this scenario in the case when the measurement signal includes noise and finite detector bandwidth. Finally, we turn to our second aspect of open quantum systems and study work extraction from an open quantum system. Here, we investigate how collective effects arising from the interaction of permutationally invariant particles with their environment affect work extraction. The analysis includes various models of work extraction, including energy output in steady states, work done against dissipative loads, and power output when coupled to a driving field.

Zusammenfassung

Der Trend zu miniaturisierten Bauteilen hat zu Systemen geführt, die so klein sind, dass sie von quantenmechanischen Effekten dominiert werden. Die Systeme interagieren mit ihrer Umgebung, um thermodynamische Aufgaben wie Kühlung oder Arbeitsentnahme durchzuführen. Dadurch wurde die Entwicklung der Quantenthermodynamik vorangetrieben. Genaue Kenntnisse über das System und hochgenaue Messungen sind erforderlich um die gewünschte experimentelle Kontrolle zu erreichen. Diese beiden Aspekte offener Quantensysteme werden in dieser Arbeit aus einer theoretischen Perspektive untersucht. Zunächst wird untersucht, wie insbesondere in Situationen mit wenigen Datenpunkten und wenn adaptive Strategien erlaubt sind, mit bayessche Techniken bessere Schätzprotokolle für die Bestimmung Umgebungsparameter eines Quantensystems gefunden werden können. Das wichtigste Beispiel ist die Qubitthermometrie, bei der eine Zweiebenenprobe mit einer Umgebung interagiert. Dieses Beispiel verdeutlicht die Wirksamkeit der bayesschen Schätzung für kleine Datensätze und ermittelt unter Verwendung bayesscher Grenzen wie die Genauigkeit mit der Anzahl der Messungen skaliert. Ebenfalls wird die auf Wechselwirkungen mit der Umgebung beruhende Sensitivität von Proben analysiert. Die Schätzung von Ratenparametern wird mit besonderem Augenmerk auf den Einfluss der Spezifikation der Vorinformationen auf die gesamte Schätzstrategie untersucht. Als nächstes wird ein weiteres Beispiel der Thermometrie betrachtet, diesmal unter Verwendung von kontinuierlich überwachten Proben. Eine adaptive Strategie die die Vorteile der bayesschen Schätzung verdeutlicht wird vorgeschlagen. Das Szenario wird weiterhin für den Fall, dass das Messsignal Rauschen und eine begrenzte Detektorbandbreite enthält untersucht. Anschließend wenden wir uns dem zweiten Aspekt zu und untersuchen die Extraktion von Arbeits aus einem offenen Quantensystem. Hier untersuchen wir, wie kollektive Effekte, die sich aus der Wechselwirkung von permutationsinvarianten Teilchen mit ihrer Umgebung ergeben, die Arbeitsextraktion beeinflussen. Die Analyse umfasst verschiedene Modelle der Extraktion von Arbeit, einschließlich der Energieabgabe an stationären Zustände, der Arbeit die gegen dissipative Kräfte geleistet wird, und der Leistungsabgabe bei Kopplung an ein Antriebsfeld.

Contents

Chapter 1

Introduction

Recent technological advances have allowed for the preparation and control of increasingly complex devices on the nanoscale. These systems can be used to perform useful thermodynamic tasks and at the same time, they are small enough that quantum effects arise in their dynamics. They include for example photovoltaic cells [1–4] and ultracold atoms [5, 6]. The description of thermodynamic tasks was developed from the perspective of classical physics for the operation of macroscopic physical systems made up of a huge number of degrees of freedom, like a gas expanding and moving a piston. This was done by reducing the description of the dynamics to a few state variables, like the temperature and pressure of the gas. From the dynamical behaviour of these observables, the well known fundamental laws of thermodynamics are derived. The field of *quantum* thermodynamics [7, 8] is concerned with how these thermodynamical laws apply to microscopic systems which are governed by quantum mechanics but are performing thermodynamic tasks.

The seminal example of one of these systems is the three level laser. Scovil and Schultz-DuBois modelled the three level laser as a heat engine that is able to reach Carnot efficiency [9, 10]. In this example, a three-level atom interacts with a hot and cold bath which generates a population inversion between two levels of the atom. This results in amplified light to be emitted when the atom is coupled to a periodic driving field. Scovil and Schultz Du-Bois showed that for a laser with zero gain, Carnot efficiency could be reached. This result has been extended and discussed theoretically from the perspective of open quantum systems where dynam-

ical operation of the transfer of energy between the environment and system can be taken into account [11, 12]. Subsequently many works on the operation of quantum engines and refrigerators were put forward [13–17] and in particular it was found that the laws of thermodynamics, initially concieved for macroscopic systems may be consistently applied to quantum systems. In particular, they are applicable to even a single particle [18–20]. However, this should not be taken to mean that thermodynamic behaviour is identical on small scales. On the macroscopic scale, fluctuations in heat and work for each single particle become insignificant and the behaviour can generally be described purely by average quantities. This is not true for systems on the scale of a few particles. Here, quantities like work and heat may fluctuate in time [21–25]. Additionally, when considering quantum systems it is not as straightforward to categorise total energetic changes into the categories of heat and work [26]. In a classical engine the working medium converts heat to work, for example, by moving a piston. The piston merely transfers the energy to the load and is unchanged by this process. In a quantum machine, the transfer of energy to the load is typically performed by a change in an external field. This field does experience a change in entropy and this change must be included in the thermodynamical description of the engine to determine how much useful energy will ultimately be transferred to the load and how much will go to heating up the quantum "piston". The study of quantum thermodynamics is not merely conceptual and there have been a number of works on thermal machines that demonstrate when quantum effects can improve their performance theoretically [27–31] as well as experimentally [32–39].

Precise control of an experiment requires accurate measurements of the system parameters. In a classical system, an ideal experiment can estimate any system parameters with arbitrary precision. In quantum systems, even an ideal experiment can only determine the probability that the system is in any particular state and in general, measurements affect the dynamics and change the underlying state in a way that classical measurements do not. However, quantum metrology which harnesses properties unique to quantum systems like entanglement or quantum coherence has also led to the possibility to estimate parameters very precisely. This has led to major experimental breakthroughs like the detection of gravitational waves [40, 41], as well as the development of highly sensitive magnetometers [42, 43] and atomic clocks [44–46].

In many contexts, relevant to thermodynamic devices, accurate knowledge of the temperature of the system is an important prerequisite since the temperature determines much of the

dynamical and steady state behaviour of the system. Quantum thermometry is therefore one of the primary examples of metrology in open quantum systems. It has been investigated both theoretically to determine fundamental bounds on the accuracy [47–52] as well as determining how non-equilibrium probes [53–57] and many-body systems can improve thermometric accuracy [58–62]. The extension of quantum metrological protocols to open systems is still an active field of research. In particular, Bayesian methods for metrology in open systems have received much recent attention [63–65] as well as techniques for improving measurement precision by adapting the measurement strategy during an experiment [47, 66].

In this thesis, we will explore both open system metrology and aspects of quantum thermodynamics. This requires a description in terms of open system dynamics. We will refer to standard textbook methods for this [67, 68]. A general outline of how the interactions with the environment are modelled throughout this thesis and a derivation of the standard master equation is given in App A. In Chapter 2, Bayesian quantum parameter estimation is introduced. This approach starts with specifying the prior information that the experimenter has. Next, we introduce how estimates and error bars for these estimates can be found either by using the specification of the prior information to derive these in a fully systematic way or by more traditional heuristic means. Optimal measurements and Bayesian bounds on the error are also introduced here, along with a way to quantify the parameter range to which a particular experiment is sensitive. Finally, we discuss the drawbacks and advantages of the different approaches to Bayesian estimation.

Chapter 3 puts these methods to the test in three examples where the parameter of interest is encoded into the probe through dissipative interactions with the environment. Here, we explore the crucial role of the choice of prior distribution, estimators and cost functions for the accuracy and computability of Bayesian estimates from the data. The main example is qubit thermometry where a two-level probe is prepared, interacts with an environment and is measured. We will compare and contrast the commonly chosen estimators and priors. Of particular interest in this example is how well the Bayesian estimation strategy captures the actual error for small amounts of data as well as how the scaling of the accuracy can be quantified with Bayesian bounds. Additionally, we study the sensitivity of the probes depending on their interaction with the environment using the sensitivity range derived in the previous chapter. The second and third examples both concern the estimation of a rate parameter. The approach taken for these

examples is to start from a prior distribution and derive all estimators and cost functions from this point. This approach is compared for two different prior distributions and measurement and probe optimisation is explored.

In Chapter 4, the Bayesian techniques are extended to thermometry using a probe that is monitored continuously in time. Here, we study the scaling of the measurement precision with measurement time. We also study adaptive estimation in this chapter which is facilitated by being able to perform a Bayesian update in each step. This along with being able to make statements about the value of a parameter and its associated error in the limit of little data is a major advantage of Bayesian estimation. We also characterise the impact of noise and a finite detector bandwidth on the measurement precision in both adaptive and non-adaptive scenarios. This step brings this example the closest of all parameter estimation examples in this thesis to being applicable in an experimental setting.

We turn to work extraction from an open system in Chapter 5. In particular, the seminal example of a three level Maser as a heat engine is extended to an ensemble of three level systems collectively coupled to the environment. That is, the environment cannot distinguish between particles which leads to permutationally symmetric dynamics. Under certain conditions we see enhanced work output for the collectively coupled ensemble and that work can be extracted outside of the temperature window that would be expected from individually coupled systems. We study three models of work extraction. First, we study the maximum energy that could be extracted from the steady state of the system. Next, we couple the system to a dissipative load. This is analogous to finding how much work can be done against the force of friction in a classical engine. However, similar to the classical case, the energy output may not consist solely of work but also heat. We therefore study how work-like this energy current is. Finally, in an example that most resembles a laser, we calculate the power output of the system when it is coupled to a driving field. This study is facilitated by tailoring tools from representation theory to this system which provides general tools for the further study of these open systems.

Chapter 2

Bayesian parameter estimation in open quantum systems

This chapter will introduce Bayesian metrology in the setting of open quantum systems. In particular, we will begin with a short introduction to the well established field of quantum parameter estimation which leads into a discussion of Bayesian estimation. There is a particular focus on the ways that Bayesian statistics makes transparent the prior assumptions about the data inherent to any estimation technique and how a suitable choice of uninformative prior is the foundation of a fully global approach to Baeysain estimation. This chapter is based on the introductory sections of publications [A] and [D].

2.1 Quantum parameter estimation

In this and the following chapters, rather than exploring the fundamental aspects of quantum measurements, we will be concerned with the estimation of parameters using systems with dynamics described by quantum mechanics. Specifically, we will focus on Bayesian single parameter estimation. To this end, we first define a general set up of a single parameter metrology protocol.

Assume that there is a system which interacts with a probe system according to some dynamical process which depends on an unknown parameter θ. The task is to determine the value

of θ. We will denote the interaction that encodes the parameter into the probe state by the superoperator $\mathcal{L}_\theta(O)$, which is a linear operator that acts on the system operators O. It may for example describe the von Neumann equation of the system under some Hamiltonian evolution. The encoding and subsequent measurement of the probe may be repeated multiple times before analysing the data and producing an estimate ϑ of the parameter. In this thesis we will explore each of these points from the perspective of a global, Bayesian framework.

First we recall some concepts from the usual approach to parameter estimation in quantum systems, what we will refer to as local estimation[1]. Here, it is implicitly assumed that the value of the parameter falls within a small range. This corresponds to the situation where we have already performed many measurements or the experiment is already very well calibrated. In the limit of many measurements, one further measurement will only change the estimate by a very small amount and so we should optimise the probe state and measurements to be sensitive to values in a small neighbourhood of θ. For a given measurement, the probability to get a particular outcome given a fixed underlying value of the parameter θ and some data that we observe x, which could be a collection of data points, is given by the likelihood function $P(x|\theta)$. The Fisher information [69, 70],

$$\mathcal{I}(\theta) = \int \mathrm{d}x P(x|\theta) \left[\frac{\partial \log(P(x|\theta))}{\partial \theta}\right]^2, \tag{2.1}$$

captures how well the experiment is able to distinguish two values of θ that are very close together on average over all possible realisations of the experiment. A large value of the Fisher information at the value of the parameter θ means that the experiment is very sensitive in a small neighbourhood of θ.

Now assume that we are able to make an unbiased estimate of the value of the parameter. This means that if the actual value of the parameter is θ^* and we estimate the value to be ϑ_x for a certain data set, ϑ_x is unbiased when it satisfies

$$\int \mathrm{d}x P(x|\theta^*)\vartheta_x = \theta^*. \tag{2.2}$$

[1]Note this is not the same as a local *measurement*. Say the probe is a composite system made up of multiple subsystems, a local measurement acts on each subsystem independently in contrast to a collective or non-local measurement which will be able to access the entire Hilbert space of the probe.

One commonly chosen estimator is the maximum likelihood estimator [71]. It is defined as the value of the parameter that maximises the likelihood function, $\vartheta_x^{(\mathrm{ML})} = \arg\max_\theta [P(x|\theta)]$. Generally, this estimator is not unbiased but it becomes unbiased in the limit of many measurements.

If an estimator is unbiased, we can relate the variance $\Delta\theta$ of the data, to the Fisher information through the Cramér-Rao bound (CRB) [72],

$$\int \mathrm{d}x P(x|\theta)\left[\vartheta-\theta\right]^2 \geq \frac{1}{\mathcal{I}(\theta)}. \tag{2.3}$$

We can therefore, minimise the spread of the data by maximising the Fisher information. This is done by designing the experiment in a way that changes the likelihood function favourably. In a quantum experiment the likelihood depends on the measurement and the state through the Born rule $p(x|\theta) = \mathrm{Tr}(\rho(\theta)M(x))$ where $M(x)$ is a measurement operator, which is an element of a POVM (positive operator valued measure) [2].

Therefore, when designing an optimal experiment for a particular probe state, the optimal POVM must be chosen. In the limit of a large number of measurements, the best POVM to choose will be one that is able to distinguish very small differences in the value of the parameter. Since we are, in principle, able to choose the POVM freely we can then maximise the Fisher information over all POVMs, which gives the quantum Fisher information. The measurement that satisfies this is a projective measurement onto the eigenbasis of the symmetric logarithmic derivative [3] which is defined implicitly by the equation,

$$\partial_\theta \rho(\theta) = \frac{1}{2}(L\rho(\theta) + \rho(\theta)L). \tag{2.4}$$

The quantum Cramér-Rao bound [73] is then given by,

$$\Delta\theta \geq \frac{1}{\sqrt{\mathcal{I}_Q(\theta)}}. \tag{2.5}$$

[2] In the simple cases that we will consider, the Hilbert space has finite dimension n and the POVM is a set of positive operators $M_i(\theta)$ with the property, $\sum_{i=1}^n M_i = \mathbb{1}$. A measurement of the probe state $\rho(\theta)$ with outcome x_i, is described by the Born rule $P(x_i|\theta) = \mathrm{tr}\{M_i\rho(\theta)\}$.

[3] E.g. if $|\psi\rangle$ is an eigenvector of L, the symmetric logarithmic derivative (SLD), $\mathrm{tr}\{\rho\,|\psi\rangle\langle\psi|\}$ would be an optimal measurement.

Here,

$$\mathcal{I}_Q(\theta) = \mathrm{tr}\{\rho(\theta)L^2\}, \tag{2.6}$$

is the quantum Fisher information (QFI). The methodology of quantum parameter estimation was largely developed around the problem of phase estimation [74, 75] and so we will briefly discuss this as the paradigmatic example before discussing thermometry as the most common example of parameter estimation in open quantum systems. In quantum phase estimation, the unknown system parameter θ is a phase encoded into the probe state by the unitary evolution under Hamiltonian H,

$$\rho(\theta) = e^{-i\theta H}\rho e^{i\theta H}. \tag{2.7}$$

We can distinguish two different protocols of probe preparation that lead to different scaling of the error (variance) in estimating θ in the asymptotic limit [74]. In the first scenario, N probes, that are identically prepared, interact with the environment and are then measured. This can also be achieved with a single probe that is reset to the initial state after each of the N measurements.

The QFI can then be used to determine what the optimal measurement would be. In this case, the QFI turns out to be independent of θ [50, 73]. Therefore, the SLD, which gives the optimal POVM is independent of the value of θ and depends only on the initial state and the operator H. The initial probe state that gives the optimal QFI is the pure state $\rho_0 = |\psi\rangle\langle\psi|$, with $|\psi\rangle = 1/\sqrt{2}(|h_0\rangle + |h_n\rangle)$, where $|h_0\rangle$ and $|h_n\rangle$ are the eigenvectors of H with minimum and maximum eigenvalues respectively [74]. In this scenario, each measurement is an identical independent trial on a single photon probe and therefore, the best scaling of the variance that can be achieved is $\Delta\theta \geq 1/\sqrt{\mathcal{I}(\theta)N}$ known as the standard quantum limit.

Alternatively, the N probes might be prepared in m lots of n probes that are in an entangled initial state. The probes then interact with the environment and a collective measurement is performed on the n probes. This is repeated m times. In this case, the best possible scaling of the variance is $\Delta\theta \geq 1/n\sqrt{m}$ which is called Heisenberg scaling and this is in fact the best possible scaling of the CRB for quantum systems [76, 77]. This is achieved by preparing the n probes in the state $|\psi\rangle = 1/\sqrt{2}(|h_0\rangle_1 \ldots |h_0\rangle_n + |h_n\rangle_1 \ldots |h_n\rangle_n)$, which, in interferometry, is the famous N00N state [78].

This state is however, very difficult to actually use in an experiment since the effects of noise can make the state unable to acquire any phase information [79]. An adaptive approach can

also be used where the basis of the measurements or other control parameters in the experiment are changed during the measurements of the N probes [80–86]. Adaptive estimation is usually facilitated through a Bayesian estimation strategy.

The evolution of the field of open system parameter estimation largely mirrors that of quantum sensing under closed system or Hamiltonian dynamics (of which phase estimation is one example but estimation of a magnetic field strength (magnetrometry) is another notable one [87, 88]). In open system parameter estimation, thermometry is the primary theoretical example [50, 89], although other quantities like decay rates may also be estimated. The main difference from parameter estimation under closed system dynamics is that the state evolution is not merely a function of the Hamiltonian. This leads to a density matrix that depends on the parameter in a non-trivial way. Initially, the main goal of this field was to extend the results from phase estimation to achieve a scaling of the variance that beats the standard quantum limit and to even achieve Heisenberg scaling in thermometry examples.

The simplest example we can consider is equilibrium thermometry, which is inspired by temperature estimation in classical systems. Here, the probe system interacts with an environment which has an unknown temperature and is allowed to come to equilibrium with the environment. This is similar to taking the temperature of a classical system where a small probe (the thermometer) is allowed to come to thermal equilibrium with an environment. If we assume the quantum probe has already reached equilibrium with the environment, its state is a Gibbs state [50, 89],

$$\rho = \frac{e^{-H/(k_B\theta)}}{\mathrm{tr}[e^{-H/(k_B\theta)}]}, \tag{2.8}$$

where H is the Hamiltonian of the probe and k_B is the Boltzmann constant. Now, if we follow the same steps as with local phase estimation, we first find the QFI. Here it is dependent on the heat capacity $\mathcal{C}(\theta) = (\langle H^2\rangle - \langle H\rangle^2)/k_B\theta^2$ of the probe, $\mathcal{I}_Q(\theta) = \mathcal{C}(\theta)/k_B\theta^2$, noting that in contrast to phase estimation it depends on θ [50]. Thus, any optimisation of the experiment based on the QFI will only be optimal for one specific temperature. This clearly shows a disadvantage of the local approach to parameter estimation that is not apparent in phase estimation. That is, if this approach is used to design an experiment an assumption has to be made about the expected value of the temperature before any data has been collected and the temperature is essentially totally unknown. Assuming that we are justified in using this approach,

say for example some data is already collected and the temperature is expected to lie in a small range, we can determine the optimal POVM and optimal probe according to the QFI. From the expression for the QFI, we see that the optimal precision is obtained by performing a projective measurement in the energy basis (eigenbasis of the Hamiltonian) of the state ρ. Additionally, the optimal probe is one with the largest heat capacity, which for an n dimensional system corresponds an effective two-level system where there is a ground state and an $n-1$ degenerate excited state [49]. Asymptotically, estimation using thermalised probes will only be able to reach the standard quantum limit, $\Delta\theta \geq 1/\sqrt{N\mathcal{I}(\theta)}$ since thermalisation will destroy any superposition that may have been initially prepared between probes and therefore all measurements are independent. Additionally, for any system with a finite energy gap, as the system temperature approaches zero, so does the heat capacity. This means that the uncertainty in the temperature diverges exponentially near zero temperature since the probe is far more likely to be measured in the ground state in every measurement even if the temperature is slightly above absolute zero [48, 52, 53, 90].

One way that quantum thermometry can achieve better scaling is through the use of non-equilibrium probes [54, 91, 92]. Additionally, an ancillary system can be included between the environment and the probe. This can be used to encode information about temperature into coherences on the probe state [93] and may also be combined with non-local measurements on multiple probes, the scaling of the error in temperature can be improved beyond the standard quantum limit [94, 95]. The experiment can also be designed from a globally optimal point of view. Here, we do not assume that the temperature of the environment is known to fall within a small range or try to achieve and asymptotically optimal scaling. Rather, we assume that the temperature it is totally unknown and could in principle take any value in $(0, \infty)$. It was found in [96], that as the range of possible temperatures is increased, the number of effective levels accessible to the globally optimal probe must increase. The idea of addressing the problem of thermometry from a global point of view was introduced in [64]. Taking a global approach to thermometry means that the strategy must be valid even in the limit of very little data. In this case, Bayesian estimation offers a framework for working with very small amounts of data. There have subsequently been many works (including publications [A] and [D]) that address quantum thermometry in a Bayesian way [63–65, 97]. Particularly because it admits an efficient way of performing adaptive thermometry [47, 66, 98].

2.2 The Bayesian framework

For the rest of this chapter we will be setting up the mathematical preliminaries for a global approach to open system parameter estimation. In the chapters that follow this will be applied to some case studies. Here, we do not require any assumptions about the range of values that the parameter can take other than those that are physically motivated, for example the temperature can only take on positive values. This approach is facilitated by the Bayesian framework [71, 99] for parameter estimation where, one defines the knowledge of the parameter θ after some measurements via a posterior distribution,

$$P(\theta|x) = \frac{P(x|\theta)P(\theta)}{P(x)}, \quad P(x) = \int_a^b d\theta P(x|\theta)P(\theta). \tag{2.9}$$

Here, the measurements have outcomes x and the parameter is known to be in the range (a, b). Each outcome of the experiment is found with probability $P(x|\theta)$ known as the likelihood which depends on the value of the parameter. Before any measurements are done, the experimenter must define their prior distribution $P(\theta)$, which quantifies all prior knowledge about the probability for the parameter to take any of the values within the range (a, b). After a measurement is made, the posterior distribution is calculated and this distribution is the new prior distribution for the next round of measurement. Over the course of more measurements, the Bernstein von Mises theorem [100] ensures that for an experiment with a likelihood that admits a non-singular Fisher information matrix, the Posterior distribution will be a sharply peaked Gaussian distribution about the actual value of the parameter and with variance equal to the Fisher information.

We contrast this briefly, to another common approach to parameter estimation known as frequentist estimation [71, 99, 101]. In this approach, one assumes that the parameter has an underlying true value and is not a random variable. However, the experimenter only has access to a sample of the full set of data which would be given if infinitely many measurements could be made. Parameters are estimated from the data by finding the value of the parameter that maximises the likelihood function. This will coincide with the actual value of the parameter once enough measurements have been taken. Frequentist estimation does not require prior informa-

tion to be specified and this was historically, a major sticking point for the acceptance of Bayesian methods by proponents of the frequentist method [99, 102]. In the asymptotic limit, the frequentist (local) and Bayesian approaches to parameter estimation converge since they will both yield an estimate equal to the actual value.

A feature of Bayesian estimation is that adaptive protocols are easily described by the iterative nature of a Bayesian update. In an adaptive protocol, the likelihood is changed during the measurement protocol. This is included in a Bayesian protocol as follows. Say the current posterior distribution after m identical measurements of the system is $P(\theta|x_1, x_2, ...x_m)$. After these measurements it is determined that the experiment should be adapted, either by changing the initial state of the probe, the way the parameter is encoded on the state or the measurement that is made. The adapted measurement is then performed which yields outcome x_{m+1}, the probability of this outcome given the new system is captured by the likelihood $\tilde{P}(x_{m+1}|\theta)$. The new posterior of the system is then simply, $P(\theta|x_1, x_2, ..., x_m, x_{m+1}) \propto P(\theta|x_1, x_2, ...x_m)\tilde{P}(x_{m+1}|\theta)$.

2.3 Uninformative prior distributions

Bayesian parameter estimation requires that any prior knowledge that an experimenter has before measurements are made is included in the prior distribution. Usually no particular outcome of the experiment is preferred and thus a so called uninformative prior is sought after. For example given an experiment with outcomes x that depend on the parameter θ, one may assume that they are totally ignorant to the value of the parameter and therefore choose a prior in which all outcomes are equally likely, that is $P(\theta) \sim c$, this is known as the the *principle of insufficient reason* [71]. However, a naive interpretation of this argument may lead to a prior that does not describe the behaviour of the parameter with respect to other parameters in the problem. In this section we will discuss ways to derive uninformative priors from a few different perspectives. First, we will discuss a situation where no assumptions are made about the experimental set up and therefore the likelihood is not known. Here, we derive a prior based on physical invariances of the parameter. Next, we will discuss a situation where the form of the likelihood function is known and derive a prior based on the ways invariances enter into it. Alternatively,

in addition to knowing the likelihood, one may insist that all parametrisations of the problem should be equally valid and construct a prior that respects this. In this approach the prior is fully specified. Several other approaches to constructing uninformative priors exist [103] that rely on arguments about defining a reference prior or state of maximum ignorance. Most notably are priors derived from the maximum entropy principle which will also be discussed in this section.

2.3.1 Prior densities based on physical transformation invariance

In the first case, we choose a prior by making the most minimal assumptions. There is no assumption about the form of the likelihood function but only intrinsic symmetries that are contained in the physics of the problem. Consider for example the estimation of the position of a classical particle. If we are truly uninformed about where the particle is, all positions in the range $-\infty < \theta < \infty$ should be equally likely and a constant prior probability distribution $P(\theta) \sim c$ is valid[4]. This is an example of a location invariant parameter [71]. These are parameters where the prior probability is invariant under translations of the parameter, that is $P(\theta + c) = P(\theta)$. Not all parameters have this type of invariance. Consider the thermometry problem from Sec. 2.1 where the system is in a Gibbs state Eq.(2.8). Here, the temperature θ of the system is scaled by its energy and thus only enters into the problem through the dimensionless Boltzmann ratio $E/k_B\theta$. To see this, note that if the temperature is changed by an order of magnitude but this ratio is held fixed, the system will behave in the same way. Therefore, it is not the absolute magnitude of θ that matters but rather its scale. A prior distribution that takes this invariance into account will have to have the property, $\int P(\theta)\mathrm{d}\theta = \int P(\gamma\theta)\mathrm{d}(\gamma\theta)$. Therefore, we get from a transformation of variables [71], that $P(\theta) = P(\gamma\theta)\partial_\theta(\gamma\theta)$ which yields $P(\theta) \propto 1/\theta$ for a prior that respects the scale invariance of the problem [64, 71, 104–107]. The examples in Chapters 3 and 4 deal with parameters with this type of invariance. Other types of parameter invariance are of course also possible. Consider the phase estimation problem from Sec. 2.1. Here, the parameter can only take values $0 \leq \theta \leq 2\pi$ and is therefore circularly symmetric. If we want a uniform prior over the angles in a circle, we would choose the prior $p(\theta) = 1/2\pi$ [108], which is still a constant but this time the range is restricted by the problem.

[4]Note that even though this prior is improper (not normalisable), it is still possible to formulate the problem correctly as long as the posterior distribution is normalised.

2.3.2 Prior densities based on transformation invariance of the likelihood

In many parameter estimation problems, the form of the likelihood function is known based on the way the parameter is measured. For example, if the measurement we perform has two possible outcomes: one (which we will call outcome a) with probability to occur q and the other with probability $1-q$ and we measure n times, we can assume that the likelihood function will be a binomial distribution. The probability to observe event X, which is outcome a, m times is given by, $P(X) = \binom{n}{m} q^m (1-q)^{(n-m)}$. The parameter we want to determine may simply be the success probability q. However, we may also want to determine an underlying parameter θ on which the success probability depends non-trivially without specifying this non-trivial dependence. Importantly, in this section we are making an assumption about the way that the data will be distributed which we did not make previously.

Assume that a given likelihood $P(x|\theta)$, which is a continuous function of both the events x and hypotheses θ is invariant under a set of group transformations of a group $\mathcal{G}$ [109],

$$P(x|\theta)dx = P(g_\gamma(x)|g_\gamma(\theta))dg_\gamma(x).$$

For example, g may perform a scale transformation so that $g_\gamma(x) = \gamma x$ corresponds simply to multiplication by a constant γ. Additionally, for any pair of hypotheses θ and θ' assume there is a unique group element γ such that $\theta = g_\gamma(\theta')$. Now following the derivation in [110], we will define a prior that respects this invariance. First, define a Bayesian confidence interval. That is, the volume over an area $A(x)$ in the space of hypotheses θ given some data x and Posterior distribution $P(\theta|x)$, which yields,

$$C = \int_{A(x)} d\theta P(\theta|x). \tag{2.10}$$

This means with probability C the hypothesis lies within area A given the observed data. Then if, for a given volume, the area is invariant under action of the group, $A(g_\gamma(x)) = g_\gamma(A(x))$, there is a prior density that will preserve the invariance property that was identified by the set of

transformations. It can be written down as

$$P(\theta) \propto \left[\left. \frac{\partial(g_\gamma(\theta))}{\partial\gamma} \right|_{\gamma=u} \right]^{-1} \tag{2.11}$$

where u is the identity element of the group. This prior is uninformative since the only assumption that was made is that it must correctly reflect the symmetries of the likelihood function [111].

In cases where there is no obvious symmetry or the symmetry group is highly complicated a different approach to deriving an uninformative prior given a known likelihood may be desired. This is the approach taken in the next section.

2.3.3 Prior densities based on information geometry

The space of probability distributions defined by the likelihood $P(x|\theta)$ to observe hypothesis θ given data x is a Riemannian manifold with metric [72],

$$\mathcal{D}(\theta_1, \theta_2) = \left| \int_{\theta_1}^{\theta_2} d\theta g(\theta) \right|, \tag{2.12}$$

which defines how different two distributions are from each other. The quantity $d\theta g(\theta)$ is an integration measure on the probability space. If we choose the metric that gives the geodesic between two hypotheses in the probability space, we get the Fisher information metric, $g(\theta)d\theta = \sqrt{\mathcal{I}(\theta)}d\theta$, where

$$\mathcal{I}(\theta) = \int dx\, p(x|\theta) \left[\frac{\partial}{\partial\theta} \log(p(x|\theta)) \right]^2 \tag{2.13}$$

is the Fisher information. Using the change of variables formula for probability distributions, the Fisher information transforms under smooth reparameterisation as $\mathcal{I}(\theta) = \mathcal{I}(\eta)(\partial\eta/\partial\theta)^2$ [112]. Therefore, the measure is invariant under changes in parameterisation, that is, $d\theta g(\theta) = d\eta(\theta) g(\eta(\theta))$. If it is possible to define the group G of all transformations that leave the likelihood invariant as defined in the previous section, this integration measure will be invariant under (left) action by G [109]. With this metric defined, the volume element of the manifold is $p(\theta) = \sqrt{\mathcal{I}(\theta)}$, which is parameterisation invariant and invariant under left group action by G therefore it is a

uniform distribution on the manifold of probability distributions. Thus, $p(\theta)$ is a valid choice as an uninformative prior according to the principle of insufficient reason. The choice of prior using the principle $p(\theta) \propto \sqrt{\mathcal{I}(\theta)}$ is also known as *Jeffrey's rule*.

These concepts from classical information geometry can be extended to quantum parameter estimation [73, 113]In quantum mechanics, the probability space corresponds directly to the space of general states defined by density operators. The likelihood is given by $p(x|\theta) = \mathrm{Tr}(\rho(\theta)M(x))$, where $M(x)$ is a positive operator valued measure (POVM) and $\rho(\theta)$ is the state. If we want to define a metric on the probability space, the metric will explicitly depend on the measurement that is made. A natural choice would be the POVMs that maximise the Fisher information, since these will be the measurements most sensitive to the parameter. As we mentioned in Sec. 2.1, the optimal measurements are projective measurments onto the eigenbasis of the symmetric logarithmic derivative (SLD) $L(\theta)$ [114, 115],

$$L(\theta)\rho(\theta) + \rho(\theta)L(\theta) = 2\partial_\theta \rho(\theta) \tag{2.14}$$

which yields the quantum Fisher information for a single parameter

$$\mathcal{I}_Q(\theta) = \mathrm{Tr}(\rho(\theta)L(\theta)^2) \tag{2.15}$$

The metric with volume element $p(\theta)d\theta = \sqrt{\mathcal{I}_Q(\theta)}d\theta$ is the quantum Fisher information metric [63, 116] and we can use this as the uninformative prior in a quantum problem.

2.3.4 Prior densities based on maximising the entropy

Finally, the uninformativeness of a prior may be decided not by insisting that the distribution be uniform in some sense but rather that it assumes the least amount of information about the hypothesis [106,107]. For a distribution that takes discrete values this concept of ignorance is quantified by the Shannon entropy [117] $S = -\sum_{i=1}^{N} P(\theta_i)\log(P(\theta_i))$ where the distribution $P(\theta)$ with maximum entropy is the most ignorant to the unknown parameter [103, 107, 118]. The entropy is maximised with respect to constraints known as testable prior information. Any constraint that the prior distribution can be verified to obey is testable information, for example that a distribution has a particular mean or variance. Commonly, these constraints consist of m

constraints on the moments of the prior distribution $\langle\theta\rangle = m_1 \ldots \langle\theta^n\rangle = m_n$ of the distribution. If there is no such prior information then the distribution the maximises the entropy is uniform.

For continuous parameters, the entropy is defined by an integral and therefore, must be specified with respect to a base probability density μ on the probability space and prior density $p(\theta)$ to yield, $S = -\int d\mu\ p(\theta)\log(p(\theta)/\mu(\theta))$ [71]. Maximising the entropy is then equivalent to minimising the Kullback-Leibler divergence. This then yields the same prior as the previous section [119]. That is, $p(\theta)\mathrm{d}\theta \propto \sqrt{\mathcal{I}(\theta)}\mathrm{d}\theta$.

To see this, first define the Kullback-Leibler divergence[5] between two densities $q(\theta)$ and $p(\theta)$,

$$D[p(\theta), q(\theta)] = \int d\mu\ p(\theta)\log\left(\frac{p(\theta)}{q(\theta)}\right), \tag{2.16}$$

with respect to the base measure μ. Then for two densities parameterised by the parameters θ and $\theta + d\theta$ the Taylor expansion of $D[p(\theta), p(\theta + d\theta)]$ taken to second order gives,

$$D[p(\theta), p(\theta + d\theta)] = \frac{1}{2}\int d\mu\ p(\theta)\frac{d^2}{d\theta^2}\log(p(\theta)), \tag{2.17}$$

which is proportional to $\mathcal{I}(\theta)$ when we use a uniform measure. On the surface, this approach to obtaining an uninformative prior may seem the most attractive, since it is the only one that directly addresses the amount of assumed information. However, it has a number of drawbacks [102, 103, 119] with the most obvious drawback being the selection of the measure μ. Thus, the choice of how to select the prior has merely been deferred to choosing the measure. One could choose a uniform or Lebesgue measure [71, 103] or a measure that depends on the invariances of the problem [107].

2.3.5 Consequences of the choice of prior density

We have now seen four ways that the prior can be constructed in an uninformed way. However, in the cases where these approaches lead to different priors it is not clear which prior is preferable. In cases where the symmetries of the problem are not easily identified and the calculation

[5] $D[p(\theta_1), p(\theta_2)] \neq D[p(\theta_2), p(\theta_1)]$

of the Fisher information is not too complicated the prior derived from the Fisher information is preferred [71, 103, 120]. One property of this prior is that the maximum probability occurs where the likelihood is most sensitive to a change in the value of θ. This amounts to assuming that the most likely outcome of an experiment is the value of the parameter where the apparatus is most sensitive. For this to be true, enough data must already have been obtained so that, the assumption $\theta_{\max}/\theta_{\min} \sim 1$ is justified. In parameter estimation of open systems, the interpretation of the quantum Fisher information metric as a natural metric on the state space is potentially problematic when the parameter that must be estimated is an environmental parameter. The state space of the environment may be very complex or completely unknown to the experimenter.

Priors derived from underlying symmetries in the parameter [104] or likelihood [71, 109, 110] are often simpler and avoid the assumption that the true value lies at the most sensitive point of the apparatus and also do not require knowledge of the state space. This may lead to simpler calculations of point estimators and cost functions as is detailed in the next sections. Additionally, the initial state of the system can be optimised without affecting the prior. This facilitates a fully global form of quantum parameter estimation.

The construction of uninformative prior distributions can lead to priors that are not normalisable. For example both the constant prior $P(\theta) \propto c$ for $\theta \in (-\infty, \infty)$ as well as the prior $P(\theta) \propto 1/\theta$ for $\theta \in (0, \infty)$ are improper priors. Improper priors should be avoided when possible because they can lead to unnormalised posterior distributions. This problem can be avoided in practice by restriciting the domain for which the prior has non-zero values [71]. The prior can then be normalised over this restricted domain.

2.4 Cost functions and point estimators

The probability to obtain the possible values of the parameter are completely captured by the posterior distribution once the likelihood and prior are given. However, it is often more useful to have a point estimate of the parameter along with an error estimate that describes the spread of the posterior distribution which is defined by the average of a cost function over the posterior distribution. These cost functions can also be used to define error bars which capture the uncer-

tainty in the estimate. This is an essential aspect of any experiment and it is a particular advantage of the Bayesian approach that error bars can be defined in a consistent way from the posterior distribution. In this thesis we will discuss two ways that one can arrive at the cost functions underlying the error estimates and point estimators. The first (Sec. 2.4.1, Sec.2.4.2 and Sec. 2.4.1 and [A]) is using commonly chosen figures of merit like the mean, median and mode of the distribution to arrive at estimators and cost functions. The second approach (sec 2.4.3 and [D]) is to determine the cost function and estimator from the specified prior distribution. This results in a streamlined formalism that also includes probe and measurement optimisation based on the prior.

2.4.1 Commonly chosen cost functions

The optimal Bayesian point estimator is found by minimising a particular cost function on average over the full range of the prior. This leaves the choice of cost function open. A familiar example of a cost function is the square distance between two hypotheses, θ_1 and θ_2 ,

$$c^{(2)}(\theta_1, \theta_2) = (\theta_1 - \theta_2)^2, \tag{2.18}$$

which allows us to quantify the spread of the distribution with the variance with respect to an estimate of the parameter $\vartheta_x^{(2)}$,

$$\varepsilon^{(2)}(\vartheta_x) = \int d\theta P(\theta|x) c^{(2)}(\vartheta_x, \theta) = \langle c^{(2)}(\vartheta_x, \theta) \rangle. \tag{2.19}$$

By minimising this average loss, following the steps shown in App. B, we get the mean of the posterior as the optimal estimator $\vartheta_x^{(2)} = \langle \theta \rangle$.

The cost function could however be specified in a variety of other ways. Commonly this function is chosen from a few standard choices which each have their own optimal estimator associated with them [71]:

- The square distance $c^{(2)}(\vartheta_x, \theta)$ with estimator $\vartheta_x^{(2)} = \langle \theta \rangle$.
- The absolute distance $c^{(1)}(\vartheta_x, \theta) = |\vartheta_x - \theta|$, which gives an optimal estimator $\vartheta_x^{(1)}$

equivalent to the median of the posterior distribution which is defined by,

$$\int_0^{\vartheta_x^{(1)}} \mathrm{d}T\, P(T|n) = \int_{\vartheta_x^{(1)}}^{\infty} \mathrm{d}T\, P(T|n) \stackrel{!}{=} \frac{1}{2}. \tag{2.20}$$

This estimator is invariant under any re-parametrization of θ.

- Finally, the box-error or 1-0 loss, $c^{(\mathrm{ml})}(\theta_x, T) = -\delta(\vartheta_x - \theta)$ which penalises all deviations by the same amount. The estimator associated with this cost function, $\boldsymbol{\vartheta}^{(\mathrm{md})}$ is the mode of the posterior, $\vartheta_x^{(\mathrm{md})} = \arg\max_\theta P(\theta|x)$ which for a constant prior gives the maximum likelihood estimator. This estimator is also invariant under any re-parametrization of θ.

Note that one requires the posterior distribution to have finite first and second moments. For an improper prior, the square and absolute cost functions defined above may diverge and thus, care must be taken to properly normalise the posterior distribution. As we will show in the qubit thermometry example in the next chapter, when the parameter we want to estimate is not a location parameter but rather has a different form of symmetry, the standard cost functions defined above may not capture the actual uncertainty correctly. These cost functions depend on the absolute scale of the parameter compared to the estimate. In cases where we want to compare different scales one needs a relative error measure. Therefore these cost functions can be modified to accommodated this. For example, the square and absolute distances can be modified to relative deviations. In this way, the small absolute deviations in the parameter are penalised more strongly when the estimated value of the parameter is small rather than large. For example, we can modify the square and absolute deviations to give,

- $c^{(2\mathrm{r})}(\vartheta_x, \theta) = (\vartheta_x/\theta - 1)^2$, which results in the relative mean estimator $\vartheta_x^{(2\mathrm{r})} = \langle\theta^{-1}\rangle/\langle\theta^{-2}\rangle$. This error captures relative deviations and will also not diverge in the case that the parameter range is unrestricted.

- $c^{(1\mathrm{r})}(\vartheta_x, \theta) = |\vartheta_x/\theta - 1|$, which gives the estimator $\vartheta_n^{(1\mathrm{r})}$ given by the median of the re-normalized posterior distributions $\propto P(\theta|n)/\theta$.

This is the approach that was used to derive a relative cost function in [A]. The calculation of the estimators by minimising the posterior loss is shown in App B. In general, the cost function

leads to a error estimate like Eq.(2.19),

$$\varepsilon^{(f)}(\vartheta_x^{(f)}) = \int d\theta P(\theta|x) c^{(f)}(\vartheta_x^{(f)}, \theta) = \langle c^{(f)}(\vartheta_x^{(f)}, \theta) \rangle, \tag{2.21}$$

which quantifies the spread of the estimator with respect to the posterior distribution.

2.4.2 Credibility regions

Another way to define an error bar associated with a particular estimate defined in the range $[a, b]$ is by defining the credibility region around the point estimate [121]. Here we see that in a Bayesian estimation scenario, no point estimator needs to be specified. The posterior distribution can be summarised by the credibility region.

This is an error estimate that uses the cumulative posterior probability

$$\int_a^{\theta_x^{Y\%}} d\theta P(\theta|x) = Y\% \tag{2.22}$$

such that with, for example 90%, probability the estimate lies between the two points $\theta_x^{5\%} \lesssim \vartheta_x \lesssim \theta_x^{95\%}$. This error estimate is independent of any particular point estimate or parametrisation. It is also able to capture the skewness of the distribution about any particular point estimate as the points $\theta_x^{5\%}$ and $\theta_x^{95\%}$ don't have to be equidistant from ϑ.

2.4.3 Cost functions from underlying symmetry of the parameter

Although it is common practice to choose the cost function heuristically, when the parameter we want to estimate is isomorphic to a location parameter, the prior distribution can be used to motivate the choice of cost function [104].

First, let Φ be a location parameter. This means that if we have no information about the parameter the prior will be a flat density, $P(\Phi) \propto 1$ on the entire range $(-\infty, \infty)$. Now, say we are interested in estimating the parameter Θ. If there exists a function $f : \Theta \to f(\Theta)$ such that $f(\Theta) = \Phi$, Θ is a location isomorphic parameter.

We then obtain the prior with respect to the parameter θ using the change of variables rule

for a statistical parameter [71] and the condition for an ignorant prior for a location parameter $P(f(\theta)) \propto 1$,

$$P(\theta) = P(\Phi)\frac{\mathrm{d}f(\theta)}{\mathrm{d}\theta}. \tag{2.23}$$

This yields,

$$P(\theta) = \frac{\mathrm{d}f(\theta)}{\mathrm{d}\theta}. \tag{2.24}$$

The prior measures the hypothesis space and so we can define a distance in the hypothesis space with

$$\left|\int_{\theta'}^{\theta} \mathrm{d}z P(z)\right| = |f(\theta) - f(\theta')|. \tag{2.25}$$

The cost function $c^{(f)}(\vartheta_x, \theta) = (f(\vartheta_x) - f(\theta))^2$, respects the location invariance of $f(\theta)$ and we have chosen a square distance to aid in optimisation of the estimator and later the measurement operator. The optimal Bayesian point estimator given this cost function is

$$\vartheta_x^{(\mathrm{f})} = f^{-1}\left[\int d\theta P(\theta|x) f(\theta)\right]. \tag{2.26}$$

An estimator and cost function derived from the prior in this way will automatically respect the symmetries of the problem. However, they are not designed to avoid diverging moments of the distribution and for them to be calculated one may need to restrict the range of the parameter one integrates over. This is something that must be considered in particular if one has an improper prior distribution.

This approach can also be applied if the prior is determined from the Fisher information since this prior is parameterisation invariant and respects all symmetries of the problem. Furthermore, the metric structure that this prior gives to the hypothesis space means that $\int \mathrm{d}\theta \sqrt{\mathcal{I}_Q(\theta)}$ is a location parameter. Therefore, we get the cost function,

$$c^{(I)}(\vartheta_x, \theta) = \left(\int_{\theta'}^{\theta} \mathrm{d}z \sqrt{\mathcal{I}_Q(z)}\right)^2 = (f(\theta) - f(\theta'))^2, \tag{2.27}$$

where the function f is now defined implicitly through the integral of the prior. Unfortunately,

in many cases the cost function $c^{(I)}(\vartheta_x, \theta) = (f(\vartheta_x) - f(\theta))^2$, and optimal estimator $\vartheta_x^{(\mathrm{I})}$ may be tedious to calculate in practice when the integral in Eq (2.27) has to be computed numerically.

2.4.4 Rate estimation in exponential distributions

We end this section with a short example that demonstrates the technique of determining an invariance in the problem, finding a prior and thus a cost function and estimator.

Consider an experiment where we detect events randomly in a particular time frame. In this example, each event is independent and we are interested in estimating the waiting time between events. This type of process, known as a Poisson process, is ubiquitous in physics . For example it is associated with the arrival of a photon at a detector after spontaneous emission [68, 122] an example that we will explore in more detail in sec. 3.2 or as we will study in Chapter 4, with the transition rate between the levels of a two-level probe in thermal contact with an environment.

In this first simplified example, we will assume that we perform N runs of an experiment with likelihood to measure an outcome at time t given by $P(t|\theta) = \exp(-t/\theta)/\theta$. We note here, that the rate parameter $\theta > 0$ is a scale invariant parameter since its absolute magnitude is not what defines its size but rather how it is scaled by the value of the waiting time t. That is, for a fixed time t, when $t/\theta > 1$, θ is considered to be large and it is considered to be small if $t/\theta < 1$. We can define the location isomorphic function that respects this invariance $f(\theta) = \log(\theta)$. Next, we use this function to define the prior, $P(\theta) \propto \partial_\theta \log(\theta) = 1/\theta$.

Alternatively, we could use the information geometry approach to determine the prior $P(\theta) \propto \sqrt{\mathcal{I}(\theta)}$ with the Fisher information of the likelihood, $\mathcal{I}(\theta) = 1/\theta^2$, which yields the prior, $P(\theta) \propto 1/\theta$. Hence, the prior is the same in both cases. This is an improper prior and so we will restrict to the range $\theta \in [a, 1/a]$. The optimal error and point estimator based on the cost function $c^{(2\mathrm{l})}(\theta_1, \theta_2) = \log^2(\theta_1/\theta_2)$ are then calculated from the prior (see sec. 2.4.3) to give,

$$\vartheta_x^{(2\mathrm{l})} \pm \Delta\vartheta_x^{(2\mathrm{l})} = t \exp\left[\langle \log\left(\frac{\theta}{t}\right)\rangle\right] \pm 2\vartheta_x^{(2\mathrm{l})} \sinh\left[\frac{1}{2}\sqrt{\langle c^{(2\mathrm{l})}(\vartheta_x^{(2\mathrm{l})}, \theta)\rangle}\right], \tag{2.28}$$

where the expectation values are taken with respect to the posterior distribution.

From this example it might seem that it doesn't matter whether the prior distribution is chosen based on physical invariances in the parameter or on the information geometry of the

problem. However, in slightly more complicated scenarios, the prior derived from the Fisher information can become significantly more complicated. The next two case studies will show this and other incompatibilities between the streamlined global Bayesian approach and using a prior based on the Fisher information.

2.5 Quantifying the actual uncertainty

As we have already noted, when asymptotically many measurements are made, the central limit theorem ensures that the posterior distribution will be a Gaussian distribution with a vanishingly small variance and a mean value equal to the true value of the parameter. In this case, the choice of prior and estimator are no longer important as all priors and estimators will converge to the same point for such a distribution. However, in the limit of finite data, the posterior distribution may be very broad. The shape of this distribution will still be very dependent on the prior. Additionally, different choices of estimator and cost function will give a different summary of the results. It is therefore, useful to determine if the estimators that are usually relied on, the mean, median and mode, as well as their associated errors reliably capture the true value and spread of the data. In fact, when the distribution is so broad, the actual value of the estimator may not even be meaningful. It may be more relevant that the associated error bars are broad enough and correctly capture the uncertainty in the measurement. We will show in Chapter 3 that in the example of temperature estimation of a qubit the errors associated with the mean and median do not always correctly quantify the actual error made. To facilitate this discussion, two methods for quantifying the actual deviation are now introduced.

2.5.1 Average deviation from the true value

Given an experiment with likelihood $P(x|\theta)$ and true value of the parameter $\tilde{\theta}$ the root mean square deviation from the true value given a particular estimate ϑ_x is,

$$\bar{\epsilon}^{(\mathrm{RMS})}(\tilde{\theta}) = \sqrt{\int dx P(x|\tilde{\theta})(\vartheta_x - \tilde{\theta})^2}. \tag{2.29}$$

This deviation is the "actual" error bar that will be associated with an estimate for a given value of $\tilde{\theta}$. It captures the spread of the estimator given the average realisation of the data at each $\tilde{\theta}$. However, in a real experiment it is not possible to know what this error would be because we do not have access to the true value of the parameter. Therefore, the error that is estimated in Eq. (2.21) should capture this spread as well as possible. To quantify this, first we define the average estimate

$$\bar{\vartheta}^{(f)}(\tilde{\theta}) = \int dx P(x|\tilde{\theta}) \vartheta_x^{(f)}, \tag{2.30}$$

which gives a weighted average of the estimate over all possible outcomes of the experiment based on how likely they are to occur given a particular true value for θ. Similarly, we define the average error estimated from the posterior associated with these parameter estimates,

$$\bar{\epsilon}^{(f)}(\tilde{\theta}) = \int dx P(x|\tilde{\theta}) \int d\theta P(\theta|x) c^{(f)}(\vartheta_x^{(f)}, \theta) \tag{2.31}$$

For large amounts of data $\bar{\epsilon}^{(f)}(\theta)$ will coincide with $\bar{\epsilon}^{(\mathrm{RMS})}(\theta)$ but it will be considered an advantage when comparing different estimation strategies if the Bayesian errors also capture the average spread for finite data.

2.5.2 Confidence intervals

A standard frequentist way to define an error over many realisations of an experiment is a confidence interval. The $y\%$ confidence interval about the actual value of the parameter θ, is the interval $[\theta_a, \theta_b]$ of parameter values such that $P(\theta_a < \theta < \theta_b) \geq y\%$. The 90% confidence interval, for example, covers the range of values of the parameter the experimenter could expect in 90% of experiments [123].

The Bernstein von Mises theorem gives the result that in the limit of many measurements, the credibility region defined in Eq.(2.22) should coincide with a confidence interval defined for a particular true temperature [100]. For limited data, when the estimator and prior are chosen well, the credibility and confidence intervals for a given true value of θ should both cover the actual value of the parameter as well as the average estimated value.

2.6 Comparing priors and estimators

It is tempting to compare the average performance of different estimation methods by comparing the average absolute value of the error that is achieved. However, this poses some problems. First: if the prior distributions are different, the posterior distribution will also differ resulting in a different summary of the data captured by the error. Therefore, errors can only be compared if they are based on the same prior distribution. Secondly, different error estimators may estimate the error in different units and therefore, the errors should be converted to the same units before comparison.

Assuming that the general method to derive an estimator given in Sec. 2.4.3 is followed, different choices of prior will lead to different estimators and their associated cost function. However, even if the data sets are fixed, there is no sense in comparing different magnitudes in the errors because the information contained in the posterior distributions is different because different priors are assumed. In this setting, different choices of prior and thus estimators can only be compared to their own bounds (discussed in Sec. 2.7) or alternatively, by how much of a spread they produce in the estimated values over several experiments and whether the error bars estimated from the posterior distribution match this spread on average.

On the other hand, if the prior is fixed, the posterior distribution for a particular data set is also fixed. Therefore, if we choose a cost function and estimator by heuristic means as in Sec. 2.4.1, it is now possible to compare the performance of the different estimators in an average realisation of the experiment. This can be done by determining whether the average spread of the estimates aligns with the average error bars determined by the cost function. In simulations where the true value of the parameter is known, the estimated value from different estimators can also be compared to the true value and thus, different estimators can be compared to each other.

In this scenario, the cost functions must be converted to have the same units so that the comparison is valid. Furthermore to facilitate the comparison of the cost functions with confidence intervals and credibility regions, we scale the cost functions by a prefactor so that for a Gaussian posterior each error estimate would match the 90% credibility region.

To do this we assume the posterior distribution is Gaussian with mean ϑ and variance $\sigma^2 =$

$\left\langle c^{(2)}(\vartheta^{(2)},\theta)\right\rangle$. The distribution is symmetric and so we must then determine the value of n in $P[\vartheta - n\sigma \leq \theta \leq \vartheta + n\sigma] = 0.9$. That is,

$$\begin{aligned} 0.9 &= \int_{\vartheta-n\sigma}^{\vartheta+n\sigma} dx \frac{1}{\sqrt{2\pi\sigma}} \exp \frac{(x-\vartheta)^2}{2\sigma^2} \\ &= \int_{-n}^{n} dz \frac{1}{\sqrt{2\pi}} \exp \frac{-z^2}{2} \end{aligned} \tag{2.32}$$

where, in the second line we have performed the standard change of variables $(x-\vartheta)^2/\sigma^2 \to z$ to get the integral in terms of the error function:

$$0.9 = \text{erf}\left[\frac{n}{2\sqrt{2}}\right]. \tag{2.33}$$

We can now solve for n to get that $n \approx 3.29$ which is the value we will use to compare square errors to the credibility regions in Chapter 3. For absolute errors, we use the standard result known as the asymptotic relative efficiency of the mean to the median [70] which gives that for a standard normal distribution $\left\langle c^{(2)}(\vartheta^{(2)},\psi)\right\rangle = \sqrt{\pi/2}\left\langle c^{(1)}(\vartheta^{(2)},\psi)\right\rangle$. Therefore absolute errors must be scaled by $n \approx 4.12$ to compare them with the 90% confidence interval in the limit of many measurements.

2.7 Optimal measurements and global bounds

One of the problems with defining the optimal measurement in terms of the Quantum Fisher information (see Eq.(2.4)) is that in some situations, like equilibrium thermometry. The optimal POVM depends on the parameter θ. Therefore measurement optimisation using this method is only really useful when we are already in the local limit where the value of the parameter is already approximately known. An alternative approach was developed in [104, 105] for priors derived from the physical symmetries of the parameter (see Sec. 2.3.1). In [D] we extend this and apply the same logic to the prior derived from the QFI. The approach is outlined here but a detailed proof of these results is given in [105]. We will define bounds on the error and obtain a Bayesian optimal POVM which relies only on knowledge of the prior distribution.

The global error that we optimise is,

$$\bar{\epsilon}^{(f)} = \int dx \int d\theta P(x|\theta)P(\theta)c^{(f)}(\vartheta_x, \theta), \tag{2.34}$$

where we restrict to the cost functions derived from the prior that we obtained in Sec. 2.4.3. Given the state of the probe $\rho(\theta)$ and POVM element $M(x)$ we can apply the Born rule, $P(x|\theta) = \mathrm{tr}(M(x)\rho(\theta))$ to obtain,

$$\begin{aligned}\bar{\epsilon}^{(f)} &= \mathrm{tr}\left\{\int dx M(x) \int d\theta \rho(\theta) P(\theta)(f(\vartheta_x) - f(\theta))^2\right\} \\ &= \int d\theta P(\theta) f(\theta)^2 \\ &\quad + \mathrm{tr}\left\{\int d\theta P(\theta)\rho(\theta) \int dx M(x) f(\vartheta_x)^2 - 2\int d\theta P(\theta)\rho(\theta) f(\theta) \int dx M(x) f(\vartheta_x)\right\}\end{aligned} \tag{2.35}$$

The first term is fixed by the prior and the last line is simplified using Jensen's inequality[6] which is saturated for projective measurements and we assume we make these here. Thus, we want to find the measurement that minimises,

$$\mathrm{tr}\left\{\int \mathrm{d}\theta P(\theta)\rho(\theta)\left[\int \mathrm{d}x M(x) f(\vartheta_x)\right]^2 - 2\int \mathrm{d}\theta P(\theta)\rho(\theta) f(\theta) \int \mathrm{d}x M(x) f(\vartheta_x)\right\}. \tag{2.36}$$

The resulting minimum error is achieved by projective measurements onto the eigenbasis of the operator $\mathcal{S}$ which solves the Lyaponov equation [105, 124],

$$\mathcal{S}\int d\theta P(\theta)\rho(\theta) + \int d\theta P(\theta)\rho(\theta)\mathcal{S} = 2\int d\theta P(\theta)\rho(\theta) f(\theta). \tag{2.37}$$

We represent this POVM as $S(x)$. This defines the minimum error,

$$\bar{\epsilon}^{(f)}_{\min} = \int d\theta P(\theta) f(\theta)^2 - \mathrm{tr}\left\{\int \mathrm{d}\theta P(\theta)\rho(\theta)\left[\int \mathrm{d}x S(x) f(\vartheta_x)\right]^2\right\}, \tag{2.38}$$

[6] $\int dx M(x) f^2(\vartheta_x) \geq \left(\int dx M(x) f(\vartheta_x)\right)^2$

to which other strategies that do not employ the globally optimal POVM (here represented by $S(x)$) can be compared.

We can also define the worst case scenario, which is when the estimation strategy yields no information about the parameter beyond that which is contained in the prior. This allows us to determine how well sub-optimal strategies perform compared to the best and worst cases. In this case, the error will be the prior uncertainty [104] which is just the prior variance of f,

$$\bar{\epsilon}_{\mathrm{p}}^{(f)} = \int \mathrm{d}\theta P(\theta) f^2(\theta) - \left[\int \mathrm{d}\theta P(\theta) f(\theta)\right]^2. \tag{2.39}$$

The maximum information is gained by the globally optimal measurement with $\bar{\epsilon}_{\mathrm{min}}^{(f)} = \bar{\epsilon}_{\mathrm{p}}^{(f)} - \mathcal{G}^{(f)}$ where we defined the precision gain by this measurement as,

$$\mathcal{G}_{\mathrm{max}}^{(f)} = \mathrm{tr}\left(\left[\int \mathrm{d}x S(x) f(\vartheta_x)\right]^2 \int d\theta P(\theta)\rho(\theta)\right) - \mathrm{tr}\left(\int \mathrm{d}x S(x) f(\vartheta_x) \int d\theta P(\theta)\rho(\theta)\right)^2. \tag{2.40}$$

In situations where the optimal measurement is difficult to perform, the gain of other non-optimal measurements $M(x)$ can then be compared against this gain to determine how close they are to the optimum.

2.8 Bayesian bounds

The Cramér-Rao bound in Eq. (2.3), is also applicable to Bayesian error measures, Eq. (2.21), as long as the underlying estimator $\vartheta^{(f)}$ is unbiased. Bayesian estimators are generically biased for small amounts of data and thus this bound may only apply in the asymptotic case of many measurements where the estimator becomes unbiased. We present the example for the relative square error which will feature prominently in both Chapter3 and 4. In this example, the relative square cost function $c^{(2\mathrm{r})}(\theta_x, \theta) = (\theta_x/\theta - 1)^2$ is used with estimator $\vartheta_x^{(2\mathrm{r})} = \langle\theta^{-1}\rangle/\langle\theta^{-2}\rangle$. To get a Bayesian CRB we first define the average cost function over the prior and all possible experiments,

$$\bar{\epsilon}^{(2\mathrm{r})}(\vartheta_x) = \int \mathrm{d}\theta P(\theta) \int dx P(x|\vartheta_x) \left(\frac{\vartheta_x}{\theta} - 1\right)^2. \tag{2.41}$$

Then the Bayesian CRB is [125, 126],

$$\bar{\epsilon}^{(2\mathrm{r})}(\vartheta_x) \geq \int \mathrm{d}\theta \frac{P(\theta)}{\theta^2 \mathcal{I}(\theta)}, \tag{2.42}$$

or for a non-relative square error,

$$\bar{\epsilon}^{(2)}(\vartheta_x) \geq \int \mathrm{d}\theta \frac{P(\theta)}{\mathcal{I}(\theta)}. \tag{2.43}$$

These inequalities are usually only valid in the limit of many measurements since it requires that the estimator is unbiased. When the estimator is unbiased, this bound is tight.

In a similar way to the optimal POVM derived in the previous section, we can define a Bayesian benchmark for the variance of the estimates. This is known as a van Trees inequality and it is defined [127]

$$\bar{\epsilon}^{(\mathrm{RMS})} \geq \frac{1}{\sqrt{\int d\theta P(\theta)\mathcal{I}(\theta) + I_0}}, \quad I_0 = \int d\theta P(\theta) \left[\partial_\theta P(\theta)\right]^2, \tag{2.44}$$

where $\mathcal{I}$ is the Fisher information given in Eq. (2.1).

For parameters that are scale invariant and therefore better captured by a relative error $(\bar{\epsilon}^{(\mathrm{rRMS})})^2 = (\bar{\epsilon}^{(\mathrm{RMS})})^2/(\tilde{\theta})^2$, this inequality can be modified to give,

$$\bar{\epsilon}^{(\mathrm{rRMS})} \geq \frac{1}{\sqrt{\int d\theta P(\theta)\theta^2\mathcal{I}(\theta) + I_0}}, \quad I_0 = \int d\theta P(\theta)\theta^2 \left[\partial_\theta P(\theta)\right]^2. \tag{2.45}$$

This relative bound is derived in App. C.

2.9 Conclusions

In this chapter we have introduced different approaches to Bayesian quantum parameter estimation. Since the usual Cramér-Rao bound is only valid asymptotically, we also present a Bayesian bound for both absolute and relative errors that can be used when there is very little data. Additionally we introduce the Bayesian CRB and a global bound on the posterior error. We also

discuss ways to compare different estimators and errors as well as give a global sensitivity range for experiments with discrete outcomes.

Broadly speaking, we can identify two methodologies for applying Bayesian techniques. In the first methodology, the prior distribution is chosen in a way that is as uninformative as possible by basing it on invariances in the physical parameter or the Fisher information. The information contained in the posterior distribution is then summarised by a cost function and its associated optimal estimator. The decisions about which cost function and prior to use may be motivated by practical considerations like convenience of calculation. The measurement strategy is optimised separately based on considerations like optimising the quantum Fisher information.

On the other hand we may approach parameter estimation in a fully global and formal way. In this case, the prior distribution should be chosen with as few assumptions as possible and this choice determines the cost functions and estimators that are used to summarise the information contained in the posterior distribution. Additionally, these are then used to optimise the experiment for global estimation by choosing a globally optimal POVM which may not be the same as the locally optimal POVM found from the QFI. The most uninformed approach, which should be favoured if no prior information about the parameter can be assumed, would be to base the choice only on physical symmetries present in the estimation problem. Basing the prior on symmetries present in a family of likelihoods or a particular likelihood, requires that more assumptions are made.

In the next two chapters these approaches will be put to use in a series of case studies. Here, we will see how these two approaches compare in terms of ease of calculation as well as how well they capture the available data.

Chapter 3

Parameter estimation with discrete measurements

In this chapter we will explore two case studies that exemplify Bayesian metrology in open systems. In both case studies we focus on estimation protocols where probes are prepared and then measured in a two step process. The goal is to provide examples for the Bayesian analysis introduced in the previous chapter in a theoretically simple setting but which can also be applied to experimental platforms. In the previous chapter two approaches to Bayesian estimation were identified. On the one hand we identified a formal approach where, the choice of prior determines the cost functions and estimators that are used to summarise the information contained in the prior. The case study in Sec. 3.2, which will be published in [D], explores this approach and the focus is on the consequences of the choice of prior distribution. In Sec. 3.1, we take the more heuristic approach introduced in the previous chapter. Here, the prior distribution is fixed and we explore how commonly chosen cost functions and their associated estimators like the median and mean compare to the ones that would be chosen from the more streamlined approach. The example in this section is also the most experimentally relevant. We consider the case of qubit thermometry with both equilibrium and non-equilibrium probes and find that non-equilibrium probes perform better for global estimation.

3.1 Qubit thermometry

The next case study we consider is qubit thermometry. In Sec. 2.1 we discussed how this is the paradigmatic example of open system metrology. The example we consider here allows for both equilibrium and non-equilibrium probes while still being simple enough for the full Bayesian analysis to be applied. It is inspired by the thermometry protocol introduced in [94]. We will compare the heuristic choices of cost functions and estimators given that the prior distribution is fixed. We will first consider the case where the Fisher information is used to derive the prior according to Jeffreys' rule Sec. 2.3.3. We then briefly compare this with the fully global approach using the prior derived from scale invariance of the parameter. We will also illustrate the use of the sensitivity range that is presented in Sec. 3.1.5 to analyse the advantages that can be obtained by non-equilibrium probes for global temperature estimation.

3.1.1 Qubit model

A two-level probe with ground and excited states $|0\rangle$ and $|1\rangle$ and non-interacting Hamiltonian $H_s = E/2(|1\rangle\langle 1| - |0\rangle\langle 0|)$, with $\hbar = 1$, is resonantly coupled to a thermal, Bosonic reservoir of unknown temperature T for a fixed time τ. This interaction is assumed to be weak and we make the Born-Markov and secular approximations so that we can model this interaction with the standard Markovian master equation derived in App. A. In the system-bath interaction picture, this yields

$$\dot{\rho} = \frac{\gamma}{1 - e^{-E/k_B T}} \left(\mathcal{D}[|0\rangle\langle 1|]\rho + e^{-E/k_B T} \mathcal{D}[|1\rangle\langle 0|]\rho \right), \tag{3.1}$$

with $\mathcal{D}[\hat{A}]\rho = \hat{A}\rho\hat{A}^\dagger - \{\hat{A}^\dagger\hat{A}, \rho\}/2$. The probe is prepared in the ground state as this corresponds to the most temperature sensitivity, which was shown in [94].

After the coupling time τ, the probe is in the state,

$$\rho(\tau) = q_\tau(T) \, |1\rangle\langle 1| + (1 - q_\tau(T)) \, |0\rangle\langle 0| \tag{3.2}$$

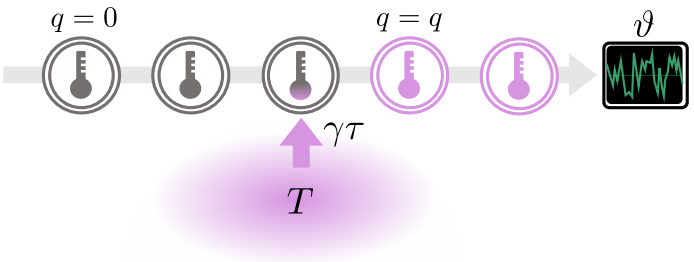


Figure 3.1: Sketch of the qubit thermometer. N ground-state qubits exchange heat with a bosonic reservoir at temperature T and rate γ, each for a duration τ. This partial thermalization brings the qubit ensemble to a mixed state with excitation probability q_τ. The experimenter infers the temperature from the measured number of excitations by means of a Bayesian estimator $\boldsymbol{\vartheta}$. Adapted from [A].

where, $q_\tau(T)$ is the excitation probability,

$$q_\tau(T) = \frac{1 - e^{-\gamma\tau \coth(E/2k_BT)}}{1 + e^{E/k_BT}}. \tag{3.3}$$

The probe is then decoupled, measured and replaced by a new probe in the same initial state. The setup is depicted in Fig 3.1. After N repetitions the number of times the probe was measured to be in the excited state is used to infer the temperature. For long coupling times, $\gamma\tau \gg 1$, the probes will equilibrate to the Gibbs state (2.8), with excitation probability, $q_\infty(T) = 1/(1+e^{E/k_BT})$. Before discussing any measurements we can determine some limiting cases. In the low temperature limit, $k_BT \ll E$, even non-equilibrium probes will suffer from the exponentially supressed sensitivity that is already well known for thermal probes [48]. In this limit, the excitation probability scales as $q_\tau(T) \sim e^{-E/k_BT} \to 0$ and thus, no data can be obtained as the probe is stuck in the ground state in every trial.

For thermal probes at high temperatures, $k_BT \gg E$, the excitation probability $q_\infty(T) \approx 1/2 - E/4k_BT$ saturates at $1/2$. This will also degrade the sensitivity since it will be difficult to tell two very high temperatures apart as they will both have approximately half of the measurements in the excited state. On the other hand, the high temperature limit for a finite τ yields

$$q_\tau(T) \xrightarrow{k_BT \gg E} \left(\frac{1}{2} - \frac{E}{4k_BT}\right)\left(1 - e^{-2\gamma\tau k_BT/E}\right). \tag{3.4}$$

Therefore, for sufficiently sufficiently small $\gamma\tau$, the excitation probability will no longer saturate to $1/2$.

3.1.2 Priors

In this case study we will mostly be concerned with a comparison of cost functions and estimators given a fixed prior. To this end, we will primariliy use the prior $P(\theta) \propto \sqrt{\mathcal{I}_Q(\theta)}$ which is explained in more detail in Sec. 2.3.3. We choose this prior since it decays to zero at $T \to 0$ and $T \to \infty$, which allows us to use the Bayesian bound to analyse the outcomes of the experiment given in Eq. 2.45 and derived in App. C. The quantum and classical Fisher information are the same for the qubit probe because there are no coherences. For a probe which has interacted for time τ the Fisher information is,

$$\mathcal{I}_Q(\theta) = \mathcal{I}(\theta) = \frac{E^2 \left(2e^{\frac{E}{\theta k}} \left(e^{\gamma\tau \coth\left(\frac{E}{2\theta k}\right)} - 1\right) + \gamma\tau \left(e^{\frac{E}{\theta k}} + 1\right) \operatorname{csch}^2\left(\frac{E}{2\theta k}\right)\right)^2}{4\theta^4 k^2 \left(e^{\frac{E}{\theta k}} + 1\right)^2 \left(e^{\gamma\tau \coth\left(\frac{E}{2\theta k}\right)} - 1\right) \left(e^{\gamma\tau \coth\left(\frac{E}{2\theta k}\right) + \frac{E}{\theta k}} + 1\right)}. \tag{3.5}$$

For thermal qubits ($\tau \to \infty$) we get the thermal Fisher information,

$$\mathcal{I}_Q^{(\mathrm{th})}(\theta) = \frac{E^2 \operatorname{sech}^2\left(\frac{E}{2\theta k}\right)}{4\theta^4 k^2}. \tag{3.6}$$

Next, if we decide to only make assumptions on the level of invariances in the parameter, we notice that temperature is a scale invariant parameter since it is scaled by the energy. Using the prior introduced in Sec. 2.3.1 with the location isomorphic function of a scale parameter $f(\theta) = \log(\theta)$ [105] we determine the prior distribution that respects this invariance from $\partial_\theta f(\theta) = 1/\theta$. Therefore, the prior based on symmetries of the parameter is $P(\theta) \propto 1/\theta$. This prior also diverges at $T \to 0$ which makes it unsuitable for deriving the relative van Trees bound (see Sec.2.8) which we use in later sections unless it is made smooth at some cut-off. These different choices of prior are compared in Sec.3.1.8, where results from using the constant prior are also shown to demonstrate that an unjustified use of a "flat" prior should be avoided. Unless otherwise state the rest of this section deals with the prior $P(\theta) \propto \sqrt{\mathcal{I}_Q(\theta)}$.

Cost function	Estimator	Error estimate ($\varepsilon_n^{(i)}$)
$c^{(1)}(\theta_n, T) = \lvert\theta_n - T\rvert$	median of posterior	$4.12\,\langle c^{(1)}(\vartheta_n^{(1)}, T)\rangle$
$c^{(1\mathrm{r})}(\theta_n, T) = \lvert\theta_n/T - 1\rvert$	median of $P(T\lvert n)/T$	$4.12\,\vartheta_n^{(1\mathrm{r})}\langle c^{(1\mathrm{r})}(\vartheta_n^{(1\mathrm{r})}, T)\rangle$
$c^{(2)}(\theta_n, T) = (\theta_n - T)^2$	$\vartheta_n^{(2)} = \langle T\rangle$	$3.29\,\sqrt{\langle c^{(2)}(\vartheta_n^{(2)}, T)\rangle}$
$c^{(2\mathrm{r})}(\theta_n, T) = (\theta_n/T - 1)^2$	$\vartheta_n^{(2\mathrm{r})} = \langle T^{-1}\rangle/\langle T^{-2}\rangle$	$3.29\,\vartheta_n^{(2\mathrm{r})}\sqrt{\langle c^{(2\mathrm{r})}(\vartheta_n^{(2\mathrm{r})}, T)\rangle}$
$c^{(2\mathrm{l})}(\theta_n, T) = \ln^2(\theta_n/T)$	$\vartheta_n^{(2\mathrm{l})} = \frac{E}{k_B}\exp\left[\langle\log\left(\frac{k_B T}{E}\right)\rangle\right]$	$2\vartheta_n^{(2\mathrm{l})}\sinh\left[\frac{3.29}{2}\sqrt{\langle c^{(2\mathrm{l})}(\vartheta_n^{(2\mathrm{l})}, T)\rangle}\right]$
$c^{(\mathrm{md})}(\theta_n, T) = -\delta(\theta_n - T)$	$\vartheta_n^{(\mathrm{md})} = \arg\max_\theta P(\theta\lvert n)$	not applicable

Table 3.1: A summary of the cost functions and their associated optimal Bayesian estimators that will be used in the qubit thermometry case study. In the last column the error estimates in units of temperature that will be used to compare the average error made for N measurements to 90% confidence intervals and the 90% credibility regions of the posterior are given in the final column.

3.1.3 Estimators and cost functions

For this case study we will primarily be comparing different error estimates $\varepsilon_n^{(i)}$ based on cost functions and their associated estimators and then converted to units of temperature, as well as 90% confidence intervals and 90% credible regions as is discussed in sections 2.4.2 and 2.5.2. We will therefore scale all of the error estimates so that in the asymptotic limit they coincide with 90% confidence intervals and we will also express all of them in units of temperature. This is summarised in Table 3.1 The cost functions in the table come from the heuristic approach to estimation where the choice of cost function is independent of the prior and is based on error quantifiers that usually work well (see Sec. 2.4.1). This is the approach that is often taken in Bayesian estimation and as we will show some choices are not justified for particular parameters.

The reason for this becomes clear if we consider the arguments in sec 2.4.3. For example, if we have chosen the prior $P(\theta) \propto 1/\theta$ based on scale invariance of the temperature, the cost function is fixed to $c^{(f)}(\vartheta_x, \theta) = (f(\vartheta_x) - f(\theta))^2$, which in the case of a scale invariant parameter with location isomorphic function $f(\theta) = \log(\theta)$, gives the cost function $c^{(2\mathrm{l})}(\theta_n, T) = \log^2(\theta_n/T)$. We notice here that the logarithmic cost function takes into account that a scale

parameter must not be penalised on absolute deviations but rather relative deviations. Therefore, we should expect that for this thermometry example, the relative estimators will perform better than the mean, median and mode of the posterior.

Additionally, the two relative cost functions have the added benefit of not having diverging moments on an unrestricted temperature range if the prior is not improper. If we were to determine the cost function from the prior that depends on the Fisher information using Eq.(2.25) we have, $f(\theta) = \int \mathrm{d}\theta \sqrt{\mathcal{I}_Q(\theta)}$. For equilibrium probes this yields,

$$f(\theta) = -\frac{4\theta^2 k_B^2 \mathrm{Li}_2\left(-e^{-\frac{E}{k_B\theta}}\right) + E\left(-4\theta k_B \log\left(e^{-\frac{E}{\theta k_B}}+1\right) + E\tanh\left(\frac{E}{2\theta k_B}\right) - E\right)}{2\theta^2 k_B E}. \tag{3.7}$$

Here $\mathrm{Li}_2(x)$ is the polylogarithm function, $\sum_{k=1}^{\infty} x^k/k^2$. The location isomorphic function is even more complicated for non-equilibrium probes.

For this case study we will not include the cost function $c^{(f)}(\vartheta_x, \theta) = (f(\vartheta_x) - f(\theta))^2$, rather, an example of a cost function derived from the prior based on the Fisher information is explored in Sec. 3.2.

3.1.4 Optimal measurements

First, in the local approach to measurement optimisation, as is explained in Sec. 2.1 the optimal measurement is to perform a projective measurement in the eigenbasis of the symmetric logarithmic derivative. That is, we solve for L in Eq.(2.4) but since the state is diagonal in the energy basis, so is the SLD. Therefore, optimal measurements would be in the energy basis of the probe. This is unsurprising since the probe has no coherences and thus measurements that are not in the diagonal basis cannot extract any information from off-diagonal elements in the density matrix (there are none).

Next as is discussed in sec. 2.7, the global approach to measurement optimisation requires us to specify the prior distribution and the location isomorphic function $f(\theta)$. If we assume that $P(\theta) = 1/\theta$, the location isomorphic function is $f(\theta) = \log(\theta)$. Then, to find the globally

optimal POVM, we must solve for $\mathcal{S}$ in the Lyaponov equation [128],

$$\mathcal{S}\int\frac{\mathrm{d}\theta}{\theta}\rho(\theta)+\int\frac{\mathrm{d}\theta}{\theta}\rho(\theta)\mathcal{S}=2\int\mathrm{d}\theta\frac{\log(\theta)}{\theta}\rho(\theta), \tag{3.8}$$

and the optimal POVM will be the eigenvectors of $\mathcal{S}$. Although the integral looks much more complicated, because the state is diagonal, the optimal measurement turns out to be a projective measurement in the energy basis once more. Since we have arrived at the conclusion that regardless of the approach, optimal measurements are made in the diagonal basis, we can immediately write down the likelihood function for the populations of the density matrix. The likelihood for n excitations in N independent trials follows a Bernoulli chain,

$$P_\tau(n|T)=\binom{N}{n}q_\tau^n(T)\left[1-q_\tau(T)\right]^{N-n}, \tag{3.9}$$

where, $q_\tau(T)$ is the probability to be in the excited state at time τ when the environment has temperature T.

3.1.5 Sensitivity range

A global approach to estimation excludes no values of the unknown parameter. This is desirable, however any actual apparatus will only be accurate within a certain parameter range. This limitation of the apparatus is captured by the likelihood. For example if a parameter takes possible values in the range $(T_{\text{min}}, T_{\text{max}})$, at sufficiently small values of the actual parameter the apparatus will not be able to distinguish this small value from T_{min}.The range to which the apparatus is able to distinguish the measured value of the parameter from the extreme points will be referred to as the sensitive range.

We will estimate the sensitive range by comparing the likelihoods at the extreme points in the support of the temperature. Each measurement represents one bit of data, hence, two values of the parameter, T_1 and T_2 are barely distinguishable if the relative base-2 entropy between their likelihoods,

$$D(T_1\|T_2)=\sum_{n=0}^{N}P(n|T_1)\log_2\frac{P(n|T_1)}{P(n|T_2)}, \tag{3.10}$$

measures no less than 1 bit. This will be zero only when $T_1 = T_2$ and will have a maximum value of N. We can thus define a sensitive range to be between the boundaries T_a and T_b at which $D(T_{\min}\|T_a) = 1$ and $D(T_{\max}\|T_b) = 1$.

For the Bernoulli chain Eq.(3.9), we get

$$D(T_1\|T_2) \;=\; Nq_\tau(T_1)\log_2\frac{q_\tau(T_1)}{q_\tau(T_2)} + N[1-q_\tau(T_1)]\log_2\frac{1-q_\tau(T_1)}{1-q_\tau(T_2)}.$$

It is zero only if $T_1 = T_2$, and otherwise a positive number of at most N bits. We can thus restrict our view on temperatures between the boundaries T_0 and T_∞ at which $D(0\|T_0) = 1$ and $D(\infty\|T_\infty) = 1$. In terms of the excitation probability this condition gives,

$$q_\tau(T_0) = 1 - 2^{-1/N}, \quad q_\tau(T_\infty) = \frac{1-\sqrt{1-4^{-1/N}}}{2}. \tag{3.11}$$

Therefore for a fixed interaction time τ and number of experiments N we can find the temperatures T_0 and T_∞ that we can reasonably estimate.

3.1.6 Numerical comparison of estimators and cost functions

In this section we will now compare the performance of the different estimators and cost functions from table 3.1. We fix the prior distribution which for a given set of outcomes fixes the posterior distribution. Therefore, the comparison captures how well the cost functions and estimators summarise the same posterior distribution. In this section we assume the prior distribution $P(\theta) \propto \sqrt{\mathcal{I}}$.

To make the comparison as simple as possible we will further assume that each probe comes to equilibrium with the environment before it is measured, that is, $\gamma\tau \gg 1$. Additionally, we will consider a situation where $N = 200$ probes are measured to obtain a single data set and determine the value of an estimator $\vartheta_n^{(f)}$ and its associated cost function $\varepsilon_n^{(f)}$. This is simulated at fixed actual temperatures in the range $k_BT/E \in [1/200, 200]$. The simulations require integrals to be evaluated numerically and for this we also restrict the integrals to the range $k_BT/E \in [1/200, 200]$. Note that this is a large range given that only 200 measurements are performed and this justifies the use of global estimation.

Of course, no statements can be made about how a particular estimator or cost function performs from a single data set. Therefore we must average over data sets which can be done in two ways. Either, the estimates and their associated error over many random trials at a particular temperature T are calculated and averaged, or we can consider the probability to obtain each possible data set for a fixed actual temperature. The latter is more efficient when only a few probes are considered as is the case here. The probability to obtain a given data set is obtained from the likelihood function and can be used to calculate the weighted averages of the estimator and cost function,

$$\bar{\vartheta}(T) = \sum_n P(n|T)\vartheta_n, \quad \bar{\varepsilon}(T) = \sum_n P(n|T)\varepsilon_n. \tag{3.12}$$

Now, since we are performing a simulation, we also know the actual value of the temperature and so we can quantify the true deviation of the estimate from the actual value as discussed in Sec. 2.5. This gives the Eq. (2.29) which we repeat here in discrete form for clarity,

$$\bar{\epsilon}^{(\mathrm{RMS})}(\tilde{\theta}) = 3.29\sqrt{\sum_n P(n|\tilde{\theta})(\vartheta_n - \tilde{\theta})^2}. \tag{3.13}$$

We have scaled this quantity by 3.29 so that it can be compared to the 90% confidence interval as is explained in Sec. 2.6. Note that in the figures of this section, temperatures that fall outside of the sensible range for estimation given the number of measurements (see section 3.1.5) are shaded over by light grey bars and thus we truncate the figure at $k_B T/E = 10$ since any temperatures higher than this fall outside of our sensible range.

We begin by analysing the bias $\bar{\vartheta}(T)/T$ of each estimator in Fig 3.2. The bias quantifies if the estimator tends to over or under estimate the temperature compared to the actual value of the temperature. This figure suggests that the mean estimator introduces a large bias in the estimate at higher temperatures. The opposite is true of the relative estimators and the mode which tend to under estimate the temperature at high values of the actual temperature. However, it is not clear from the bias how these estimators differ from each other or whether these estimators out perform the logarithmic estimator.

Next we compare how much the estimators vary between experiments using the average root

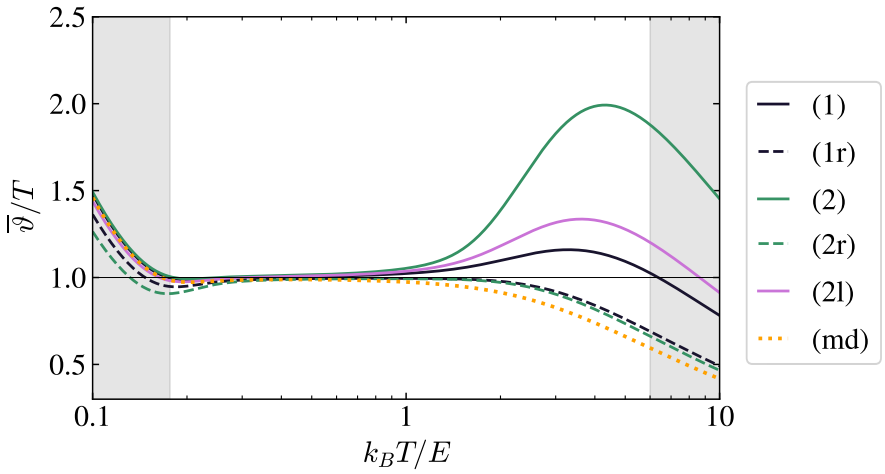


Figure 3.2: Outcome-averaged estimated temperatures $\bar{\vartheta}(T)$ from (3.12) relative to the true temperature T in units of Ek_B, for 200 fully thermalised qubits. The Bayesian estimators and their associated error measures are indicated in the legend and summarised in Table 3.1. The grey regions indicate temperatures outside the detectable range given by (3.11). Adapted from [A].

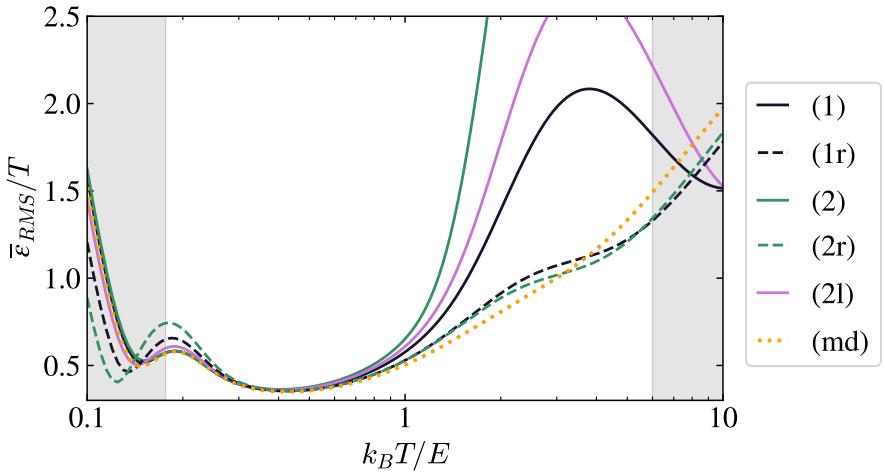


Figure 3.3: Average RMS deviations (3.13) relative to the true temperature T in units of Ek_B, for 200 fully thermalized qubits. The Bayesian estimators and their associated error measures are indicated in the legend and summarised in Table 3.1. The grey regions indicate temperatures outside the detectable range given by (3.11). Adapted from [A].

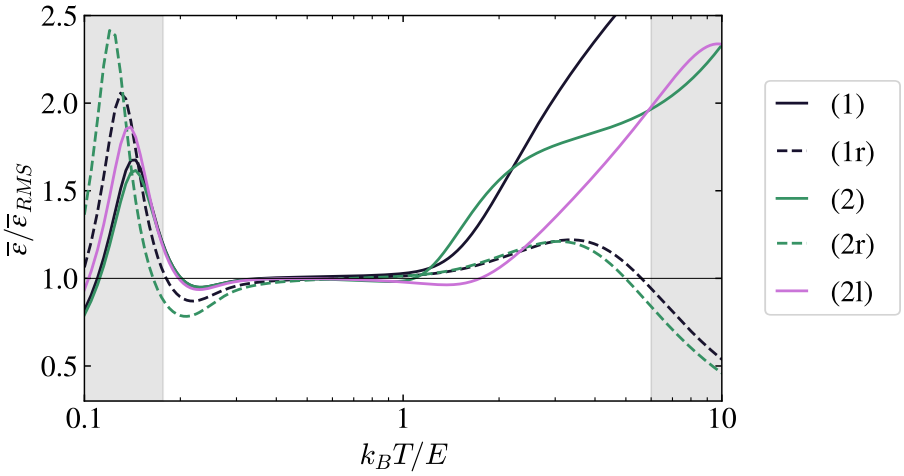


Figure 3.4: Average error measures $\bar{\varepsilon}(T)$ from (3.12) relative to the RMS deviations for true temperature T in units of Ek_B, for 200 fully thermalized qubits, excluding the mode estimator. The Bayesian estimators and their associated error measures are indicated in the legend and summarised in Table 3.1. The grey regions indicate temperatures outside the detectable range given by (3.11). Adapted from [A].

mean square error (see Eq. (3.13)) relative to the true temperature in Fig 3.3. Here we see that all estimators perform similarly and have their minimum spread for intermediate temperatures about the point $k_BT/E \lesssim 1$ where the experiment has the maximum local sensitivity. This is because for these temperatures even though only 200 measurements are performed, the posterior is sharply peaked. However, for high temperatures, the two relative eestimators and the mode have much less spread than the mean, median and logarithmic estimators. Finally, in Fig. 3.4, we compare the average spread of the estimates about the true temperature to the average estimate of the cost function. This quantifies how well the cost function captures the actual spread. Again, the relative cost functions perform better over the whole range of temperatures. The mode estimator doesn't have a meaningful cost function and so it is omitted from this figure. Additionally, we see that for the relative and logarithmic estimators, the bias in Fig. 3.2, contributes relatively little to the total error.

In Fig. 3.5 we compare the estimators in more detail. The bias of each estimator is shown in the first row with the solid curves. The 90% credible region (see Eq(2.22)) is denoted by

the dashed black lines and the 100%, 90% and 50% confidence intervals are shown in dark to light green shades. Here we see that indeed, around the most sensitive range of temperatures, the credible regions and confidence intervals match well. Which is expected from the Bernstein von Mises theorem once the posterior distribution has converged sufficiently towards a sharply peak Gaussian distribution. The bottom row compares the relative RMS deviations $\bar{\varepsilon}_{\mathrm{RMS}}(T)/T$ (solid blue) to the 90% credibility ranges $\bar{\varepsilon}^{90\%}$ (dashed black), the measured errors $\bar{\varepsilon}/T$ (dotted red), and the 90% confidence interval (green shade). The hope is that the relative RMS deviation matches the measured error as well as possible since this would indicate that on average the summary of the posterior captures the actual spread of the data correctly. First, the 90% credibility range does not depend on the estimator and so it is the same for each figure. Unfortunately, due to the shape of the posterior at high temperatures, it tends to over estimate the actual spread of the data significantly except for the median estimator. Considering these limitations, the two relative estimators seem to perform best all round. Therefore, when the cost function and hence, estimator are chosen on heuristic grounds for this experiment one of these would be the better choice than the standard choice of median or mean.

3.1.7 Asymptotic scaling of the error

In this section we will discuss how different cost functions will scale with the number of measurements made. Again, we will restrict our attention to equilibrium probes. Additionally, since the moments of the distribution diverge when we integrate over all $T \geq 0$, we restrict to a finite range $T \in [T_1, T_2]$ and compare the average cost function over all priors,

$$\mathcal{C}_{\text{fin}} = \frac{1}{\mathcal{M}} \int_{T_1}^{T_2} \mathrm{d}\theta \, P(\theta) \sum_{n=0}^{N} P(n|\theta) c(\vartheta_n, \theta), \tag{3.14}$$

where we renormalize the prior $P(\theta) = \sqrt{\mathcal{I}(\theta)}$ accordingly by

$$\mathcal{M} = \frac{4}{\pi} \left[\arcsin \sqrt{q_\tau(T_2)} - \arcsin \sqrt{q_\tau(T_1)} \right]. \tag{3.15}$$

This scaling is demonstrated for the different cost functions in Fig. 3.6 for the range of integration set to $k_B T/E \in [0.1, 10]$. This figure seems to suggests that the logarithmic and relative

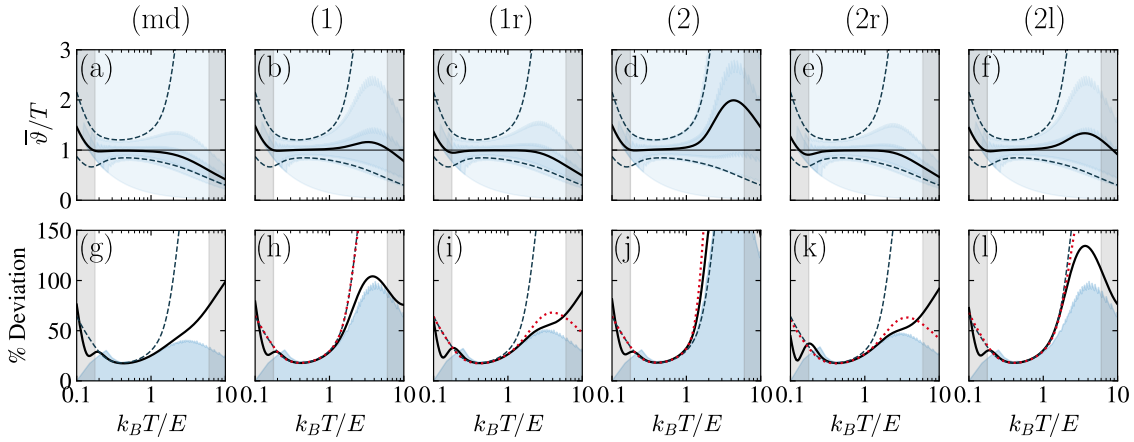


Figure 3.5: Temperature estimates and errors for $N = 200$ thermal qubits. Each two-panel column shows the results for the respective Bayesian estimator referenced by the label on top: the mode in (a,g), the median in (b,h), the relative median in (c,i), the mean in (d,j), the relative mean in (e,k), and the logarithmic mean in (f,l). In the first row, the solid curves show the average biases $\bar{\vartheta}(T)$, the dashed lines delimit the 90% credible regions, and the shaded regions from light to dark blue mark the 100%, 90%, and 50% confidence intervals, all relative to the true temperature T. The second row shows the average RMS deviation from the actual temperature T (solid lines), the 90% credible region widths $\bar{\varepsilon}^{90\%}$ (dashed line), the average measured errors $\bar{\varepsilon}(T)$ (red dotted line), and the 90% confidence intervals (blue shaded region), all given as a percentage with respect to T. The grey bars delimit the detectable temperature range. Adapted from [A].

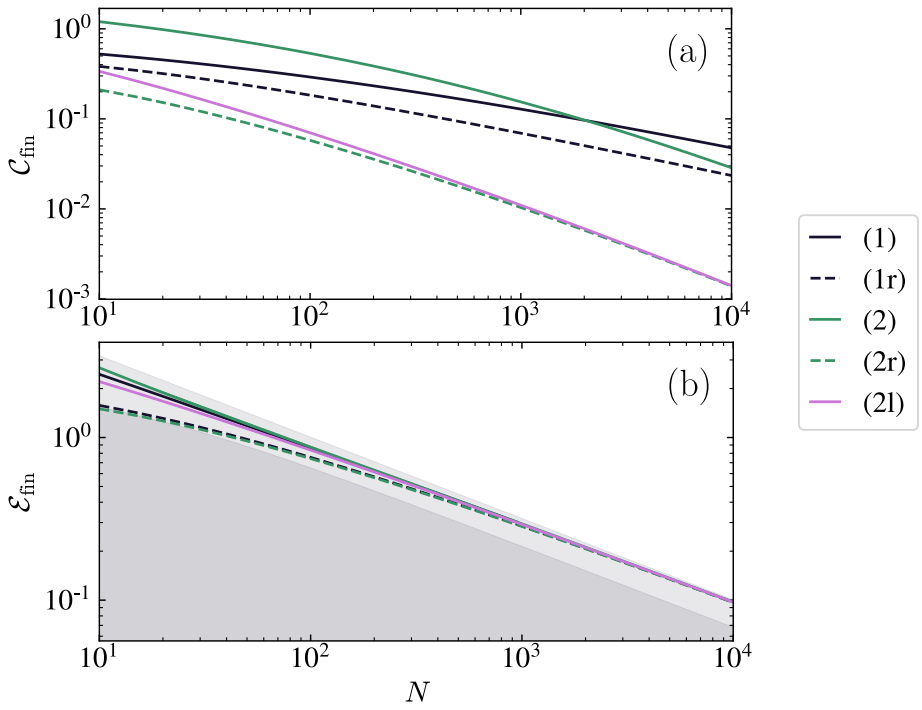


Figure 3.6: (a)The scaling of the average costs (3.14) over the temperature range $k_BT/E \in [0.1, 10]$ with probe number N for probes that are brought to thermal equilibrium. We evaluate this for the estimators listed in the legend and since the cost functions vary in dimension of T, we set $E/k_B \equiv 1.$. The scaling of the outcome-averaged relative errors (3.16) with probe number N, assuming the finite temperature range $k_BT/E \in [0.1, 10]$. The estimators are compared to the T-averaged Cramér-Rao bound (3.17) (light shade) and the global error benchmark (3.20) (dark) at thermal equilibrium. Adapted from [A].

mean estimators result in drastically lower average cost functions. However, this plot actually serves as a cautionary tale since all the plotted cost functions depend on different parameterisations of the error and therefore differ in dimension.

A proper comparison is made by converting the measured errors into relative temperature errors $\varepsilon_n/\vartheta_n$ and averaging them over all outcomes and temperatures in the prior range to obtain the dimensionless figure of merit,

$$\mathcal{E}_{\text{fin}} = \sum_{n=0}^{N} P(n)\frac{\varepsilon_n}{\vartheta_n}. \tag{3.16}$$

Note that to maintain consistency, both ε_n and ϑ_n must be evaluated over the same restricted range, $k_BT/E \in [T_1, T_2]$ when averages are taken over the posterior distribution. The scaling of this quantity with N for each of the estimated errors is shown in Fig. 3.6. First we will address the asymptotic scaling. Asymptotically all the average estimated errors scale like $1/\sqrt{N}$. This is expected since the posterior distribution will be sharply peaked. We compare this scaling to the Cramér-Rao bound for the RMS deviation of unbiased estimators, $\bar{\varepsilon}_{\text{CRB}}(T) = 3.29/\sqrt{\mathcal{I}(T)}$ (which we scale by 3.29 to match the 90% confidence interval and our scaled error estimates). In order to compare this to the error estimates, we convert this to a prior averaged relative deviation,

$$\mathcal{E}_{\text{CRB}} = \int_{T_1}^{T_2} \mathrm{d}T \frac{P^{(0)}(T)\bar{\varepsilon}_{\text{CRB}}(T)}{\mathcal{M}T} = \frac{6.58\ \ln\left(T_2/T_1\right)}{\pi\sqrt{N}\mathcal{M}}. \tag{3.17}$$

This bound is depicted by the light grey shade in Fig 3.6. Asymptotically, all of the error estimates converge to this bound which is expected since the bias that each estimator exhibits vanishes for large enough N. However, when asymptotic scaling has not yet been reached, for up to about 10^3 probes, the error estimates each have slightly different scaling. We compare this scaling to the scaling of the relative van Trees inequality [127] discussed in Sec.2.8,

$$\begin{aligned} \mathcal{E}_{\text{rRMS}} &\geq \frac{3.29}{\sqrt{\int_0^\infty \mathrm{d}T\ P^{(0)}(T)T^2\mathcal{I}(T) + \mathcal{I}_0}}, \\ \mathcal{I}_0 &= \int_0^\infty \mathrm{d}T\ P^{(0)}(T)T^2\left[\partial_T \ln P^{(0)}(T)\right]^2, \end{aligned} \tag{3.18}$$

where we use the prior averaged relative RMS error,

$$\mathcal{E}_{\text{rRMS}}(\boldsymbol{\vartheta}) = 3.29\sqrt{\mathcal{C}^{(2\text{r})}(\boldsymbol{\vartheta})} = \sqrt{\int_0^\infty \mathrm{d}T\, P^{(0)}(T)\frac{\bar{\varepsilon}^2_{\text{RMS}}(T)}{T^2}}. \tag{3.19}$$

This bound is unfortunately not tight, which is shown by the dark shaded region in Fig. 3.6. For small N the relative errors are well captured by this bound but then consistently under estimated for larger N. For the equilibrium case with Jeffrey's prior, the integrals in Eq.(3.18) can be explicitly solved and the benchmark simplifies to,

$$\mathcal{E}_{\text{rRMS}} \geq \frac{3.29}{\sqrt{(N+1)\pi^2/8 - N + 1}} \approx \frac{6.81}{\sqrt{N + 9.56}}, \tag{3.20}$$

which depends only on the number of probes and not the energy gap of the probes, E.

3.1.8 Numerical comparison of priors

In this section we will briefly demonstrate the importance of selecting an appropriate prior when only few measurements are performed. To this end the average RMS deviation of the estimators relative to the true temperature is compared for three priors in Fig.3.7. the priors we compare are (a): $P(\theta) \propto \sqrt{\mathcal{I}(\theta)}$ which is the same as the one we have been using up to now, (b): $P(\theta) \propto 1/\theta$ and (c): $P(\theta) \propto 1$. Here we see that for the small range of temperatures where the thermometer is most sensitive, the effect of the prior has already been lost and the posterior distributions will all have converged to the same point and will thus yield the same errors. However, for temperatures outside of this range, the spread of the estimates is larger for both prior (b) and (c). On one hand, this should be expected for the scale invariant prior as it assumes information only about symmetries of the parameter and nothing about the likelihood. Therefore, in some sense it is more "uninformed". However, the constant prior is not physically motivated and leads to even more scattered results at high temperatures. One can see that in a scenario where it is not difficult to calculate, prior (a) may lead to more accurate estimates. However, as is the case in this study, the fully global approach where the cost function is determined by the prior can lead to unwieldly integrals when prior (a) is chosen. Thus heuristic choices of cost functions and their corresponding optimal estimators might be chosen instead.

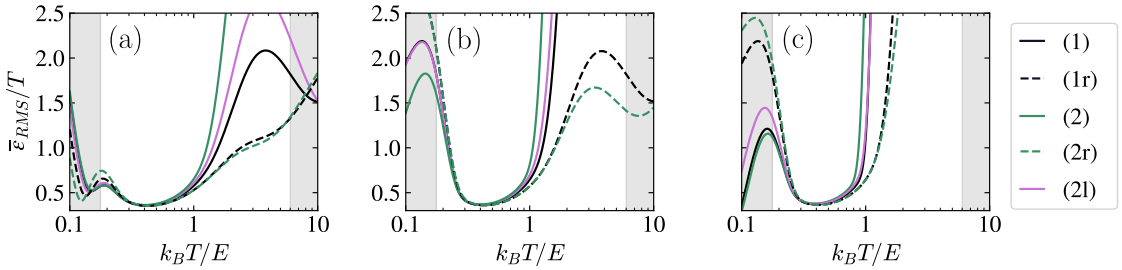


Figure 3.7: Influence of the prior distribution on the estimators' accuracies for 200 fully thermalized qubits as a function of actual temperature T in units E/k_B. We compare the RMS deviations from the true temperature T for (a) The prior $P^{(0)}(T) \propto 1/\sqrt{\mathcal{I}(\theta)}$, (b) $P^{(0)}(T) \propto 1/T$ and (c) the constant prior over the range $k_BT/E \in [0.01, 200]$. The grey bars delimit the detectable temperatures determined by Eq.(3.11). Adapted from [A].

3.1.9 Non-equilibrium probes

We finish this case study with an analysis of non-equilibrium probes where, the probes interact for a short time with the environment such that $\gamma\tau \lesssim 1$. First we consider the effect that shorter coupling times have on the sensitive range determined by Eq.(3.10). In Fig. 3.8 (a), the boundary values T_∞ and T_0 as well as their ratio is shown as a function of $\gamma\tau$ for $N = 200$ probes. This figure shows that size of the sensitivity range given by the difference of these two boundary values increases and shifts toward higher temperatures as τ decreases. Additionally, for equilibrium probes (large $\gamma\tau$), the range is constant. However, Fig. 3.8(b) shows that this increased range comes at the cost of local accuracy. Here we plot the maximum value of the Fisher information $\max_T \mathcal{I}_\tau(T)$ for a particular interaction time τ relative to the maximum of the equilibrium Fisher information, $\max_T \mathcal{I}_\infty(T)$. This quantity increases with $\gamma\tau$ which means the highest local sensitivity is achieved by equilibrium thermometry.

This result is also shown by numerical simulation in Fig. 3.9. Here we fix $N = 200$ but compare three interactions times. We compare the performance at different temperatures in terms of the average RMS deviation of equilibrium probes ($\gamma\tau \to \infty$), to probes that interact for short-times ($\gamma\tau = 0.1$ and $\gamma\tau = 0.01$) and have reached a non-equilibrium state. The solid part of each curve denotes the temperatures that fall within the sensitive range defined in Sec. 3.1.5. These results agree with the ones shown in Fig. 3.8. A longer interaction time leads to a

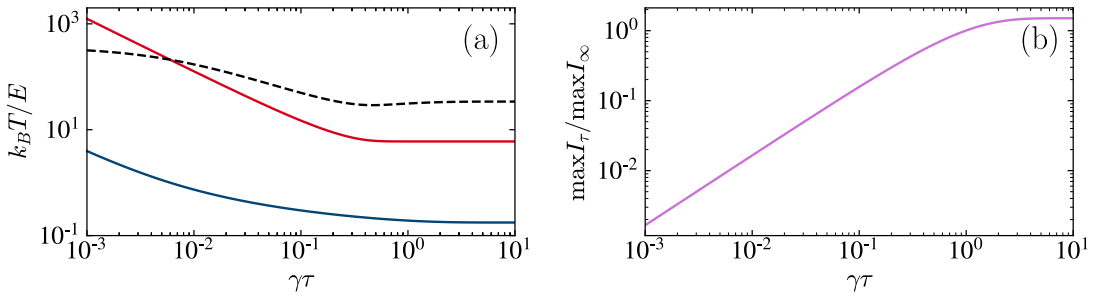


Figure 3.8: (a) Lowest and highest detectable temperature, T_0 (blue) and T_∞ (red) in units E/k_B determined from Eq.(3.11). The ratio T_∞/T_0 (dashed) as a function of thermalization time $\gamma\tau$ for 200 qubits is also shown. The range of temperatures to which the thermometer is sensitive increases for a shorter interaction time. This comes at the cost of local accuracy (b) since the highest attainable temperature resolution relative to the resolution at equilibrium, $\max_T \mathcal{I}_\tau(T)/\max_T \mathcal{I}_\infty(T)$ is reduced with decreasing $\gamma\tau$. Adapted from [A].

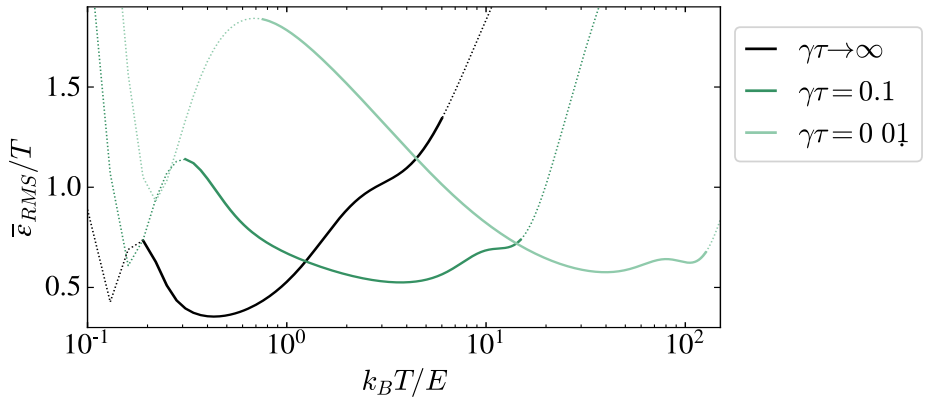


Figure 3.9: The average RMS deviation relative to the actual temperature for the relative mean estimator ϑ_{2r} (refer to Table 3.1) is shown for for 200 measurements at each temperature T in units E/k_B. We compare the case of fully thermalized probes (black line) to finite thermalization times, $\gamma\tau = 0.1$ (dark green) and $\gamma\tau = 0.01$ (light green). The solid stretches extend over the respective detectable temperature ranges given by (3.11). Adapted from [A].

higher sensitivity near the locally optimal temperature but a shorter interaction time extends the range of temperatures that the probes is sensitive to and increases the value of the temperature to which the probe is most sensitive.

This result is useful in two contexts, on the one hand, total experiment time instead of number of measurements made may be the resource that an experimenter wants to optimise. In this case, it makes sense to perform more measurements with a very short interaction time. On the other hand, at the beginning of the experiment, the range of temperatures that the reservoir could take might not be well localised, in this case, a shorter measurement time increases the range of temperatures that can be detected. This could be coupled with an adaptive scheme where the energy gap of the probe and the interaction time are changed once more information is obtained about the temperature so that advantage can be taken of the local sensitivity once enough data is collected to justify the local approach. The non-equilibrium results from this chapter inspire the thermometry strategy in Chapter 4. Here, a two-level probe in contact with the system of interest is monitored continuously, and therefore is not reset after each measurement which results in measurements that are not independent from each other. An adaptive strategy is employed to improve the sensitivity of the thermometer.

3.2 Estimation of the rate of decay by spontaneous emission

The third case study will focus on the systematic approach explained in sec.2.4.3. We will compare two choices of prior distribution on the one hand we use the prior based on physical invariances in the parameter and on the other using Jeffrey's prior based on the Fisher information. Then we discuss the performance of the associated Bayesian optimal measurements and estimators.

The example that we consider is the estimation of the rate of spontaneous emission into the vacuum of a photon from a two-level system prepared in a superposition state. This textbook example [68, 122] was already studied for the prior derived from symmetries of the parameter in [104, 105] but we revisit it here since it is a simple example that illustrates the differences between the estimation strategies obtained from different prior distributions and is similar to

the thermometry example studied above. As we will see, even in this simple example, the prior derived from the Fisher information can yield unwieldly cost functions and estimators and introduces conceptual problems when we attempt probe optimisation.

Once again we consider a qubit with free Hamiltonian $E/2(|1\rangle\langle 1| - |0\rangle\langle 0|)$ interacting dissipatively with an environment. We prepare the probe in the superposition state $|\psi(\eta)\rangle = \sqrt{1-\eta}\,|0\rangle + \sqrt{\eta}\,|1\rangle$ and detect if a photon is emitted into the environment by the probe or not. The emission of a photon by the probe happens at an unknown rate $\Gamma = 1/\tau$, where τ is the lifetime of the excited state. The standard Master equation [1] for the probe is,

$$\dot{\rho} = -i\frac{E}{2}[|1\rangle\langle 1| - |0\rangle\langle 0|, \rho] + \frac{1}{\tau}\mathcal{D}[|0\rangle\langle 1|]\rho \tag{3.21}$$

with $\mathcal{D}[\hat{A}]\rho = \hat{A}\rho\hat{A}^\dagger - \{\hat{A}^\dagger\hat{A}, \rho\}/2$ and the solution for the state of the probe at time t given the initial state $|\psi(\eta)\rangle$ is,

$$\begin{aligned}\rho_t(\tau) =&(1-\eta e^{-t/\tau})\,|g\rangle\langle g| + (\eta e^{-t/\tau})\,|e\rangle\langle e| \\ &+ \sqrt{\eta(1-\eta)e^{-t/\tau}}(|e\rangle\langle g| + |g\rangle\langle e|).\end{aligned} \tag{3.22}$$

This problem is similar to the rate estimation example that we opened the chapter with, the lifetime of the state τ is scaled by the total time t and hence the symmetry function and prior that are derived from the scale invariance of the parameter are once again, $P(\theta) \propto 1/\theta$ and $f_s(\theta) = \log(\theta)$.

However, we have a physical model of the state and thus can calculate the QFI of this state,

$$\mathcal{I}_Q(\theta) = \frac{t^2 e^{-t/\theta}}{\theta^4} + \frac{t^2\left(1-e^{-t/\theta}\right)}{\theta^4\left(e^{t/\theta}-1\right)^2} \tag{3.23}$$

to get the prior $P(\theta) \propto \sqrt{\mathcal{I}_Q(\theta)}$. Formally, the symmetry function obtained from the information geometry approach is,

$$f_I(z) = \int dz\sqrt{\mathcal{I}(z)} \tag{3.24}$$

[1] see App. A for how this is derived

In this case study this integral must be evaluated numerically. This is a minor drawback of this approach which was also encountered for non-equilibrium probes in Sec. 3.1.2.

In this section we will follow the systematic approach to obtain error estimates and point estimates and for ease of reference we repeat these functions here. We compare the approach based on f_s and f_I which yield error estimates,

$$c^{(s)}(\vartheta_x, \theta) = \log^2\left(\frac{\vartheta_x}{\theta}\right), \quad c^{(I)}(\vartheta_x, \theta) = \left|\int_{\theta'}^{\vartheta_x} \mathrm{d}z\sqrt{\mathcal{I}(z)}\right|^2 = (f_I(\vartheta_x) - f_I(\theta'))^2, \tag{3.25}$$

and point estimates,

$$\vartheta_x^{(\mathrm{s})} = \theta_u \exp\left[\int d\theta P(\theta|x) \log\left(\frac{\theta}{\theta_u}\right)\right] \quad \vartheta_x^{(\mathrm{I})} = f_I^{-1}\left[\int d\theta P(\theta|x) f_I(\theta)\right], \tag{3.26}$$

where θ_u is a constant with units of t introduced so that the argument in the logarithm is dimensionless. For the approach based on the Fisher information, the accuracy of the point estimator and error estimate are dependent on the accuracy of the integrals for each z in the range of times in the prior. We can now compare the performance of each of these examples on the same data set to determine if the conceptual differences lead to actual differences in performance for this example.

First, in Fig 3.10(a) we compare the dependence of the prior distributions obtained from each approach. Here we see that for $\tau/t \to 0$, the distributions have very different behaviour, the scale prior diverges at this point while the prior obtained from f_I decays to zero. Next, recall that globally optimal error $\bar{\epsilon}^{(f)}_{\mathrm{min}} = \bar{\epsilon}^{(f)}_p - \mathcal{G}^{(f)}$ (Eq. (2.38)) can be written in terms of the variance of the prior Eq. (2.39) and the precision gain Eq. (2.40). We can quantify the effect of the prior on an estimation strategy by looking at the relative precision gain, $\bar{\epsilon}^{(f)}_{\mathrm{min}}/\bar{\epsilon}^{(f)}_p = 1 - \mathcal{G}/\bar{\epsilon}^{(f)}_p$. In Fig 3.10(b) the relative precision gain of each approach is shown in terms of the integration interval the integration interval $[1/a, a]$. We see that for a small interval the precision gain is similar for both approaches but for a larger interval which is more consistent with a truly global approach, the approach based on the scale invariant function f_s has a higher intrinsic precision gain. This suggests that this approach assumes less information *a priori* than the approach based on information geometry.

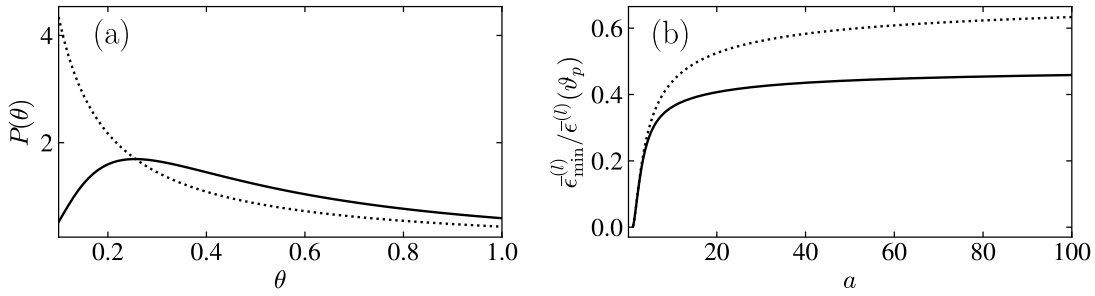


Figure 3.10: Here we compare the prior distributions and study the effect of the prior range for priors derived from $f = \log(\theta)$ (dotted line) and $f = \sqrt{\mathcal{I}_Q(\theta)}$ (solid line) in both sub-figures. In (a) The priors are shown as a function of θ. (b) The precision gain from a single measurement relative to the prior uncertainty is shown as a function of the prior range, $1/a < \theta < a$. For a small range both parameterisations of the prior perform similarly as their dependence on θ is similar for large values of θ. However, as the prior range is expanded, the prior derived from $f = \log(\theta)$ has a higher precision gain. To be published in [D].

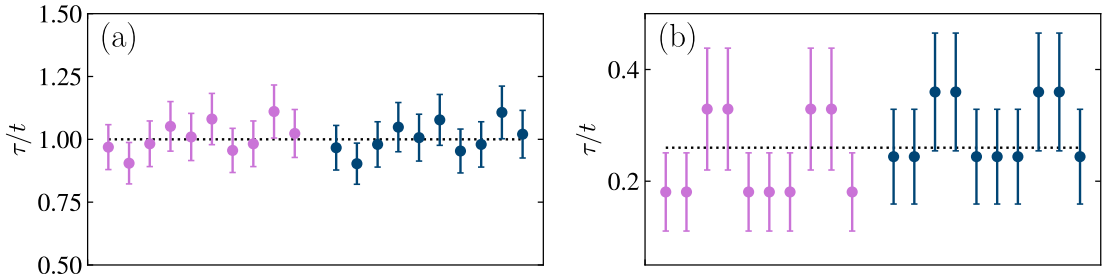


Figure 3.11: The variability of the estimates of the lifetime τ obtained through the the approach with a prior derived from $f = \log(\theta)$ (purple points) and $f = \sqrt{(\mathcal{I}(\theta))}$ (dark blue points) with initial state prepared with superposition parameter $\eta = 1$ and measurements performed in the energy basis of the qubit. Ten rounds of (a) $N = 200$ measurements and (b) $N = 20$ measurements were simulated for an actual scaled lifetime of (a) $\tau/t = 1$ and (b) $\tau/t = 0.26$, given a fixed measurement time. The average estimate and error obtained from the posterior were determined from each approach on the same data. The actual value of the lifetime for fixed time t is shown by a dotted line in each figure. To be published in [D].

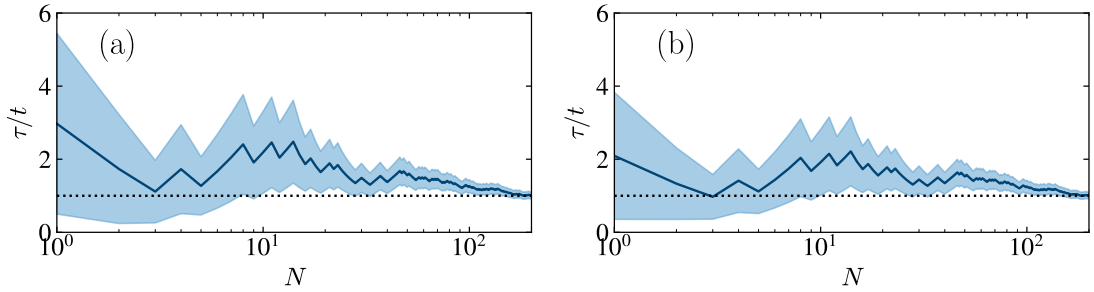


Figure 3.12: . The estimated lifetime scaled lifetime τ/t for fixed time t is shown here as a function of the number of measurements by the blue line. The estimated error ϵ is given by the blue shaded regions. Here we compare the two approaches applied to the same simulated data derived from different priors $f = \log(\theta)$ (a) and $f = \sqrt{\mathcal{I}(\theta)}$ (b). The actual value of the lifetime is shown by a dotted line in each figure. To be published in [D].

Next, we compare the performance of the two approaches without measurment optimisation. We prepare the probe in the excited state and measure in the energy basis of the probe. In Fig 3.11 we compare the estimates along with their error bars for ten rounds of (a) $N = 200$ and (b) $N = 20$ measurements at a fixed measurement time τ/t. We also show in Fig 3.12 how the estimate converges to the actual value of the lifetime over the course of a single run of $N = 200$ measurements. These results show that the difference in the two approaches almost disappears for a large number of measurements. For the case where fewer measurements are made, the approach based on the Fisher information has estimates that are skewed towards higher values of τ and the spread in estimates is slightly larger for the scale invariant approach. This is due to the dominance of the prior for small data sets and that small values of θ dominate the scale invariant prior, which is not the case for the prior derived from the Fisher information. This difference disappears for larger data sets.

Finally we study how probe and measurement optimisation affects the estimation by finding the globally optimal quantum measurement based on the prior for both priors. For both approaches this POVM will depend on the value of the superposition parameter η that was chosen

for the initial state. The optimal measurement is determined by solving the equation

$$\mathcal{S}\int d\theta P(\theta)\rho(\theta) + \int d\theta P(\theta)\rho(\theta)\mathcal{S} = 2\int d\theta P(\theta)\rho(\theta)f(\theta), \tag{3.27}$$

and finding the eigenvectors of the operator $\mathcal{S}$.

Note that the prior based on the Fisher information depends on the state and therefore, so does the prior. Thus, by optimising the initial state of the probe, we are in turn optimising the prior and cost function. This introduces a major conceptual problem. The prior should be specified based on the available prior information and assumptions of the experimenter and optimising it after it has already been specified is a circular approach. Therefore, we will only consider probe optimisation based on the prior $P(\theta) \propto 1/\theta$ since this prior does not depend on the state and therefore does not suffer from this same conceptual problem.

First, we determine the optimal initial state of the probe. We do this by determining the value of η that minimises the average error $\bar{\epsilon}$. For the prior determined by $f(\theta) = \log(\theta)$, the prior error does not depend on the initial state. Therefore the prior variance ϵ_p is fixed and we need to maximise the precision gain $\mathcal{G}$. The maximum over the global range $0 < \theta < \infty$ is achieved by $\eta = 1$. In a scenario where the prior range is restricted, for example in the local estimation limit, the maximum gain will depend on this range and a smaller gain is obtained for a smaller prior range of θ. This is shown in Fig 3.13(a). Here the prior range is restricted to $1/a < \theta < a$ for three choices of a and the gain is plotted with respect to the parameter η. This result reflects the fact that when the prior is restricted to a small range, the assumption that has been made is that the value of the parameter is already well known and thus is well captured by the prior error.

We now compare the globally optimal measurement strategy to a variety of strategies that deviate from the globally optimal one in different ways. First, instead of performing the optimal measurement by projecting onto the eigenbasis of $\mathcal{S}$, we perform the following physically motivated measurement with POVM elements $M_Y = (1 - e^{-2t/\tau_0})\,|e\rangle\langle e|$ (for the detection of a photon at time t) and $M_N = e^{-2t/\tau_0}\,|e\rangle\langle e| + |g\rangle\langle g|$ (for no detection of a photon) [122]. This measurement depends on the actual value of τ which we denote τ_0. A better accuracy is expected when $\tau_0/t \ll 1$ rather than for $\tau_0/t \gg 1$ as the former corresponds to the situation where it is likely that a photon has already been emitted since the lifetime is short [105]. This is compared to the global minimum error (purple dotted line) and prior variance (black dotted

line) which represent the best and worst case scenarios of any estimation strategy in Fig 3.13(b). We see that the physically motivated measurement (solid purple curve) is only able to achieve the minimum error for small τ_0.

Additionally, we study the case where the probe is not optimised. This is depicted in Fig 3.13(b) with a suboptimal superposition parameter $\eta = 1/3$. The minimum error, $\epsilon_{\rm sub}$, for suboptimal η is of course larger (red dotted line in the figure) and the physically motivated measurement with suboptimal η (red solid curve) performs in a similar way to the one for an optimal probe where it is most informative for small τ_0 but for larger τ_0 the gain in precision is not much better than is expected from the prior variance.

Another POVM that we could use is the locally optimal measurement, that is a projective measurement onto the eigenbasis of the symmetric logarithmic derivative in Eq (2.4). This measurement also depends on τ_0. We compare this strategy for $\eta = 1/3$ to the physically motivated measurement and suboptimal minimum error in figure Fig 3.13(b). Here we see that this strategy is able to reach the bound $\epsilon_{\rm phys}$ for some values of θ and does perform better than a physically motivated POVM but not as well as the fully global approach.

3.3 Conclusions

In this Chapter we have explored two examples of Bayesian estimation of environmental parameters. In both cases, the parameters can be identified as scale parameters in the system and thus we can derive the same prior for all of them based on this symmetry.

In the first case study we consider quantum thermometry. Here a two level probe interacts with a Bosonic bath of unknown temperature and measurements are made on either a fully or partially thermalised probe. When we determine the cost function and estimator from the prior distribution, we find that the optimal cost function for equilibrium probes is a non-trivial but still an analytic function. Although this function can in principle be computed numerically, we focus for the rest of this case study on the heuristic approach where the prior and cost functions are chosen separately. Thus, we compare different cost functions and their estimators and fix the prior distribution to be $P(\theta) \propto \sqrt{I(\theta)}$. We find that the commonly chosen mean, median and mode considerably over estimate the error at high temperatures. However, the two novel

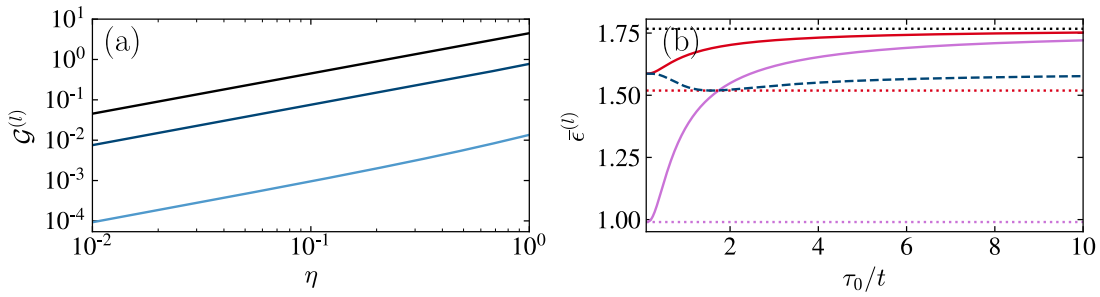


Figure 3.13: The approach using the prior derived from $f = \log(\theta)$ can be used to get a optimal measurement.(a) The improvement in the actual error over the prior variance quantified by the gain $\mathcal{G}$ monotonically increases with the possible values of η with a maximum gain achieved for $\eta = 1$. In this figure the prior range, $1/a < \theta < a$ is varied with $a = \{2, 10, 100\}$ going from the light to dark blue lines. (b) The prior uncertainty (black dotted line) and minimum error $\bar{\epsilon}_{\min}$(light purple dotted line) define the worst and best case scenario after one measurement. Here three different estimation strategies are compared to these bounds for varying actual values of the lifetime τ/t. The solid lines show the behaviour of the average error when the measurement is made with the physical POVM defined by $M_Y = (1 - e^{-2t/\tau_0})\,|e\rangle\langle e|$ (for the detection of a photon at time t) and $M_N = e^{-2t/\tau_0}\,|e\rangle\langle e| + |g\rangle\langle g|$. The probe is prepared with a state with either the globally optimal superposition parameter $\eta = 1$ (purple curve) or $\eta = 1/3$ (red curve). The dashed blue curve shows the error obtained when the locally optimal POVM predicted by the QFI is used with $\eta = 1/3$. Finally, the best possible error that can be obtained by the globally optimal POVM with $\eta = 1/3$ is shown by the red dotted line. To be published in [D].

relative errors Eq. (2.4.1) that are presented in Chapter 2 achieve a consistently better accuracy in quantifying the actual variability of the data, throughout the range of the prior distribution. We also compare all error estimates for either the scale invariant prior or the constant prior and find that the divergence problem is worse at high temperatures.

The error estimates are also compared to the CRB and relative van-Trees bound. We find that although it is not a tight bound, the van-Trees bound captures the scaling of the error with number of measurements well in the limit of little data. However, for large amounts of data, the error scales as expected with the CRB. Finally, using the sensitivity region that was derived in Sec. 3.1.5, we were able to determine that by using partially thermalised probes, the range of temperatures that the probe is sensitive to is extended. This comes at the cost of maximum accuracy. These results motivate the continuous measurement scenario explored in Chapter 4 where an adaptive strategy is also studied.

In the second case study we consider the estimation of the lifetime of the excited state of a qubit. This is a simple estimation scenario adapted from one of the case studies in [105]. Here we apply the formal approach to the selection of cost functions from the priors to use it to contrast the prior derived from the scale invariance of the lifetime with the prior derived from the Fisher information. In this case study we find that the estimated error in both approaches is very similar when averaged over many experiments. However, the calculation of the cost function and estimator has to be done numerically when the prior derived from the Fisher information is used. We also study measurement and probe optimisation. Here we uncover a conceptual difficulty in optimising the probes or potentially the POVM for a prior based on the Fisher information. Optimising the initial state of the probe or the POVM changes the likelihood function which in turn will change the prior distribution. This should not happen since the prior is supposed to be fixed at the start of the problem. The prior derived from physical invariances of the probe does not depend on the likelihood and thus probe optimisation does not pose this same problem there.

The case studies in this chapter show how each approach to Bayesian estimation can be used while highlighting potential pitfalls. In general the Bayesian approach where the choice of prior distribution determines the cost functions and estimators, may be preferred to the approach where these choices are made independently since it leads to a clear way to optimise the probes and POVM and also leads to global Bayesian bounds on the error. However, depending on the

example at hand, there are several pragmatic choices that may need to be made. For example, if an underlying symmetry of the parameter cannot be found, the prior derived from the Fisher information may be a suitable default option. However, this can lead to cost functions and estimators that are difficult to calculate and in this case the standard choices explored in the thermometry case study might be more suitable. On the other hand when symmetries are clear as they are in these case studies, the prior derived from these symmetries and the associated cost functions are often simpler to calculate. These results could be extended in future to multiparameter scenarios where the optimal Bayesian POVM and bounds may not be the same. Additionally, it may be of interest to find estimation strategies for environmental parameters with multiple free parameters that can enter into the strategy for example by including interactions between probes and adaptive measurements.

Chapter 4

Parameter estimation by continuous monitoring

In the previous chapter, the main examples of Bayesian estimation were for discrete measurements, where a probe is prepared and subsequently measured. In this chapter the same framework is applied to continuous measurements, based on the work published in [B].

4.1 Motivation

In this chapter, we consider a scenario where the temperature information is obtained from the probe continuously while the probe remains in contact with the environment. Rather than quantifying the performance of the protocol in terms of number of measurements, we will be concerned with the total measurement time. Continuous monitoring of a probe allows for measurements that are less invasive than measuring the probe and completely resetting the state. This technique has been applied in several other contexts, in particular with measurement feedback, for state stabilisation [129,130] and for magnetometry both experimentally, [131–133] and theoretically [88, 134, 135]. The fundamental limits of Bayesian estimation for continuous measurements in quantum systems have also been explored [136–138]. Furthermore, the temperature is an environmental parameter and thus its estimation is fundamentally different from the estimation of a magnetic field or other Hamiltonian parameters. We will also make use of the

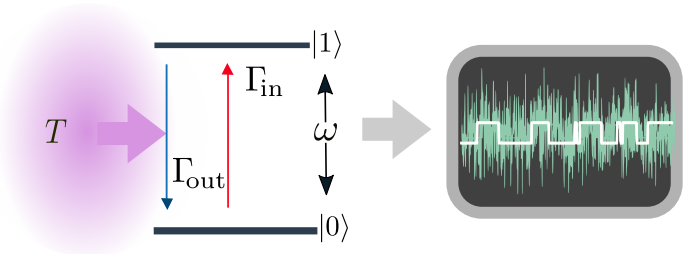


Figure 4.1: In the continuous thermometry scenario that we consider, a two-level probe interacts with a bath of unknown temperature T. The coupling to the bath induces transitions between the levels of the probe with rates Γ_{in} and Γ_{out}. The energy gap between the two levels can be changed to improve the sensitivity of the probe to a particular temperature. The state of the probe evolves in a telegraph-like trajectory shown by the white line. However, in an experiment involving weak continuous measurements, the actual signal observed is noisy with a finite bandwidth as shown by the green line. Adapted from [B].

Bayesian framework to describe adaptive strategies to improve the temperature sensitivity.

4.2 Model

Our model probe system is the same as the one considered in previous discrete model discussed in Chapter 3. It consists of a two-level probe prepared in a state with no coherences which stochastically jumps between the ground and excited states (denoted $|0\rangle$ and $|1\rangle$) due to its interaction with the bath at unknown temperature T. This is illustrated in Fig. 4.1. Again, the probe has Hamiltonian $H_S = \omega/2(|1\rangle\langle 1| - |0\rangle\langle 0|)$. By continuously monitoring the state of the probe, we will get information about the temperature of the bath because the rate with which the probe transitions from the ground to the excited state and vice versa is determined by this temperature. We assume that we are able to detect these jumps without measurement back-action, that is, the measurement of the probe does not affect its state. This is achieved by measuring in the energy basis of the probe [139].

First, we define the dynamics of the probe state. The density matrix will be diagonal at all times since the probe is prepared in an initial state with no coherences. Therefore, we need only consider the populations of the density matrix. The probability that the probe is in the ground state at some time t is $p_0(t)$ and similarly we denote the excited state probability by $p_1(t) =$

$1 - p_0(t)$. The rate of transition from the ground state into the excited state is generically given by Γ_{in} and similarly the rate of transition from the excited state to the ground state is denoted Γ_{in}. We can describe the dynamics with the classical rate equation,

$$\partial_t \begin{pmatrix} p_0(t) \\ p_1(t) \end{pmatrix} = \begin{pmatrix} -\Gamma_{\text{in}} & \Gamma_{\text{out}} \\ \Gamma_{\text{in}} & -\Gamma_{\text{out}} \end{pmatrix} \begin{pmatrix} p_0(t) \\ p_1(t) \end{pmatrix}, \tag{4.1}$$

which can be obtained by simplifiying the equation derived for a generic qubit A.24 to the present situation where only diagonal elements are relevant. Given the initial state $p_1(0) = p_1$, we solve this equation to give the probability to be in the excited state in terms of this initial state, the rates and the elapsed time,

$$p_1(t) = e^{-(\Gamma_{\text{in}}+\Gamma_{\text{out}})t} \left[p_1 - \frac{\Gamma_{\text{in}}}{\Gamma_{\text{in}} + \Gamma_{\text{out}}} \right] + \frac{\Gamma_{\text{in}}}{\Gamma_{\text{in}} + \Gamma_{\text{out}}}. \tag{4.2}$$

The rate equation and solution above apply generically for a system that transitions into and out of a state with two different rates. However, in the physical picture of a two-level system in contact with a thermal reservoir, this can be made more specific. The transitions that happen in the two-level system are due to the exchange of energy with the thermal bath and therefore the rates of the transitions must obey the local detailed balance relation,

$$\frac{\Gamma_{\text{in}}}{\Gamma_{\text{out}}} = e^{-\beta\omega}. \tag{4.3}$$

Here, $\beta = 1/k_B T$ is the inverse temperature. We set the Boltzmann constant $k_B = 1$ and note that ω is the energy difference between the two states of the probe. We will consider two commonly used bath models. The first is a Bosonic thermal bath with with an Ohmic spectral density $\kappa(\omega) = \kappa'\omega$ with ω below the cut off frequency, the rates of this bath in terms of the Bose-Eistein distribution $n_B(\omega) = [e^{\beta\omega} - 1]^{-1}$ are,

$$\Gamma_{\text{in}} = n_B(\omega)\kappa(\omega), \quad \Gamma_{\text{out}} = [n_B(\omega) + 1]\kappa(\omega). \tag{4.4}$$

The second bath model is a bath of Fermions with a flat spectral density that is thus, independent of the frequency.The Fermionic bath rates in terms of the Fermi-Dirac distribution $n_F(\omega) =$

$[e^{\beta\omega}+1]^{-1}$ are,

$$\Gamma_{\text{in}} = n_F(\omega)\Gamma, \quad \Gamma_{\text{out}} = [1 - n_F(\omega)]\Gamma. \tag{4.5}$$

While we focus on thermometry in this chapter, this technique could be applied to any situation where the parameter of interest is encoded in the rates Γ_{in} and Γ_{out} like the bath spectral density $\kappa(\omega)$ [140]. Additionally, this setup could be realised in a variety of experimental scenarios, for example, the probe could be a quantum dot that is continuously monitored by coherent reflection spectroscopy [141].

4.2.1 Probability to measure a particular trajectory

In an ideal experiment over the time interval $[0, \tau]$, the apparatus will continuously record a trajectory $\nu_\tau = \{n(t)|t \in [0,\tau]\}$, where $n(t) \in \{0,1\}$ denotes the occupation of the state at all times t. Generically, a trajectory will resemble a series of step functions with the probe in some initial state, then jumping to another state after some time, remaining there for some time and jumping back and so on. We can derive the probability density that we observe this trajectory given the transition rates as follows,

First, discretise time into steps $\delta t = \tau N$. The discrete trajectory, $\nu_N = [n_j]_{j=0}^N$, is then a sequence of snapshots, $n_j = n(j\delta t)$, of whether the system is in state 0 or 1. This process is Markovian and therefore the likelihood to observe a particular trajectory is,

$$P(\nu_N|T) = p_0 \prod_{j=0}^{N-1} p(n_{j+1}|n_j) = (p(1|0))^k \, (p(0|1))^l \, (p(0|0))^q \, (p(1|1))^r \, . \tag{4.6}$$

Here, k, l, q, r denote the number of times a particular transition is observed from one state to another. For example, a transition from state 0 to state 1 is observed k times and the situation where the state was intially in state 0 and was observed in state 0 after one time step was observed q times. Now, we integrate the rate equation, Eq. (4.1) over time δt assuming the initial state is either 0 or 1, to get the following conditional probabilities,

$$p(1|0) = \frac{\Gamma_{\text{in}}}{\Gamma_{\text{in}} + \Gamma_{\text{out}}} \left[1 - e^{-(\Gamma_{\text{in}}+\Gamma_{\text{out}})\delta t}\right]$$

$$p(0|1) = \frac{\Gamma_{\text{out}}}{\Gamma_{\text{in}} + \Gamma_{\text{out}}} \left[1 - e^{-(\Gamma_{\text{in}}+\Gamma_{\text{out}})\delta t}\right]$$
$$p(0|0) = 1 - p(1|0), \qquad p(1|1) = 1 - p(0|1). \tag{4.7}$$

In the limit, $\delta t \to 0$, in which we recover, $\nu_N \to \nu_\tau$, we can expand the conditional probabilities given above to linear order in δt,

$$p(1|0) = \delta t \Gamma_{\text{in}}, \qquad p(0|1) = \delta t \Gamma_{\text{out}}, \tag{4.8}$$

as long as $(\Gamma_{\text{in}} + \Gamma_{\text{out}})\delta t \ll 1$, which occurs for small δt and finite temperatures. Substituting this back into the discretised likelihood gives us,

$$p(\nu_\tau|T) = (\delta t \Gamma_{\text{in}})^k (\delta t \Gamma_{\text{out}})^l (1 - \delta t \Gamma_{\text{in}})^q (1 - \delta t \Gamma_{\text{out}})^r . \tag{4.9}$$

In the limit $\delta t \to 0$, the probability for the state to jump is infinitessimally smaller that the probability to stay in the same state therefore, $k, l \ll q, r$. This means we can write the total time spent in state 1 as $r\delta t = \tau_1$ and similarly the total time in the ground state as $q\delta t = \tau - \tau_1$. Thus, in this limit we get,

$$(1 - \delta t \Gamma_{\text{out}})^r = (1 - \tau_1 \Gamma_{\text{out}}/r)^r \to e^{-\Gamma_{\text{out}}\tau_1}. \tag{4.10}$$

By appling the same logic to the term involving q we obtain,

$$\mathcal{P}(\nu_\tau|T) := \frac{1}{(\delta t)^{l+k}}\, p(\nu_\tau|T) = p_0 \Gamma_{\text{in}}^k \Gamma_{\text{out}}^l e^{-\Gamma_{\text{in}}(\tau-\tau_1)-\Gamma_{\text{out}}\tau_1}. \tag{4.11}$$

Note that because t is a continuous variable, $\mathcal{P}(\nu_\tau|T)$ is a probability density with units that will depend on the number of jumps that are observed in a given interval.

4.3 Bayesian approach

We will analyse the performance of this thermometer using the Bayesian techniques that have been developed so far in this thesis (see Chapter 2). The main goal is to determine the asymptotic

scaling of two measurement approaches. One one hand, we will prepare globally optimal probes and on the other the probes will be adapted during the experiment. As we will have to simulate many trajectories and the formal aspects of Bayesian estimation are already well explored in the previous section, we will employ the following pragmatic approach.

First, we select the simplified relative square cost function as out posterior averaged error measure, that is,

$$\varepsilon^{(2r)}(\vartheta, \nu_\tau) \coloneqq \int d\theta P(\theta|\nu_\tau) \left(\frac{\vartheta(\nu_\tau) - \theta}{\theta}\right)^2. \tag{4.12}$$

We will consider this error averaged over all possible trajectories to get a figure of merit that is independent of the trajectory,

$$\bar{\varepsilon}^{(2r)}(\vartheta) \coloneqq \int d\nu_\tau \rho(\nu_\tau) \varepsilon^{(2r)}(\vartheta, \nu_\tau). \tag{4.13}$$

The optimal estimator is then (see App B),

$$\vartheta_{\mathrm{R}}(\nu_\tau) \coloneqq \frac{\int \mathrm{d}\theta P(\theta|\nu_\tau)/\theta}{\int \mathrm{d}\theta P(\theta|\nu_\tau)/\theta^2}. \tag{4.14}$$

We assume for the sake of simplicity in our simulations that the temperature is restricted to the range $T_{\min} \leqslant T \leqslant T_{\max}$ and that within this range all temperatures are equally likely. This means that the distribution $P(\nu_\tau)$ is also constant in this range.

4.4 Fisher information

Although we will focus on Bayesian analysis of the data, we will also quantify the asymptotic accuracy of this measurement scheme. In the limit of a long measurement time the Bernstein von Mises theorem applies even for this Markov process [142] so that the posterior distribution will be a sharply peaked Gaussian about the actual value of the temperature and therefore, the Bayesian Cramér-Rao bound will apply. For convenience, we repeat the Cramér-Rao bound for

the relative Bayesian error which we define in Sec. 2.8

$$\bar{\epsilon}^{(2\mathrm{r})}(\vartheta_x) \geq \int \mathrm{d}\theta \frac{P(\theta)}{\theta^2 \mathcal{I}(\theta)}. \tag{4.15}$$

To properly quantify this bound, we must derive the Fisher information of our trajectory. We introduce the following notation to make this calculation more convenient: denote the initial state of the probe by, $\boldsymbol{p} = (p_0, p_1)^{\mathrm{T}}$ and let K be the 2×2 matrix elements $K_{nm} = p(n|m)$. We also introduce the column vector $\boldsymbol{F}^{(N)}$ with elements $F_n^{(N)} = \langle \delta_{n_N,n} [\partial_T p(\nu_N|T)]^2 \rangle$. Note that we can write the likelihood iteratively as, $p(\nu_N|T) = p(n_N|n_{N-1}) p(\nu_{N-1}|T)$ with $p(\nu_0|T) = p_{n_0}$, we get the following expression from the Fisher information,

$$\begin{aligned}
\mathcal{I}[p(\nu_N|T)] &= (1,1)\boldsymbol{F}^{(N)} \\
&= \sum_{n_N} \sum_{\nu_{N-1}} p(n_N|n_{N-1}) p(\nu_{N-1}|T) \left\{ \partial_T \ln \left[p(n_N|n_{N-1}) p(\nu_{N-1}|T) \right] \right\}^2 \\
&= \sum_{n_N} \sum_{\nu_{N-1}} p(n_N|n_{N-1}) p(\nu_{N-1}|T) \left[\partial_T \ln p(n_N|n_{N-1}) + \partial_T \ln p(\nu_{N-1}|T) \right]^2 \\
&= (1,1)\mathsf{J}\mathsf{K}^{N-1}\boldsymbol{p} + 2(1,1)(\partial_T \mathsf{K})(\partial_T \mathsf{K}^{N-1}\boldsymbol{p}) + (1,1)\mathsf{K}\boldsymbol{F}^{(N-1)}
\end{aligned} \tag{4.16}$$

where, in the last line we introduce the matrix J with elements $J_{nm} = K_{nm}(\partial_T \ln K_{nm})^2$. Now using the fact that K is a stochastic matrix with $(1,1)\mathsf{K} = (1,1)$ and $(1,1)\partial_T \mathsf{K} = 0$, we can simplify this expression and iterate this process down to the initial probe state,

$$\mathcal{I}[p(\nu_N|T)] = (1,1)\mathsf{J}\mathsf{K}^{N-1}\boldsymbol{p} + \mathcal{I}[p(\nu_{N-1}|T)] = \ldots = (1,1)\mathsf{J}\sum_{j=0}^{N-1} \mathsf{K}^j \boldsymbol{p} + \mathcal{I}[p_{n_0}]. \tag{4.17}$$

Note that depending on probe preparation the initial state may also depend on the temperature. Next, we use that the matrix K has right eigenvalues, $\lambda_1 = 1$ and $\lambda_2 = e^{-(\Gamma_{\mathrm{in}}+\Gamma_{\mathrm{out}})\delta t}$ with corresponding (unnormalized) right-eigenvectors $\boldsymbol{v}_1 = (\Gamma_{\mathrm{out}}, \Gamma_{\mathrm{in}})^{\mathrm{T}}$ and $\boldsymbol{v}_2 = (1,-1)^{\mathrm{T}}$ to get

the following expression,

$$\mathsf{K}^j \boldsymbol{p} = \frac{\boldsymbol{v}_1}{\Gamma_{\rm in}+\Gamma_{\rm out}} + \left(\frac{\Gamma_{\rm in}}{\Gamma_{\rm in}+\Gamma_{\rm out}} - p_1\right) e^{-(\Gamma_{\rm in}+\Gamma_{\rm out})j\delta t}\boldsymbol{v}_2, \tag{4.18}$$

which holds also for $j = 0$.

We substitute this into (4.17) and carrying out the geometric sum in the first term to get,

$$\mathcal{I}[p(\nu_N|T)] = \frac{(1,1)\mathsf{J}}{\Gamma_{\rm in}+\Gamma_{\rm out}}\left[N\boldsymbol{v}_1 + \frac{1-e^{-(\Gamma_{\rm in}+\Gamma_{\rm out})N\delta t}}{1-e^{-(\Gamma_{\rm in}+\Gamma_{\rm out})\delta t}}(\Gamma_{\rm in}p_0 - \Gamma_{\rm out}p_1)\boldsymbol{v}_2\right] + \mathcal{I}[p_{n_0}]. \tag{4.19}$$

The vector $(1,1)\mathsf{J}$ contains the Fisher information of the transition probabilities,

$$(1,1)\mathsf{J} = (\mathcal{I}[p(n|0)], \mathcal{I}[p(n|1)]) = \left(\frac{[\partial_T p(1|0)]^2}{p(1|0)[1-p(1|0)]}, \frac{[\partial_T p(0|1)]^2}{p(0|1)[1-p(0|1)]}\right). \tag{4.20}$$

In the continuum limit $\delta t \to 0$, the transition probabilities are given by eq.(4.8), which we substitute in to give,

$$(\mathcal{I}[p(n|0)], \mathcal{I}[p(n|1)]) = \left(\left[\frac{\partial_T \Gamma_{\rm in}}{\Gamma_{\rm in}}\right]^2 \Gamma_{\rm in}\delta t, \left[\frac{\partial_T \Gamma_{\rm out}}{\Gamma_{\rm out}}\right]^2 \Gamma_{\rm out}\delta t\right). \tag{4.21}$$

Inserting (4.21) into the Fisher information expression (4.19) and expanding to linear order in δt at fixed $\tau = N\delta t$, we arrive at

$$\begin{aligned}\mathcal{I}[\tau,T] = \mathcal{I}[p_{n_0}] &+ \left[p_0\frac{(\partial_T\Gamma_{\rm in})^2}{\Gamma_{\rm in}} + p_1\frac{(\partial_T\Gamma_{\rm out})^2}{\Gamma_{\rm out}}\right]\frac{1-e^{-(\Gamma_{\rm in}+\Gamma_{\rm out})\tau}}{\Gamma_{\rm in}+\Gamma_{\rm out}} \\ &+ \frac{\Gamma_{\rm in}\Gamma_{\rm out}}{\Gamma_{\rm in}+\Gamma_{\rm out}}\left[\left(\frac{\partial_T\Gamma_{\rm in}}{\Gamma_{\rm in}}\right)^2 + \left(\frac{\partial_T\Gamma_{\rm out}}{\Gamma_{\rm out}}\right)^2\right]\left[\tau - \frac{1-e^{-(\Gamma_{\rm in}+\Gamma_{\rm out})\tau}}{\Gamma_{\rm in}+\Gamma_{\rm out}}\right],\end{aligned} \tag{4.22}$$

Here, the large data limit corresponds to long times. Hence, we consider the long time Fisher information. To do this, we drop all the terms in Eq. (4.22) that do not grow with the total time τ,

$$\mathcal{I}[\tau,T] = \frac{\Gamma_{\rm in}\Gamma_{\rm out}}{\Gamma_{\rm in}+\Gamma_{\rm out}}\tau\left[\left(\frac{\partial_T\Gamma_{\rm in}}{\Gamma_{\rm in}}\right)^2 + \left(\frac{\partial_T\Gamma_{\rm out}}{\Gamma_{\rm out}}\right)^2\right]. \tag{4.23}$$

Note that these two results can be applied to the estimation of any parameter that is encoded in a rate. The derivatives of the rates with respect to temperature would just need to be replaced by derivatives with respect to the other parameter. Using the expressions that we have for the rates (4.4) and (4.5), we get the long time Fisher information for interaction with either a Bosonic,

$$\mathcal{I}[\tau, T] = \kappa(\omega)\tau \frac{\omega^2}{8k_B^2 T^4} \frac{\cosh(\beta\omega)}{\sinh^3(\beta\omega/2)\cosh(\beta\omega/2)} \tag{4.24}$$

or Fermionic,

$$\mathcal{I}[\tau, T] = \Gamma\tau \frac{\omega^2}{8k_B^2 T^4} \frac{\cosh(\beta\omega)}{\cosh^4(\beta\omega/2)} \tag{4.25}$$

reservoir.

4.5 Adaptive vs non-adaptive strategies

Our next task is to design the measurement protocol to be as sensitive as possible to the temperature. The first thing that might come to mind, based on the approach taken in the previous chapter, is optimising the initial state of the probe. However, especially given that we do not consider any coherences in this example, the effect of preparing the state in the ground or excited state is vanishingly small for long measurement times. The size of the energy gap, however is vitally important in determining the temperature with the maximum accuracy even at long times. This is apparent since the long time Fisher information Eq. (4.23) still depends on the energy gap. The strategy that we employ is to optimise the gap for all possible true temperatures according to the prior distribution so that asymptotically we achieve the smallest possible value of the average posterior error, $\bar{\varepsilon}^{(2r)}[\vartheta]$. We do this by finding the asymptotically optimal energy gap, i.e., the one that minimises,

$$\omega^*_{\text{n.ad.}} := \arg\min_{\omega} \int \frac{dT\rho(T)}{T^2\mathcal{I}[\tau, T]}. \tag{4.26}$$

Note that although this is an asymtotically optimal strategy because we use the Bayesian Cramér-Rao bound to derive it, it is optimal on average over the entire prior range and thus is global in that sense. In Fig. 4.2 the blue solid line shows the average relative error (Eq. (4.13)) for a probe

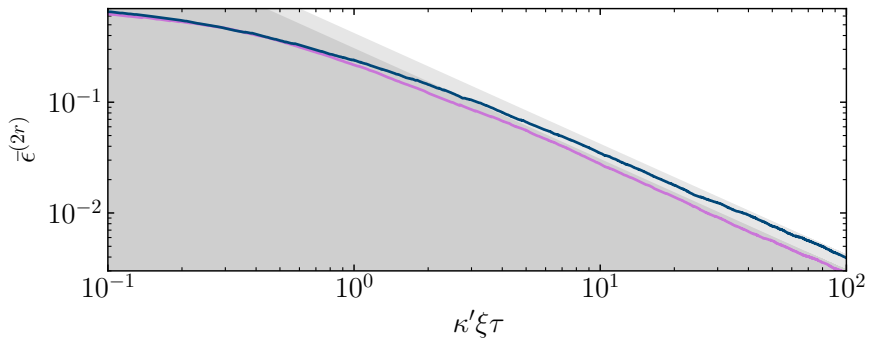


Figure 4.2: The average relative error $\bar{\epsilon}^{(2\mathrm{r})}$ of thermometry performed on a Bosonic bath as a function of measurement time τ. The error is calculated from 10^4 trajectories generated from randomly sampled temperatures in the range $[T_{\min}, T_{\max}]$. The error was calculated for both non-adaptive (dark blue) and adaptive (light purple) scenarios. The Cramér-Rao bound of the adaptive strategy is indicated by a dark grey shade and for the non-adaptive by the lighter grey shade. Both strategies reach their respective asymptotic scaling and the adaptive strategy is able to out-perform the non-adaptive one even for non-asymptotic times. Temperature and energies are expressed in units of ξ while τ has units of ξ^{-1}. The coupling κ' is dimensionless. We have set $T_{\min} = 0.1\xi$, $T_{\max} = 10\xi$, $\xi = 1$, and $\kappa' = 1$. In the adaptive strategy, the energy gap is changed in every simulated time step $d\tau = 10^{-3}$. Adapted from [B].

initialised with the globally optimal gap as a function of the measurement time for 10^4 randomly sampled trajectories. Here, we observe that asymptotically, the accuracy reaches the Bayesian Cramér-Rao bound indicated by the light grey shade. This means that this is an asymptotically optimal strategy.

Next, we assume that the energy gap of the probe can be changed during the course of the experiment. Again, we do not perform an optimisation for any given τ but one that is optimal for long times. To do this, we assume that at some time t we have observed the trajectory ν_t and the estimate of the temperature is $\vartheta(\nu_t)$. We then change the gap to be the one that will maximise the Fisher information for the estimated temperature,

$$\omega^*_{\text{ad.}}(\nu_t) \coloneqq \arg\max_{\omega} \tilde{T}^2(\nu_t)\mathcal{I}[\tau, \vartheta(\nu_t)]. \tag{4.27}$$

We can then use the expressions for long time Fisher information of the reservoirs, see Eq. (4.24) and Eq. (4.25), to find the optimal gap in each case which gives,

$$\omega^*_{\text{ad.}}(\nu_t) \approx 2.4750\, \tilde{T}(\nu_t), \tag{4.28}$$

for a Bosonic bath and,

$$\omega^*_{\text{ad.}}(\nu_t) \approx 2.6672\, \tilde{T}(\nu_t), \tag{4.29}$$

for a Fermionic bath.

This strategy will tune the gap to the fixed optimal gap for the actual underlying temperature for long enough times and the average error will converge to the Cramér-Rao bound with the asymptotic Fisher information calculated for the optimal gap at each theta,

$$\bar{\epsilon}^{(2\text{r})}(\vartheta_x)_{\text{ad.}} \geq \int \mathrm{d}\theta \frac{P(\theta)}{\theta^2 \mathcal{I}_{\text{ad.}}[\theta]}. \tag{4.30}$$

This strategy is not guaranteed to be optimal for transient times where optimal Bayesian strategies like the one considered in [66] may give a better performance. Unlike the strategy that we will consider, these strategies requires the optimisation to be calculated in each time step depending on the current posterior. It would thus require more computational power than the simple

Strategy	Optimal Gap	Optimal value of $T^2\mathcal{I}[\tau,T]$	$\int_{T_{\min}}^{T_{\max}} dT\,\rho(T)\,[T^2\mathcal{I}[\tau,T]]^{-1}$
Non adaptive with initially optimised gap	$\approx 0.4595\,\xi$	—	$0.4133\,(\kappa'\xi\,\tau)^{-1}$
Adaptive asymptotically thermal	$\approx 2.4750\,T$	$1.5430\,\kappa'\tau T$	$0.3015\,(\kappa'\xi\,\tau)^{-1}$

Table 4.1: The optimal adaptive and non-adaptive strategies and the scaling of these strategies for estimation of the temperature of a Bosonic bath. For the non adaptive strategy in the first row, the energy gap of the probe is chosen to be the one that minimises $\bar{\epsilon}^{(2\mathrm{r})}$ in Eq. (4.13) and then is left unchanged for the entirety of the measurement process. The adaptive strategy summarised in the second row starts with the same initial gap but after each time step it is tuned to the optimal value by setting $\vartheta = T$ in the second column. In the asymptotic limit we can compare the scaling of the error $\bar{\epsilon}^{(2\mathrm{r})}$ and we see that $\left[\bar{\epsilon}^{(2\mathrm{r})}(\mathrm{ad.}) - \bar{\epsilon}^{(2\mathrm{r})}(\mathrm{n.\ ad.})\right]/\bar{\epsilon}^{(2\mathrm{r})}(\mathrm{ad.}) \approx 40\%$. Temperature and energies are expressed in units of ξ while τ has units of ξ^{-1}. The coupling κ' is dimensionless. In the examples we have set $T_{\min} = 0.1\xi$, $T_{\max} = 10\xi$, $\xi = 1$, and $\kappa' = 1$.

strategy we adopt here where only the current estimate is required. In Fig 4.2, we demonstrate the results summarised in Table 4.1 for thermometry of a Bosonic bath averaged over 10^4 trajectories generated for randomly sampled temperatures in the range of the prior, $[0.1, 10]\xi$. The expected asymptotic scaling is reached by both strategies and the adaptive strategy can outperform the non-adaptive strategy even for short times.

4.6 Noisy measurements

Up to now we have only studied an idealised scenario where the measurements can be performed projectively. However, many experimental platforms like a quantum dot [143], on which a two level thermometer could actually be implemented will give noisy trajectories rather than the step-like trajectory considered up to now. The measurement device may also have a finite bandwidth which introduces a delay into the signal. This corresponds to a scenario where, as the interval between measurements tends to zero, so does the information gained per measurement. For an infinite bandwidth, The measured signal will consist of the actual telegraph like signal which, since we measure in the energy basis is $\langle\sigma_z\rangle$ plus Gaussian white noise. In a single time step an

Strategy	Optimal Gap	Optimal value of $T^2\mathcal{I}[\tau, T]$	$\int_{T_{\min}}^{T_{\max}} dT\,\rho(T)\,[T^2\mathcal{I}[\tau, T]]^{-1}$
Non adaptive with initially optimised gap	$\approx 1.5401\,\xi$	—	$132.79\,(\Gamma\,\tau)^{-1}$
Adaptive asymptotically thermal	$\approx 2.6672\,T$	$0.3795\,\Gamma\tau$	$2.6350\,(\Gamma\,\tau)^{-1}$

Table 4.2: The optimal adaptive and non-adaptive strategies and the scaling of these strategies for estimation of the temperature of a Fermionic bath with flat spectral density (i.e., $s = 0$). The strategies are the same as Table. 4.1. Here the adaptive strategy can improve the asymptotic precision more than an order of magnitude compared to the non-adaptive one. This effect is mainly due to the choice of a flat spectral density, that is more appropriate for fermionic baths. Again, all the parameters are expressed in terms of ξ, such that temperature T, frequency ω, and the coupling Γ have dimension ξ, while time τ has the dimension ξ^{-1}. For this table we have set $T_{\min} = 0.1\xi$, $T_{\max} = 10\xi$, and $\Gamma = \xi = 1$.

increment in the measured signal is defined by,

$$dz(t) = \langle \sigma_z \rangle\, dt + \frac{1}{\sqrt{\lambda}} dW. \tag{4.31}$$

Here, dW is known as a Wiener increment and it is scaled by a factor $1/\sqrt{\lambda}$ which determines how noisy the measurement is. That is, for $\lambda \to \infty$ the measurement is perfect but for $\lambda \to 0$ the measurement will be very noisy. This factor is called the measurement strength. The Wiener increment is a Gaussian random variable which has a mean value of 0 and a variance dt. Measurements that generate a signal like this are modelled by the Gaussian measurement operator [68, 139],

$$A(z) = \left(\frac{2\lambda\delta t}{\pi}\right)^{\frac{1}{4}} e^{\lambda\delta t(z-\sigma_z)^2}. \tag{4.32}$$

In the limit $\lambda\delta t \to 0$ we have a continuous measurement. The signal we observe at time t is thus z_t consisting of noisy fluctuations around the actual value 0 or 1 of the state of the system.

4.6.1 Noisy measurements with a finite bandwidth

Now, we want to include the bandwidth of the detector into this description. We assume noisy trajectory with a finitie bandwidth consists of outcomes D_t that are read out of the detector at time t. We model the finitie bandwidth as a low pass frequency filter. Thus, the relation between the unfiltered signal and the filtered signal D_j is defined as [144],

$$D_j = \gamma\delta t \sum_{k=0}^{j} e^{-\gamma(j-k)\delta t} z_k, \tag{4.33}$$

where, $1/\gamma$ is the delay introduced in the detector from the finite bandwidth, where all frequencies larger than γ are dampened. We will denote the trajectory that is observed as, $\nu_\tau = \{D_t | t \in [0, \tau]\}$.

We can model the measurement without filtering with the measurement superoperator,

$$\mathcal{M}(z)\rho := A(z)\rho A^\dagger(z) \tag{4.34}$$

with measurement operators $A(z)$ defined in Eq. (4.32). Now, we add the frequency filter to get the measurement operator for outcome D having observed D' in the previous time step. Due to this filter, the signal in step t_{j+1} depends on the previous signal D' and the current signal D through $z_{j+1} = (D - D'e^{-\gamma\delta t})/\gamma\delta t$. Substituting this in to the measurement superoperator Eq. (4.34) we define the measurement operator, $\gamma\delta t\mathsf{M}(D|D') = \mathcal{M}((D - D'e^{-\gamma\delta t})/\gamma\delta t)$, with,

$$\mathsf{M}(D|D') = \sqrt{\frac{2\lambda}{\pi\gamma^2 dt}} \begin{pmatrix} e^{-\frac{2\lambda}{\gamma^2 dt}(D - D'e^{-\gamma dt})^2} & 0 \\ 0 & e^{-\frac{2\lambda}{\gamma^2 dt}[D - (D'e^{-\gamma dt} + \gamma dt)]^2} \end{pmatrix}.$$

The system evolution can then be written down stroboscopically where the system interacts with the environment and is then measured,

$$\boldsymbol{P}_t(\nu_t) = \mathsf{M}(D_t|D_{t-dt})e^{\mathsf{W}dt}\boldsymbol{P}_{t-dt}(\nu_{t-dt}), \tag{4.35}$$

where $\boldsymbol{P}_\tau(\nu_\tau) = (p_0(\nu_\tau), p_1(\nu_\tau))^{\mathrm{T}}$ is a column vector with $p_j(\nu_\tau)$ the joint probability of

occupying state $j \in \{0, 1\}$ at time τ and observing ν_τ. W is the rate matrix (Eq. (4.1)) that describes the probes evolution under interaction with the bath. The likelihood to observe the noisy trajectory is,

$$\rho(\nu_\tau|T) = (1, 1) \cdot \boldsymbol{P}_\tau(\nu_\tau). \tag{4.36}$$

In the limit of $\lambda \gg \gamma \to \infty$ we recover the ideal measurements and obtain the probability to observe a trajectory given in Eq.(4.11). Thus, we can determine the posterior distribution using Bayes rule as before by updating the prior with this likelihood function in each step and normalising the distribution.

The simulation of an experiment is then done as follows,

- The trajectory is simulated according to the Poisson process in Eq. 4.1 for some true temperature. Filtered white noise is them added on top of this trajectory. This gives the measurement record $D(t)$.
- The measurement operator using $D(t)$ as well as the rate equation in Eq. (4.1) are use to calculate the likelihood in Eq. (4.36).
- The prior is updated using the likelihood and prior from the previous step using Bayes rule.

4.6.2 Kushner-Stratonovich equation

The evolution of the prior distribution under continuous measurements with a finite bandwidth and Gaussian noise can also be derived using Ito-Stochastic calculus [145, 146]. That is, we use Itos rule that $(dW)^2 = dt$. This allows us to derive the Kushner Stratonovich equation which is used in Bayesian filtering of noisy trajectories [88, 147].

In the non-noisy case, the probability to occupy the ground state is monitored which is equivalent to monitoring the difference in probability to occupy the ground and excited states, and we would get the signal, $r = \langle z \rangle = 2p_0(t) - 1 = \{-1, 1\}$. Recall that for the noisy signal we can define the Wiener increment,

$$dW = 2\sqrt{\lambda}(r - \langle z \rangle)dt. \tag{4.37}$$

When we introduce the finite bandwidth γ, the outcome $D(t)$ read out on the detector at time t is related to the infinite bandwidth signal by,

$$D(t) = \int_{-\infty}^{t} ds\, \gamma e^{-\gamma(t-s)t} r(s). \tag{4.38}$$

This is the continuous version of Eq.(4.33). The observed outcome from the experiment is the solution to the stochastic differential equation,

$$dD = \gamma(2p_0(t) - 1 - D)dt + \frac{\gamma}{2\sqrt{\lambda}} dW. \tag{4.39}$$

By expanding the likelihood $P(dD|T)$ (Eq. 4.35) to first order in dt and normalising the probability distribution we obtain the evolution of the ground state population [144],

$$\begin{aligned} dp_0(t) =& (\Gamma_{\text{out}} p_1(t) - \Gamma_{\text{in}} p_0(t) + 8\lambda p_0(t) p_1(t)(2p_0(t) - D - 1))dt \\ &+ \frac{8\lambda}{\gamma} p_1(t) dD. \end{aligned} \tag{4.40}$$

We now derive and expression for the evolution of the state of knowledge about T. The change in this state of knowledge after an infinitesimal time step of the experiment is

$$dP(T, t|\nu_\tau) = P(T, t + dt|\nu_\tau, dD) - P(T, t|\nu_\tau). \tag{4.41}$$

The evolution can then be broken up into an instantaneous part due to measurement (the first term below in square brackets) and the time evolution after measurement

$$\begin{aligned} dP(T, t|\nu_\tau) =& [P(T, t|\nu_\tau, dD) - P(T, t|\nu_\tau)] \\ &+ [P(T, t + dt|\nu_\tau, dD) - P(T, t|\nu_\tau, dD)]. \end{aligned} \tag{4.42}$$

However, the state of knowledge will only evolve independently of measurements if predictions are made. Excluding prediction gives

$$dP(T, t|\nu_\tau) = P(T, t|\nu_\tau, dD) - P(T, t|\nu_\tau). \tag{4.43}$$

The posterior distribution after measurement, $P(T,t|\nu_\tau,dD)$, can be computed using Bayes Rule:

$$P(T,t|\nu_\tau,dD)=\frac{P(dD|T)P(T,t|\nu_\tau)}{\int dTP(dD|T)P(T,t|\nu_\tau)}. \tag{4.44}$$

Then, to calculate the difference in Eq. 4.43, we first expand $P(T,t|\nu_\tau,dD)/P(T,t|\nu_\tau)$ to first order in dt and to second order in dD taking the expectation of the dD^2 terms since $\langle(dD)^2\rangle=\gamma^2dt/(4\lambda)$. That is,

$$\begin{aligned}\frac{P(T,t|\nu_\tau,dD)}{P(T,t|\nu_\tau)}=&\frac{\exp(\frac{-2\lambda}{\gamma^2dt}(dD-\gamma(2p_0(t)-D-1)dt)^2)}{\int dTP(dD|T)P(T,t|\nu_\tau)}\\ =&\frac{\exp(\frac{-2\lambda}{\gamma^2}(-2\gamma(2p_0-D-1)dD+\gamma^2(2p_0-D-1)^2dt))}{\int dT\exp(\frac{-2\lambda}{\gamma^2}(-2\gamma(2p_0-D-1)dD+\gamma^2(2p_0-D-1)^2dt))P(T,t|D)}\\ =&1+\frac{4\lambda}{\gamma}((2\overline{p_0}^2-D-1-(2p_0-D-1)(2\overline{p_0}-D-1))dt\\ &+\frac{4\lambda}{\gamma}(2p_0-D-1-(2\overline{p_0}-D-1))dD+o(dt),\end{aligned} \tag{4.45}$$

where, $\overline{p_0}=\int dTp_0P(T,t|\nu_\tau)$. Therefore, the evolution of the posterior is

$$dP(T,t|D)=\frac{8\lambda}{\gamma}(p_0-\overline{p_0})(dD-\gamma(2\overline{p_0}-D-1)dt). \tag{4.46}$$

which is known as the Kushner-Stratonovich equation [147, 148].

The simulation of an experiment is then done as follows,

- The trajectory is simulated according to the Poisson process in Eq. 4.1 for some true temperature. Filtered white noise is them added on top of this trajectory. This gives the measurement record $D(t)$.
- The measurement record $D(t)$ is used to solve for the evolution of $p_0(t,T)$ at each temperature in the range of the prior.
- The prior is updated using the measurement record and the prior averaged population $\bar{\rho}_0$ which was averaged over the previous increment of the prior.

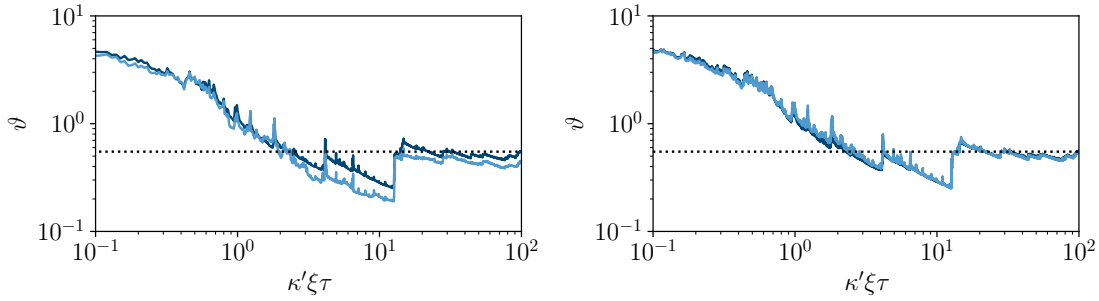


Figure 4.3: The estimated temperature for a single noisy trajectory with finite bandwith $\gamma = 10\xi$ is compared for two different time steps in the simulation. We set $dt = 10^{-2}$ in (a) and $dt = 10^{-3}$ in (b). The Kushner-Stratonovich method (light blue line) requires a smaller time step to converge than the direct calculation method (dark blue line). The actual temperature of the system is indicated by the dotted black line. The measurement strength is set to $\lambda/\gamma^2 = 0.05\xi$.

4.6.3 Comparison of simulations

We can now compare the two simulation methods to determine which is more efficient. The Kushner Stratonovich equation leads to an improved computational time because only one element of the state vector needs to be modelled to get to the Posterior distribution. However, the computation of the stochastic differential equation requires an approximation method. Here we simulate it using the second order Milstein method [145, 146]. In Fig. 4.3, we show that this additional approximation introduces numerical instability when the time step δt is not small enough. We illustrate this by comparing the estimated value of θ between simulations, for the same trajectory but setting $\delta t = 10^{-2}$ in Fig.4.3 (a) and $\delta t = 10^{-3}$ in Fig.4.3 (b). We see that the trajectory does not change between these two plots for the direct calculation method which indicates that the simulation of the equation has converged. On the other hand, the Kushner Stratonovich method only converges for the smaller time step. The decrease in the time step size by one order of magnitude increases the simulation time far more than having to calculate two values for the state in each step. Therefore, the direct method derived in Sec. 4.6.1 was used to simulate the data shown in the next two sections.

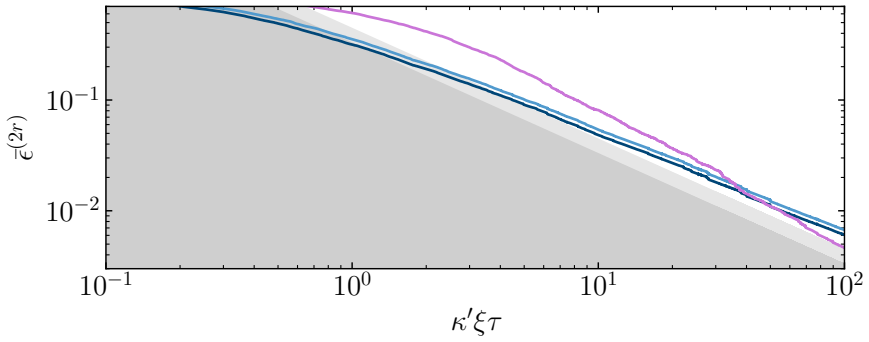


Figure 4.4: The average relative error $\bar{\epsilon}^{(2\mathrm{r})}$ of thermometry performed on a Bosonic bath as a function of measurement time τ for noisy measurments with a finite bandwith $\gamma = 10.0\xi$. The error is calculated from 10^3 trajectories generated from randomly sampled temperatures in the range $[T_{\min}, T_{\max}]$. The error was calculated for two non-adaptive scenarios with different measurement strength ($\lambda/\gamma^2 = 0.25\xi$ light blue and $\lambda/\gamma^2 = 0.50\xi$ dark blue) and one adaptive ($\lambda/\gamma^2 = 0.25\xi$ light purple) scenario. The Cramér-Rao bound of the adaptive strategy is indicated by a dark grey shade and for the non-adaptive by the lighter grey shade. Increasing the measurement strength leads to a better accuracy however, the noise prevents the Cramér-Rao bound from being reached. The adaptive strategy performs worse than the non-adaptive one at short times. Temperature and energies are expressed in units of ξ while τ has units of ξ^{-1}. The coupling κ' is dimensionless. We have set $T_{\min} = 0.1\xi$, $T_{\max} = 10\xi$, $\xi = 1$, and $\kappa' = 1$. In the adaptive strategy, the energy gap is changed in every simulated time step $dt = 10^{-3}$. Adapted from [B].

4.6.4 Performance of noisy measurements

In Fig. 4.4, contrary to what we saw for ideal measurements, we see that noisy measurements of the temperature of a Bosonic bath do not saturate the Cramér-Rao bound Eq. (4.15). The adaptive strategy, which is the same one as was used for the ideal measurement can reach the non-adaptive Cramér-Rao bound for long times. We cannot simulate for arbitrarily long times and thus we could not rule out the possibility that the adaptive noisy measurement may outperform the ideal non-adaptive measurement. In contrast to ideal measurements, the adaptive strategy actually performs worse than non-adaptive measurements for short times. The adaptive strategy is not designed to be optimal for the noisy scenario. It was designed to be optimal in the limit of long times for ideal measurements.

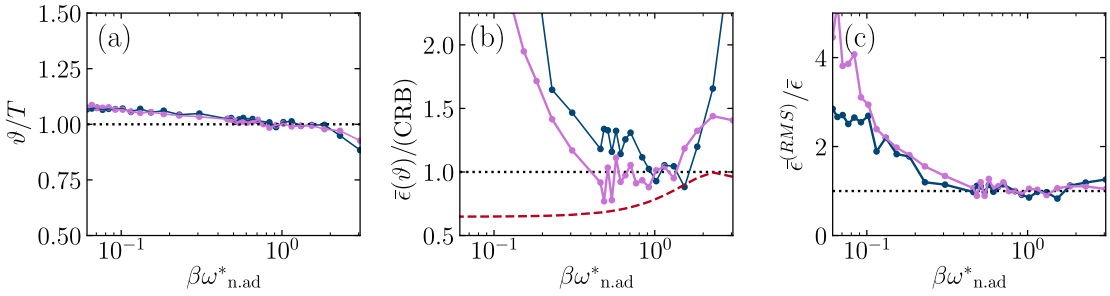


Figure 4.5: Comparison of the bias (a) relative error (b) and accuracy of the relative error relative to the true spread of the data $\bar{\epsilon}^{\mathrm{RMS}}$ in Eq. (2.29). Here $\mathrm{ad.(CRB)/n.ad.(CRB)}$ is denoted by the red dashed line. In (c) for noisy measurements with bandwidth $\gamma = 10.0\xi$ measurement strengths $\lambda/\gamma^2 = 0.25\xi$. The thermometry protocol was simulated for 35 true temperatures T in the range $[0.1, 10.0]\xi$ and with initial gap $\omega^*_{\mathrm{n.\,ad}}$. At each of the temperatures the results are averaged over a set of 200 trajectories and the points are joined with lines to guide the eye. The results are plotted at time $\kappa'\xi\tau = 100.0$. The shaded regions depict the standard deviation of the estimated temperature, true relative error and estimated error relative to the true error for each of the sets of trajectories. The blue points correspond to the non-adaptive strategy. The purple points, correspond to the adaptive strategy where new trajectories are generated because the gap of the two level system changes in each step. The strength of the measurement has the greatest effect in reducing the bias and the bias is also the main contribution to the relative error. Adapted from [B].

In Fig 4.5 (a) we see that for noisy measurements a significant estimation bias persists at low and high temperatures. The temperature is estimated for 35 different true temperatures in the range $[0.1, 10.0]\xi$. For each true temperature, the average estimate of 200 trajectories is calculated at the final time of each trajectory $\kappa'\xi\tau = 100$. At high temperatures, the temperature is over estimated and at low temperatures opposite behaviour is observed. If the measurement strength is increased, the bias is reduced for high temperatures but this is not observed at low temperatures. This is the reason behind the noisy measurements failing to saturate the Cramér-Rao bound. At low temperatures this bias arises from there not being sufficient time for transitions to be observed which results in the estimated temperature being below the actual temperature. We see that one of the ways that the adaptive strategy improves the estimation is by slightly reducing the bias. In particular, at low temperatures, being able to adjust the gap of the system allows more transitions to occur in the trajectory leading to a better estimate.

The average error $\bar{\epsilon}^{(2r)}$ of both the adaptive and non-adaptive strategy is plotted relative to the non-adaptive Cramér-Rao bound in Fig. 4.5 (b) with the ratio of the adaptive to non-adaptive Cramér-Rao bound shown for reference. Agreement with this line would be the best error an adaptive strategy could achieve. We see that for a small window of temperatures, the non-adaptive bound is achieved by even the non-adaptive strategy. This window is extended by the adaptive strategy and there is significant improvement for small temperatures where there is a large bias.

Finally, Fig. 4.5 (c) shows that the bias leads to the estimated error $\bar{\epsilon}^{(2\mathrm{r})}$ under estimating the true error $(\bar{\epsilon}^{(RMS)})^2$ when the total measurement time is not long enough. For asymptotically long times these errors must coincide due to the Bernstein von Mises theorem. From an experimental point of view, this suggests that in the range of temperatures where the bias is the main contribution to the error, the actual variability of the data is higher than what would be expected from the error estimate $\bar{\epsilon}^{(2\mathrm{r})}$.

4.7 Conclusions

In this chapter we have explored thermometry of a Bosonic and Fermionic reservoir using a two level probe under continuous measurement. The trajectory was derived exactly and we found

that it depends only on the observed number of jumps and the total time spent in the ground state. This result is important because it allowed the derivation of an analytical expression for the Fisher information. For long times, this gives a result for the Bayesian Cramér-Rao bound. This result was used to design an asymptoically optimal adaptive strategy. For ideal measurements where the state of the probe is monitored perfectly, the adaptive and non-adaptive strategies were found to reach these bounds in simulations and the adaptive strategy improved on the non-adaptive strategy even for short times.

In addition, we studied the more realistic situation where the measurements are weak which results in a noisy signal with a finite bandwidth. Analytical solutions cannot be found here, however, the results were simulated using both the Kushner stratonovich method and a direct calculation using the measurement operator to update the posterior distribution. We found that while the Kushner stratonovich method allows fewer matrix elements to be calculated, the calculation requires a smaller time step to converge. In simulations we found that when the adaptive strategy is applied to noisy measurements, the accuracy improves at long times but for shorter times the performance of the adaptive strategy is actually worse because the noisy measurements induce a bias in the estimates on which the adaptive strategy is based.

There are several future directions, first, this chapter only considered classical dynamics. If the probe were to be prepared in a state with coherences and coupled to a driving field that maintained these coherences throughout the measurement, the results of this chapter would need to be revisited to determine the optimal measurement strategy in this case. In particular, it is not clear from these results if quantum correlations will improve the measurement precision. Many body probes may also be considered. In this case one could study how correlations between these probes may be used to improve the accuracy of the measurements. The optimisation performed in the adaptive strategy is based on the Bayesian Cramér-Rao bound. This means that although it is optimal for the entire range of the prior, it is only optimal for *many* ideal measurements. Adaptive Bayesian strategies that optimise the Bayesian error like that studied in [66] may lead to a better strategy for few measurements and in noisy scenarios. However, this strategy would come at a large computational cost because of needing to simulate very many noisy trajectories. It would thus be an advantage to find a simpler, possibly heuristically motivated, adaptive strategy that performs well for few measurements but is perhaps sub-optimal.

Chapter 5

The thermodynamics of collectively coupled multi-level systems

In this chapter we will explore how collective coupling can enhance or degrade the performance of open systems compared to individually coupled counterparts. As a particular application, we will look at how much work can be extracted from such a system. A mathematical toolbox from representation theory is presented, which allows for the study of various properties of collectively coupled d-level systems. This work is based on publication [C].

5.1 Motivation

The examples considered in the first part of this dissertation are concerned with individual systems in contact with their environment. While this serves to illustrate the use of Bayesian techniques in open system metrology without additional complication, it is often the case that many-body systems can display interesting emergent phenomena that would not necessarily be expected from the simple addition of single-particle systems. For example, multiple spins may be used as a probe system with enhancements in temperature sensitivity near critical points [149, 150]. Alternatively, metrological enhancements can also be achieved through collective coupling. Collectively coupled probes reach a non-thermal steady state with the environment [56, 151]. This is useful for very low or high temperature thermometry where precision diverges with the num-

ber of probes when the probes are prepared in a Gibbs state. As is the case for a Gibbs state, the quantum Fisher information for temperature estimation is proportional to the heat capacity For collectively coupled spins as well [89, 151]. However, the heat capacity of the non-equilibrium state of the collectively coupled ensemble does not decrease as fast at high temperatures as it does for a Gibbs state. Similarly, for collectively coupled oscillator probes, the non-equilibrium steady state of the probes achieves a higher Fisher information than independent probes at low temperatures [56]. These systems are the focus of this chapter. The thermodynamic results from this were extended in [152] for n-level systems operating in an Otto cycle. This work suggests that the enhancements in the heat capacity are also present under certain conditions of initial state preparation for generic collectively coupled n-level systems. Additionally, these collective enhancements in the quantum Fisher information were also shown with a probe consisting of n oscillators for the low temperature regime in [56] and for collective spins strongly coupled to a bath in [149].

Generically, a collective system consists of many subsystems that are identical from the point of view of the environment but are not necessarily fundamentally indistinguishable. This allows for the system to display more exotic types of exchange symmetry than purely bosonic or fermionic exchange symmetry [152–154].

A description of the system that allows for efficient calculations while including interactions between particles, dynamical behaviour beyond steady state dynamics and independently operating baths, is still lacking. Such a description would allow for the detection of multiple temperatures or temperature differences as well as interesting non-equilibrium effects or the investigation of probe states that are only reachable when the particles are allowed to interact. The theoretical description of the collective dynamics for multiple levels beyond the usual collective spin description is the first obstacle and the one that we tackle here. Describing such a framework is the focus of the first parts of this chapter. The results from this chapter could be extended to systems with interactions in later works.

We first extend the description in [152] to describe a collectively coupled system simultaneously coupled to multiple heat baths prepared in any initial state. This description is then applied to a collectively coupled three level system which is inspired by the single-particle heat engine proposed by Scovil and Schultz-DuBois [9]. This single-particle engine is the paradigmatic example of a continuously operating quantum heat engine. As is the case in the metrological

setting, the advantages that may be obtained from many-body heat engines over their single-particle counterparts are not restricted to systems with collective coupling. Many-body systems like Bose-Eistein condensates or spin chains are a common extension of the basic single-particle quantum engine [155–157]. In these systems the strong non-linear interactions can be studied and utilised to affect the performance of the engine [5, 158, 159] particularly near critical points [23, 30, 159–161]. Unique properties of many-body quantum systems like many-body localisation can be used to enhance engine performance [162].

The enhancements seen for many-particle engines do not even require interactions between the individual particles making up the working medium. In our approach, the many-body effects come from the collective interaction of the particles in the working medium with the environment. The study of these types of effects is motivated in part because quantum thermal machines with this type of coupling are known to display enhanced performance similar to superradiance[1] [31, 164–168]. In these models each emission or absorption process is a superposition of single-particle events, which interfere constructively. This has also been experimentally demonstrated in a system of cold atoms [169]. Most recent studies of collective heat engines have focused only on Bosonic exchange symmetry [151, 170–175] as this gives a simpler dynamical description in terms of the well known collective spin operators. In this chapter, we will investigate how different types of exchange symmetry, or anti-symmetry affect the performance of an engine.

I will first briefly introduce quantum thermal machines in Sec 5.2 and introduce the mathematical tools necessary to describe permutationally symmetric systems in Sec. 5.3. The dynamics and steady states of collectively coupled three level systems are described in Sec. 5.4. Finally, these systems are put to use as a collective working medium and compared to individually coupled three level heat engines in Sec. 5.5.

While these results highlight interesting effects of collective coupling in their own right, the methods in this chapter also pave the way toward a description of metrology using collectively coupled systems that are not in the fully symmetric state. The methods presented here would be extendable to a description that includes interactions and the description of the system in-

[1] Superradiance is a phenomenon where the power of the light emitted by n dipoles that interact with a field collectively scales quadratically rather than linearly and is probably the most well known effect due to collective coupling in quantum systems [163].

cluding multiple baths could be useful in metrology settings where multiple parameters must be simultaneously estimated.

5.2 Introduction to quantum thermal machines

The familiar picture of a classical heat engine consists of macroscopic state parameters like the position of a piston and pressure change in a gas. In a quantum thermal machine, the working medium is usually on the scale of at most a few atoms. The quantum state parameters that can be changed in a thermodynamic process are quantities like the spacing of energy levels or the occupation of the energy levels of the system.

This allows for the definition of thermodynamic cycles in analogy to classical engines [176]. The most minimal model in these cases is a two level system. Consider for example the quantum Otto cycle [32,177]. Here the thermodynamic process is broken up into four strokes. In the first stroke, the analogue to the isochoric (constant volume) stroke in a classical engine, the energy level spacing is held fixed and the system is allowed to thermalise with a hot bath. This induces a larger population in the excited state. The next stroke is the working stroke in analogy to an adiabatic, isentropic "compression" stroke. Here, the excited state population is fixed and the difference in energy between the levels is decreased. This is followed by a stroke where the system, now with a smaller energy gap, is thermalised with a cold bath which lowers the population of the excited state. Finally, the cycle is completed by applying work to the system to bring the energy gap back to the original value. Engines based on an Otto cycle are convenient because they allow for the clear calculation of how much work is extracted from the engine compared to how much work is applied and how much heat is absorbed or emitted. As we will see in the model considered in this chapter, these quantities are not always easily defined in quantum systems. This problem shows up for example in a quantum Carnot cycle [10]. This cycle is defined by replacing the "isochoric" strokes with isothermal strokes so the system is held in thermal equilibrium with either the hot or cold reservoir during each stroke. In this case, the isothermal strokes now consist of exchanges of both work and heat with the environment and defining what part of the energy exchanged is heat and what is work requires a more thoughtful analysis.

While the breakdown of the operation of an engine into cycles may be convenient, pro-

cesses that appear in nature like photosynthesis [2] or technological applications like the operation of nanoscale thermoelectrics like photovoltaic cells may more strongly resemble continuous quantum engines [25, 178] where the operation is not broken down into strokes but rather the system is simultaneously coupled to all baths. In fact, the first theoretical proposal of a quantum thermal machine [9] was based on the three level maser. Here, a single three level system with two transitions coupled to a hot and a cold reservoir with the third allowing for work to be performed emits light when there is a population inversion in the working transition. A three level system is the simplest quantum system that can act as a continuous heat engine or refrigerator and it is the model we will build the collective engine upon. The continuous operation of a three level system as either a refrigerator or an engine are illustrated in Fig 5.1. Both models require a

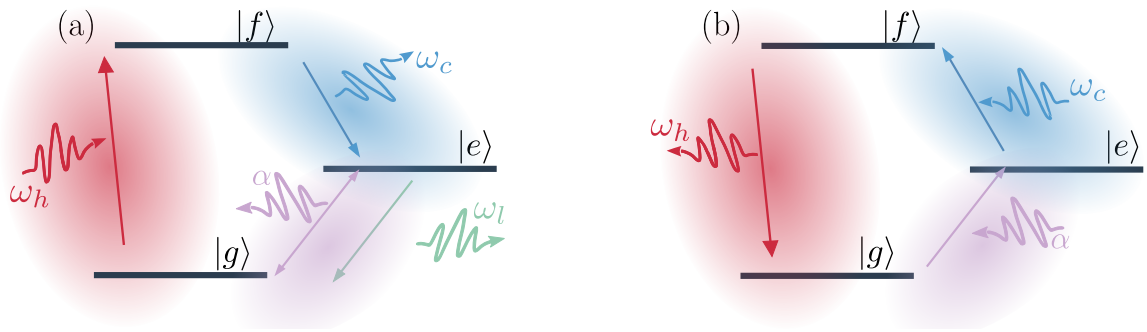


Figure 5.1: Diagram of the operation of a three level system as (a) a heat engine and (b) a refridgerator. In both cases, the system is coupled to a hot and cold reservoir with frequencies ω_c and ω through different transitions. The hot and cold reservoirs have inverse temperatures β_h and β_c respectively. Population inversion is achieved between $|g\rangle$ and $|e\rangle$ when $\omega_h\beta_h < \omega_c\beta_c$. In this case, an energy current can be emitted through the transition of frequency ω_l by coupling the system to a dissipative load or driving the transition with driving strength α (a). Alternatively, if the transition ω_l is coupled to a reservoir with inverse temperature $\beta_l < \beta_h$, or driven by an external field with driving strength α, the system may operate as a refrigerator, where heat is absorbed from the cold reservoir and dumped into the hot reservoir. Adapted from [C].

hot and cold bath and whether the system behaves as a refrigerator or engine depends on how the system exchanges energy with the environment. In both cases, the system couples to a hot bath at inverse temperature β_h with transition frequency ω_h and a cold bath at inverse temperature β_c with transition frequency ω_c. The remaining transition with frequency $\omega_l = \omega_h - \omega_c$

is either driven by a coherent driving field or allowed to couple dissipatively to another reservoir. In this chapter, we set $\hbar = 1$ and express all frequencies in units of the lasing frequency ω_l and inverse temperatures in units ω_l^{-1}.

The difference between a refrigerator or engine is determined by how this third level is coupled. In a refrigerator, heat is absorbed from the cold bath and dumped into the hot bath. This is facilitated by driving in the ω_l transition or coupling the ω_l transition to an environment with an inverse temperature $\beta_l < \beta_h$ [10,28,175,179]. For operation as an engine, which is the scenario we consider in this chapter, heat is absorbed from the hot bath and dumped into the cold bath. In the absence of any coupling to the ω_l transition, this leads to a population inversion. In this case, unitary operations can then be applied to extract work from the state [9, 180]. Another option is that the system could be allowed to output energy into a dissipative load [26, 164]. Alternatively, the transition may also be driven by an external field which gives rise to a quantum amplifier [181–185]. All three of these options will be explored in this chapter, however, we consider not only a single-particle working medium but rather an ensemble of non-interacting three level systems that couple collectively to the environment. In the absence of driving, the definition of what constitutes work or heat depends on the given context [26, 186]. This is explored Sec. 5.5.2 where we let the system emit power into a dissipative load modelled by a zero temperature bath in analogy to doing work against the force of friction. In this case it is not guaranteed that all energy emitted into this bath is work. We will qualitatively define a "work"-like energy current as one that is differentiated from heat in that it more strongly resembles laser-like light. Alternatively, when a classical driving field is included so that the system Hamiltonian is time dependent, the emitted energy can be associated with work unambiguously [11,187,188]. This is studied in Sec. 5.5.3.

5.3 Mathematical preliminaries

In this section only the mathematical details which are immediately necessary for understanding the thermodynamics of collectively coupled ensembles are discussed. Further background to representation theory is covered in more detail by textbooks like [189, 190].

The dynamics of a generic d-level system, which will also be called a particle in this chapter,

are described by unitary transformations, which are the $d \times d$ unitary matrices with determinant 1. These matrices make up the elements of the special unitary group $\mathrm{SU}(d)$. When there are n particles the permutations of these particles are described by a second group, the symmetric group S_n. The elements of this group are ordered lists of objects and operations in this group can be generated by the transposition operator $P_{i,j}$ which exchanges object i and j. The permutation matrices form a representation of this group. If the dynamics remain the same under all relabellings of particles, they are called permutation invariant dynamics. The Hilbert space is given by $\mathcal{H}^n = (\mathbb{C}^d)^{\otimes n}$. This physically corresponds to the situation described in the introduction to this chapter where an external control field or environment is not able to distinguish the particles.

5.3.1 Schur basis

When dealing with groups, it is often helpful to map the group elements onto linear operators. The group multiplication operation is then represented by matrix multiplication. In other words, there is a linear map D which acts on the elements of the group G such that for every $g, g' \in G$, $D(g * g') = D(g)D(g')$ which is called a representation of a group. If there is a non-trivial subspace of the representation such that the action of $D(g)$ on any element of the subspace stays within the subspace, then, the representation is reducible. If there is no such subspace, the representation is called irreducible. The irreducible representations of a group are called irreps for short. For the special unitary group, the Unitary matrices $U(u)$ form a representation of $\mathrm{SU}(d)$ with $u \in \mathrm{SU}(d)$. Using this, the product representation of the Hilbert space of n, d-level systems is, $U(u) = u^{\otimes n}$. This describes non-interacting unitary rotations of the n particles. The product representation is in general a reducible representation but there is a block-diagonal basis of the Hilbert space where each element of the representation,

$$U(u) = \bigoplus_{\lambda} \bigoplus_{m_\lambda} U^{\lambda}(u), \tag{5.1}$$

is a direct sum of irreps that are labelled by an index λ and may occur multiple times within the block diagonal structure, this is known as the multiplicity m_λ of the irrep [189–191].

Now, since permutations of the particles are described by the symmetric group,this basis

could also be written as a direct sum of the irreps of S_n. For general groups, this would not add any further structure to the Hilbert space but for the permutation group and special unitary group, further structure is given by Schur-Weyl duality [192]. The two group representations commute, $[U(u), P(\sigma)] = 0 \ \forall u, \sigma$ which means that the irreps of the symmetric group act on the multiplicity labels of the special unitary group and vice versa. There is a basis that simultaneously decomposes the space into the irreps of both groups such that the representations can be expressed as,

$$\begin{aligned} U(u) &= \bigoplus_{\lambda} \mathbb{1}_{\mathcal{K}^\lambda} \otimes U^\lambda(u) \\ P(\sigma) &= \bigoplus_{\lambda} P^\lambda(\sigma) \otimes \mathbb{1}_{\mathcal{H}^\lambda}. \end{aligned} \tag{5.2}$$

and applying this to the Hilbert space we get,

$$\mathcal{H}^n = \bigoplus_{\lambda} \mathcal{K}^\lambda \otimes \mathcal{H}^\lambda, \tag{5.3}$$

where, the label of the irreps $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_d]$, with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$, is associated with a partition of n into d non-negative integers such that $n = \lambda_1 + \lambda_2 + \cdots + \lambda_d$. $\mathcal{H}^\lambda$ and $\mathcal{K}^\lambda$ are the unitary and permutation subspaces which carry irreps of the special unitary group and symmetric group respectively. This basis, known as the Schur basis, is labelled $|\lambda, s_\lambda, o_\lambda\rangle$. The consequence of representing states of the Hilbert space in this basis is that operators acting on irreps of S_n change only s_λ and operators acting on the irreps of $\mathrm{SU}(d)$ change only o_λ.

We will only consider permutation invariant systems which means that any relevant observable of the system A will remain unchanged under any relabelling of the particles. In the Schur basis this observable is $A = \bigoplus_\lambda \mathbb{1}_{\mathcal{K}^\lambda} \otimes A^\lambda$. As long as only permutation invariant observables are considered, the permutation factor $\mathcal{K}^\lambda$ does not add anything to the dynamical description of the system and we can trace out the degrees of freedom s_λ from the Schur basis when calculating dynamics. We thus obtain a reduced basis $|\lambda, o_\lambda\rangle$ over the accessible Hilbert space,

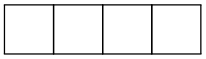
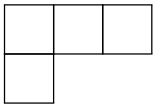
$$\mathcal{H}^\lambda_{\text{acc}} = \bigoplus_{\lambda} \mathcal{H}^\lambda. \tag{5.4}$$

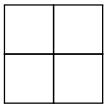
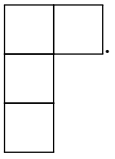
The construction of the reduced Schur basis and what it tells us about the permutation symmetry of each $\mathrm{SU}(d)$ irrep of the system is best illustrated by an example.

Example: Schur basis of four three level systems

Here, we will construct an operator that will allow us to transform an operator expressed in the standard product basis $|x_1, x_2, ...x_n\rangle$ of the un-factorised Hilbert space $\mathcal{H}^n = (\mathbb{C}^d)^{\otimes n}$, into the block diagonal basis $|\lambda, s_\lambda, o_\lambda\rangle$. First, as we noted above, Schur-Weyl duality [192] allows us to associate each irrep label λ with a partition of n into d non-negative, non-increasing, integers. This can be conveniently represented by means of a Young diagram. This is a diagram consisting n boxes arranged into d rows according to the partition. These diagrams are a useful tool for characterising the irreps of a particular system and constructing the Schur basis.

Consider a collection of four, three level systems. There are four partitions of four particles into 3 levels, $\lambda_a = (4, 0, 0)$, $\lambda_b = (3, 1, 0)$, $\lambda_c = (2, 2, 0)$, $\lambda_d = (2, 1, 1)$. These correspond to the four allowed permutation symmetry classes for $n = 4$. The Young diagrams for each of these partitions are:

$\lambda_a =$  $\lambda_b =$ 

$\lambda_c =$  $\lambda_d =$  . (5.5)

The number of ways each of these diagrams can be filled with the integers $\{1, ..., n\}$ without repetition and with numbers strictly increasing along both rows and columns gives the multiplicity m_λ of each $\mathrm{SU}(3)$ irrep, or equivalently, the dimension of the corresponding S_4 irrep [189,191,193]. The diagram filled in this way is called a standard Young tableau. For example λ_b can be filled in three ways ,

1	2	3
4		

1	2	4
3		

1	3	4
2		

. (5.6)

and the block describing this irrep will appear three times in the block diagonal Hilbert space of four three level systems each corresponding to a different filling of the Young diagram. We will get the following multiplicities for the other four irreps, $\{m_{\lambda_a} = 1, m_{\lambda_c} = 2, m_{\lambda_d} = 3\}$.

The dimension d_λ of each $\mathrm{SU}(d)$ irrep can be determined by a different filling prescription. The boxes are now filled with the numbers $\{0, ..., d-1\}$. In this case the integers must increase along columns but repetitions are allowed along rows as long as the numbers do not decrease from left to right. This filling prescription is consistent with the fact that there are no more than d rows in any given diagram.

The diagrams can also be used to construct the Schur basis:

- First, since we are interested in the reduced basis we fix the standard Young tableau to be in numerical order and construct permutation operators that permute numbers along rows and columns, e.g., set $T = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 \\ \cline{1-1}\end{array}$. This is used to construct the permutation operators, $P(r)$ and $P(c)$ that permute the numbers along rows and columns. In this example, $P(c) = \{P_{(1234)} = \mathbb{1}, P_{(2134)}\}$ and $P(r) = \{\mathbb{1}, P_{(3214)}, P_{(2134)}, P_{(2314)}, P_{(3124)}, P_{(1324)}\}$. We use the permutation notation so that for example, $P_{(2134)}$ means the first two elements in the set are switched and the last two are held fixed.

- These operators are used to construct the Young symmetriser [192] which is the projector which acts on elements of the computational basis $|x_1, x_2, x_3, x_4\rangle$ to bring them into the block diagonal structure, $|\lambda, s_\lambda, o_\lambda\rangle$. This projector[2] is

$$\mathcal{P}_{\lambda,T} \propto \left(\sum_{c \in \mathrm{Col}(T)} \mathrm{sgn}(c) P(c)\right)\left(\sum_{r \in \mathrm{Row}(T)} P(r)\right), \tag{5.7}$$

  where $\mathrm{Col}(T)$ are the permutations of the integers within each column of the diagram and $\mathrm{Row}(T)$ are the permutations of the integers in each row. If a diagram consists only of a single row, this will lead to a projection onto a fully symmetric subspace. Alternatively, if there is only a single column in the diagram the projection will be only a fully anti-symmetric subspace. In the λ_b example we will end up with mixed symmetry and the

[2]The sign of a permutation is determined by whether the minimum number of transpositions it takes to perform a permutation is even or odd.

projector, $\mathcal{P}_{\lambda,T} \propto (\mathbb{1} - P_{(2134)})(\mathbb{1} + P_{(3214)} + P_{(2134)} + P_{(2314)} + P_{(3124)} + P_{(1324)})$.

- Next, the restrictions of the permutation operators onto the relevant parts of the computational basis are constructed. To do this first compute the semi-standard tableaux. Informally each box labels a particle but each particle can be in a choice of three levels. The tableau is filled so that numbers do not decrease along rows but strictly increase along columns. In the example of λ_b, there are 15 tableaux,

$$\begin{array}{|c|c|c|}\hline 0&0&0\\\hline 1\\\cline{1-1}\end{array}\quad\begin{array}{|c|c|c|}\hline 0&0&1\\\hline 1\\\cline{1-1}\end{array}\quad\begin{array}{|c|c|c|}\hline 0&0&2\\\hline 1\\\cline{1-1}\end{array}$$
$$\begin{array}{|c|c|c|}\hline 0&0&0\\\hline 2\\\cline{1-1}\end{array}\quad\begin{array}{|c|c|c|}\hline 0&0&1\\\hline 2\\\cline{1-1}\end{array}\quad\begin{array}{|c|c|c|}\hline 0&0&2\\\hline 2\\\cline{1-1}\end{array}$$
$$\begin{array}{|c|c|c|}\hline 0&1&1\\\hline 1\\\cline{1-1}\end{array}\quad\begin{array}{|c|c|c|}\hline 0&1&2\\\hline 1\\\cline{1-1}\end{array}\quad\begin{array}{|c|c|c|}\hline 0&1&1\\\hline 2\\\cline{1-1}\end{array}$$
$$\begin{array}{|c|c|c|}\hline 0&1&2\\\hline 2\\\cline{1-1}\end{array}\quad\begin{array}{|c|c|c|}\hline 0&2&2\\\hline 1\\\cline{1-1}\end{array}\quad\begin{array}{|c|c|c|}\hline 0&2&2\\\hline 2\\\cline{1-1}\end{array}$$
$$\begin{array}{|c|c|c|}\hline 1&1&1\\\hline 2\\\cline{1-1}\end{array}\quad\begin{array}{|c|c|c|}\hline 1&1&2\\\hline 2\\\cline{1-1}\end{array}\quad\begin{array}{|c|c|c|}\hline 1&2&2\\\hline 2\\\cline{1-1}\end{array}\,. \tag{5.8}$$

The type vector $\mathbf{t}(\mathbf{x})$, which counts the number of particles in the ground state and first and second excited states, is then constructed. For example the tableau $\begin{array}{|c|c|c|}\hline 0&0&0\\\hline 1\\\cline{1-1}\end{array}$ has type vector $\mathbf{t} = (3,1,0)$ which has the span $\{|0001\rangle, |0010\rangle, |0100\rangle, |1000\rangle\}$ in the computational basis. The permutation operators $P(r)$ and $P(c)$ are then computed from the span of the type vector. e.g. $P_{(2134)} = |0001\rangle\langle 0001| + |0010\rangle\langle 0010| + |0100\rangle\langle 1000| + |1000\rangle\langle 0100|$ and then the Young symmetriser can finally be calculated. The type vector depends only on how many times each number appears in the semi-standard tableaux. Therefore, semi-standard tableaux with the same numbers will lead to the same type vectors and so are spanned by the same vectors from the computational basis. However, they lead to different Young symmetrisers since these act on the specific rows and columns.

- Perform singular value decomposition on the Young symmetriser. The first k columns of the left singular vector with non-zero singular values form the basis elements $|\lambda, o_\lambda\rangle$.

- Operators that were originally expressed in the Hilbert space $\mathcal{H}^n$ can now be projected onto this reduced basis by, $O^\lambda_{o,o'} = \langle\lambda, o_\lambda| O_n |\lambda, o'_\lambda\rangle$.

The construction of the Schur basis shows how each irrep corresponds to a different type of exchange symmetry of the particles. For example, λ_a, which consists only of a row will lead to a Young symmetriser that projects onto the subspace of the Hilbert space that describes all fully symmetric states. Additionally, the subspace given by λ_d is equivalent to the subspace of a single-particle showing that all columns with d boxes can be dropped from the description. The irreps of $\mathrm{SU}(d)$ can thus be labelled using $d-1$ labels.

5.3.2 Dynamical operators with a focus on $\mathrm{SU}(3)$

Thus far, the discussion has been applicable to any permutationally invariant d-level system but now we shall focus on a description of three level systems. To do this I will first briefly introduce some theory of Lie groups. This will allow us to define the dynamical operators of the system. These operators are already well known for two level systems as they are just the collective spin operators [189]. Their use in particle physics for three level systems is also well established [189, 194–196]. However, for a broader audience of physicists they may be unfamiliar. The dynamical operators of any d-level system are given by the generators of $\mathrm{SU}(d)$. These are the set of d^2-1 elements from which all other elements of $\mathrm{SU}(d)$ can be obtained by group multiplication. Using these operators along with their commutation relations, which define the $\mathrm{SU}(3)$ algebra, the entire group structure is defined. One way to simplify the description of the dynamics is to represent the dynamical operators in a basis where the maximum number of them can be simultaneously diagonalised. In terms of the generators, this corresponds to the largest set of commuting, Hermitian generators and gives the Cartan sub-algebra [152, 190]. These generators are represented by $d-1$ diagonal matrices. The remaining d^2-d generators are represented by non-diagonal matrices and come in pairs.

For $\mathrm{SU}(3)$ there are 8 generators in total. Using the Cartan basis, the diagonal generators, similar to the J_z operator in $\mathrm{SU}(2)$, are labelled W_z and Y [189] and for a single-particle they

are,

$$W_z^{(1)} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y^{(1)} = \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{5.9}$$

The ordering in the single-particle, computational basis is $\{|0\rangle, |1\rangle, |2\rangle\}$ so that they coincide with the Gell-Mann matrices from particle physics. The remaining six non-diagonal generators act as the $\mathrm{SU}(3)$ equivalent of the more familiar $\mathrm{SU}(2)$ raising and lowering operators. They come in pairs and are denoted $W_+ = W_-^\dagger$, $U_+ = U_-^\dagger$, and $V_+ = V_-^\dagger$. The single-particle versions are

$$W_+^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; U_+^{(1)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \; V_+^{(1)} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.10}$$

For the n-particle product representation, that is the reducible representation where the Hilbert space is not decomposed into the Schur basis, the collective generators are sums of single-particle terms: $A = \sum_{i=1}^n A_i^{(1)}$, where $A_i^{(1)}$ is any one of the above generators acting on particle i. Within a generic irrep of $\mathrm{SU}(3)$, the generators are defined by their commutators which are listed in App. D. Note also that a sub-irrep of $\mathrm{SU}(2)$ arises for the operators $W_z,\ W_\pm$. A similar sub irrep can also be defined for $U_z,\ U_\pm$ and $V_z,\ V_\pm$ with $U_z = 3/2Y + W_z$ and $V_z = 3/2Y - W_z$.

Convenient labelling of the reduced Schur basis for three level systems

The block diagonal structure obtained by decomposing the space into the Schur basis is useful for speeding up the calculation of the dynamics of larger systems. However the decomposition into this basis alone does not allow the calculation of quantities analytically. Notice that with the definitions of the operators given we cannot write down the action of the operators on a particular reduced Schur basis element. We merely know how the operators move between basis states but not what the coefficients are. Therefore, from the description of the system by the Schur basis alone we would be restricted to numerical calculation.

Fortunately it is possible to give a labelling of the reduced Schur basis which transforms conveniently under the action of the generators [194]. This basis labelling is in analogy to the $|J, m\rangle$

labelling that is familiar for collective two level systems. This basis $|(p,q),W,w,y\rangle$ is characterised by

(i) The irrep label $\lambda=(p,q)$. These numbers can be determined from the Young diagram of a particular irrep with p giving the number of single box columns and q giving the number of two box columns.

(ii) The number W which labels the eigenstates of the operator $W^2=\frac{1}{2}(W_+W_-+W_-W_+)+W_z^2$. This is the total spin-length quantum number defining the sub-irreps of $\mathrm{SU}(2)$ generated by $W_z,\ W_\pm$.

(iii) The numbers w and y label the eigenstates of W_z and Y respectively. generically, they are called the weights which are the eigenvalues of the Cartan generators of the group.

The coefficients that are obtained from the action of the generators on the states labelled in this way [194, 196] and the coefficients are summarised in App. D. Finally, note that the numbers p,q allow us to explicitly write down the dimension and multiplicity of the irrep for n particles [152, 189],

$$\begin{aligned} d_\lambda =& \frac{1}{2}(p+1)(q+p+2)(q+1) \\ m_\lambda =& \frac{2d_\lambda n!}{\left(\frac{n+q-p+3}{3}\right)!\left(\frac{n+q+2p+6}{3}\right)!\left(\frac{n-2q-p}{3}\right)!}. \end{aligned} \tag{5.11}$$

Weight diagrams

The diagonal elements of the basis, which are known as the weights of the irrep are illustrated graphically in a weight diagram. In Fig. 5.2 the weight diagram of the irrep corresponding to $(p,q)=(2,1)$ is plotted with the values of w on the x-axis and the value of y on the y-axis. Note that the number of points in a diagram is equivalent to the dimension d_λ of the irrep. The weight diagrams of $\mathrm{SU}(3)$ are either hexagonal or triangular with a $2\pi/3$ rotational symmetry (up to an axis rescaling). There are $(p+1)$ points on the top edge and $(q+1)$ points on the bottom edge of a given diagram. The right-most point (with largest w) is known as the highest-weight vector, and has coordinates $([p+q]/2,[p-q]/3)$. The right-most points with maximum

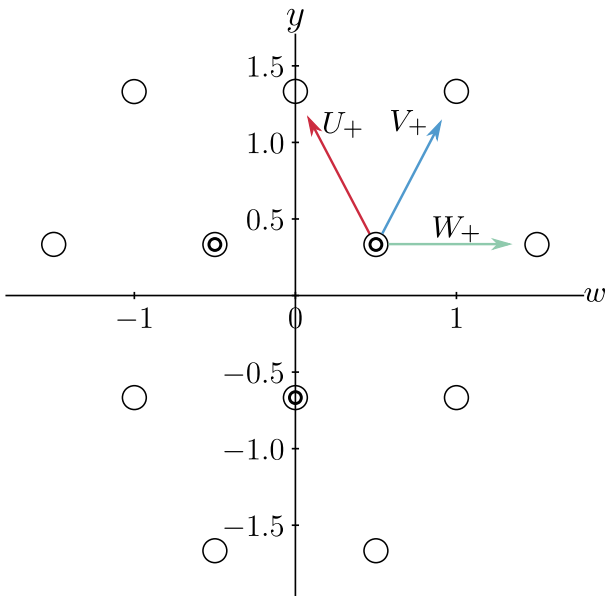


Figure 5.2: A weight diagram of one of the irreducible representations found for four three level systems, labelled $\lambda = (p, q) = (2, 1)$. The points in the diagram correspond to the reduced basis states $|(p, q), W_i, w_i, y_i\rangle$. The values of w_i are on the x-axis and y_i on the y-axis. The central states are degenerate, which is shown with concentric circles. The degenerate states have different values of W_i. The action of the $U_\pm$, $V_\pm$ and $W_\pm$ operators on a state is indicated by arrows in the diagram in the upper right hand corner. Adapted from [C].

and minimum y are then $(p/2, [p + 2q]/3)$ and $(q/2, -[2p + q]/3)$. Note that although under the action of $W_z, W_\pm$ the numbers W, w change in the same way as J and m do for a spin, there is no single operator that raises and lowers the value of y. A single action of $U_\pm$ or $V_\pm$ will change the value of w as well as y and therefore y does not change in the same way one might expect from the intuition obtained from $\mathrm{SU}(2)$. The central points in any hexagonal diagram have the same single particle expectation values with a non-interacting Hamiltonian and therefore, they have the same w and y values but differ in their W quantum numbers. The degeneracy is illustrated by concentric circles in Fig 5.2. The amount of degeneracy of a given point can be determined immediately from the shape of the diagram. The points on the perimeter of any diagram are always non-degenerate, as are all the points of a triangular diagram. The level of degeneracy increases by one going from the perimeter to the centre for each hexagonal layer in the weight diagram. Once a triangular layer is reached no additional degeneracy is added to the

points as the centre is approached [189].

5.4 Open system dynamics and the steady state

With the tools from representation theory outlined in the previous section, we are now equipped to describe the interaction of an ensemble of three level systems collectively coupled with an environment. In particular, the focus will be the thermodynamic properties of these systems when they are configured to perform as a heat engine. First the general form of the open-system model is introduced and the structure of the steady state is discussed.

5.4.1 General open system dynamics for work extraction from three level systems

The minimal model of a single-particle three level engine, on which the collective engine we study later is based, has transitions with frequencies ω_h and ω_c which are resonantly coupled to two single mode bosonic reservoirs at a high and a low temperature. The third transition $\omega_l = \omega_h - \omega_c$ is the transition from which work is extracted.

The non-interacting Hamiltonian for a single-particle engine is

$$h = \omega_l \left|e\right\rangle \left\langle e\right| + \omega_h \left|f\right\rangle \left\langle f\right|, \tag{5.12}$$

where we set $\hbar = 1$ and the energy eigenstates $\{\left|g\right\rangle, \left|e\right\rangle, \left|f\right\rangle\}$ map onto the computational basis states $\{\left|1\right\rangle, \left|0\right\rangle, \left|2\right\rangle\}$ described in sec 5.3.2. This Hamiltonian is generalized to collective systems by using the diagonal $\mathrm{SU}(3)$ generators introduced in sec 5.3.2 This can be seen in a straightforward way since in the product representation the generators are just a sum of single-particle terms. The Hamiltonian expressed as

$$H_S = \omega_l W_z - \left(\frac{\omega_c + \omega_h}{2}\right) Y. \tag{5.13}$$

is equivalent to the one given for single-particles, up to a constant shift which we can drop without changing the dynamics. Since the Hamiltonian is made up of only the diagonal gen-

erators, it is diagonal in the Schur basis with eigenvalues $E^{\lambda}_{wy} = \omega_l w - (\omega_c + \omega_h)y/2$.

The interaction of this ensemble with the thermal Bosonic reservoirs is modelled using a standard Lindblad master equation derived in App. A. For a single-particle this is

$$\begin{aligned}\dot{\rho}_{\text{sing}} =& g_v(\bar{n}_c\mathcal{D}[|e\rangle\langle f|]\rho + (\bar{n}_c+1)\mathcal{D}[|f\rangle\langle e|]\rho) \\ &+ g_u(\bar{n}_h\mathcal{D}[|g\rangle\langle f|]\rho + (\bar{n}_h+1)\mathcal{D}[|f\rangle\langle g|]\rho) \\ &+ g_w(\bar{n}_l\mathcal{D}[|g\rangle\langle e|]\rho + (\bar{n}_l+1)\mathcal{D}[|e\rangle\langle g|]\rho),\end{aligned} \tag{5.14}$$

where $\bar{n}_i = 1/(\exp(\omega_i\beta_i)-1)$, is the bosonic average occupation of the mode with frequency ω_i of the reservoir with inverse temperature $\beta_i = 1/k_B T_i$. Now, the jump operators are generalised to the collective counterparts. For example in the single-particle picture, the cold bath exchanges energy with the transition of frequency ω_c which is between the states $|e\rangle$ and $|f\rangle$. The collective counterpart to this is the collective operator $V_{\pm}$ which describes all the superpositions of the single-particle transitions between $|e\rangle$ and $|f\rangle$ in the ensemble. Transitions corresponding to energy exchange with the hot bath are described by $U_{\pm}$, leading to the master equation,

$$\dot{\rho} = g_u\mathcal{L}_U(\bar{n}_h)\rho + g_v\mathcal{L}_V(\bar{n}_c)\rho + g_w\mathcal{L}_W(\bar{n}_l)\rho. \tag{5.15}$$

where,

$$\mathcal{L}_O(\bar{n}_i)\rho = \bar{n}_i\mathcal{D}[O_-]\rho + (\bar{n}_i+1)\mathcal{D}[O_+]\rho. \tag{5.16}$$

Here, $\mathcal{D}[O]\rho = O\rho O^\dagger - \frac{1}{2}\{O^\dagger O, \rho\}$.

The dynamics are permutation-invariant since all of the operators are permutation invariant. Eq. (5.4) tells us that the accessible Hilbert space is, therefore, a direct sum of the irreps of $\mathrm{SU}(3)$ and all the dynamical operators can be expressed in this way as well.

The accessible state space will be $\rho_{\text{acc}} = \bigoplus_\lambda p^\lambda \rho^\lambda$, where each ρ_λ is normalised in the restricted Hilbert space $\mathcal{H}^\lambda$ and p^λ is a probability distribution over the index λ that is determined by the state preparation. Now, consider the situation where the system is prepared in some generic state and then expressed in the Schur basis. This state may have coherences between different blocks in the block diagonal structure but these coherences cannot be detected by any observable that also obeys permutation invariance [152] and therefore we can use the reduced Schur

basis that we have already constructed.

Since the irreps of the state cannot be mixed by the dynamics, we can calculate the dynamics for each block of the Hilbert space individually. The master equation is expressed for each block as

$$\dot{\rho}^{\lambda} = g_u \mathcal{L}_U^{\lambda}(\bar{n}_h)\rho^{\lambda} + g_v \mathcal{L}_V^{\lambda}(\bar{n}_c)\rho^{\lambda} + g_w \mathcal{L}_W^{\lambda}(\bar{n}_l)\rho, \tag{5.17}$$

where L_O^{λ} is the Lindblad term with the jump operator O replaced by O^{λ} its projection onto the λ component of the reduced Schur basis.

5.4.2 Structure of the steady state

The results that are explored in later sections, all come from the steady state behaviour of a system described by a master equation of the form Eq.(5.15). Therefore, in this section it is proven that the steady state, ρ_{∞}^{λ}, is block diagonal in the Schur basis and has no coherences in w or y. That is,

$$\langle \lambda, W, w, y | \rho_{\infty}^{\lambda} | \lambda', W', w', y' \rangle = 0 \text{ if } w \neq w', \ \lambda \neq \lambda' \text{ or } y \neq y'. \tag{5.18}$$

This result is proved in two parts. Part 1 is the statement that the steady state is block diagonal in the Schur basis. This was proved for generic d-level permutationally invariant collections of particles in theorem 1 of [152]. In Part 2, I will show that for the specific systems we consider here, where each bath couples to a different level, there are no coherences between different weight spaces of the diagonal generators. This is the proof shown in [C].

Part 1: The steady state is block diagonal in the Schur basis

The state is block diagonal with no coherences in λ,

$$\langle \lambda, W, w, y | \rho_{\infty}^{\lambda} | \lambda', W', w', y' \rangle = 0 \text{ if } \lambda \neq \lambda' \tag{5.19}$$

The steady state is also unique in each irrep if the Lindbladian has contributions from at least two of $\{U_{\pm}, V_{\pm}, W_{\pm}\}$.

Part 2: The steady state has no coherences in w or y

Under the same assumptions on the Lindbladian as part 1, define the unitary super-operator

$$\begin{aligned}\mathcal{U}_{\theta,\phi}(\cdot) &= U^{\dagger}_{\theta,\phi}(\cdot)U_{\theta,\phi} \\ &= e^{-i(\theta W_z+\phi Y)}(\cdot)e^{i(\theta W_z+\phi Y)}.\end{aligned} \tag{5.20}$$

Then [3], $[\mathcal{L},\mathcal{U}_{\theta,\phi}]=0$. Addtionally, let the state ρ be a steady state of the master equation and define the twirled state [197, 198]

$$\sigma = \int_0^{2\pi}\frac{d\theta}{2\pi}\int_0^{2\pi}\frac{d\phi}{2\pi}\mathcal{U}_{\theta,\phi}(\rho) \tag{5.21}$$

which is an average of the state ρ over the subgroup with elements $U_{\theta,\phi}$. Then,

$$\begin{aligned}\mathcal{L}(\sigma) &= \int_0^{2\pi}\frac{d\theta}{2\pi}\int_0^{2\pi}\frac{d\phi}{2\pi}\mathcal{L}(\mathcal{U}_{\theta,\phi}(\rho)) \\ &= \int_0^{2\pi}\frac{d\theta}{2\pi}\int_0^{2\pi}\frac{d\phi}{2\pi}\mathcal{U}_{\theta,\phi}(\mathcal{L}(\rho)) \\ &= 0\end{aligned} \tag{5.22}$$

where the second last line comes from $[\mathcal{L},\mathcal{U}_{\theta,\phi}]=0$. Therefore σ is also a steady state of the master equation and from part 1, there are no coherences in λ. This state, is also invariant under the action of the diagonal subgroup since $[\sigma,\mathcal{U}_{\theta,\phi}]=0$ and therefore, there are also no coherences in w and y:

$$\langle\lambda,W,w,y|\sigma|\lambda',W',w',y'\rangle = 0 \text{ if } w\neq w',\ \lambda\neq\lambda' \text{ or } y\neq y'. \tag{5.23}$$

Note that depending on the irrep there may be degenerate weight spaces that have the same w and y but different W values. This is true for any free Hamiltonian, and thus there can still be coherences in W. Since part 1 also tells us that within the Hilbert space of each irrep, the steady

[3] e.g. $e^{-i\theta W_z}U_-U_+e^{-i\theta W_z} = U_-U_+e^{i\theta W_z}/2 - U_-U_+e^{i\theta W_z}/2 + U_-U_+e^{-i\theta W_z}e^{-i\theta W_z}$. Similar relations can be derived for the other non-diagonal generators and also for Y so operators of the form $U_{\theta,\phi}$ commute with any term containing a pair of raising and lowering operators as would appear in the master equation.

state is unique, therefore the steady state σ is the only steady state and has the block diagonal structure given in Eq. (5.18).

5.4.3 Favouring specific irreps in the steady state

The structure of the steady state leads to the condition that the probabilities to occupy each block in the Hilbert space are fixed in time under evolution by an equation like (5.15). These probabilities can then in principle be chosen by the preparation of the initial state. The steady state can then be calculated for each block independently and the steady state of the entire Hilbert space can then be reconstructed by performing the direct sum over the blocks noting that the probability to occupy a particular block p^λ is fixed by the state preparation and that the multiplicity m_λ of the block must be taken into account.

One way that the system could be prepared with some specifically desired distribution over the irreps is where each particle is initially thermalised separately with an environment of some inverse temperature β_0 and then the particles are brought together into an ensemble. The initial state is then $\rho_{\beta_0} = \exp(-\beta_0 H_S)/Z_{\beta_0}$, where $Z_{\beta_0} = \mathrm{tr}[\exp(-\beta_0 H_S)] = \mathrm{tr}[\exp(-\beta_0 h)]^n$.

In the reduced Schur basis, this state becomes

$$\rho_{\beta_0,\mathrm{acc}} = \bigoplus_\lambda p^\lambda \frac{e^{-\beta_0 H_S^\lambda}}{Z_{\beta_0}^\lambda}, \quad p^\lambda = \frac{m_\lambda Z_{\beta_0}^\lambda}{Z_{\beta_0}}. \tag{5.24}$$

Now consider two limiting cases. First, if the initial temperature is $\beta_0 \to \infty$, the system will approach the ground state and the fully symmetric subspace, $(p,q) = (n,0)$,that is, $\lambda = [n,0,0]$ will have $p^\lambda \to 1$. On the other hand when $\beta_0 \to 0$, $Z_{\beta_0}^\lambda \to d_\lambda$, so $p^\lambda \to m_\lambda d_\lambda / 3^n$.

Additionally, this allows us to make an approximation for the limit of a large number of particles n by focusing on a single 'typical' irrep [152,199,200]. If we prepare the system in some product state $\rho_0^{\otimes n}$ where ρ_0 has eigenvalues in decreasing order $\{e_1, e_2, e_3\}$, the distribution p^λ is peaked at the irrep with a Young diagram where the row lengths are in proportion to the eigenvalues e_i [201, 202]. That is the i-th row has length ne_i. In the (p,q) notation this is irrep $(p,q) = (n[e_1 - e_2], n[e_2 - e_3])$. The variance of the distribution over the irreps is of order n, therefore by making this approximation, fluctuations of order $\sqrt{n}$ are ignored in p and q. Thus the calculation of the steady state in this limit is greatly sped up since we can calculate quantities

for a single typical irrep and the Hilbert space dimension of this irrep will be of the order n^3 according to Eq. (5.11).

5.5 Three models of work extraction

We will now consider three models where work can be extracted from an ensemble in its steady state. First, in Sec. 5.5.1 we explore the amount of work that can be extracted by unitary operations from a system which has been brought into the collective analogue of a population inversion. In Sec. 5.5.2, we discuss how the system can be set up so that it does work on a dissipative load. Finally, in Sec. 5.5.3, we couple the system to an external driving field as well as the hot and cold reservoirs and discuss its operation as an amplifier.

As we have established in the previous section, we can calculate the steady state quantities we are interested in within a fixed irrep and then reconstruct the entire Hilbert space from the direct sum of the irreps. Therefore, to simplify notation, the label λ will often be dropped from operators when we are implicitly working within a fixed irrep.

5.5.1 Unitary work extraction from the steady state

For a single three level particle, the thermodynamic interpretation if the condition for maser action due to population inversion between $|g\rangle$ and $|e\rangle$, is given by the single-particle SSDB master equation [9],

$$\begin{aligned}\dot{\rho}_{\text{sing}} =& g_v(\bar{n}_c\mathcal{D}[|e\rangle\langle f|]\rho + (\bar{n}_c+1)\mathcal{D}[|f\rangle\langle e|]\rho) \\ &+ g_u(\bar{n}_h\mathcal{D}[|g\rangle\langle f|]\rho + (\bar{n}_h+1)\mathcal{D}[|f\rangle\langle g|]\rho).\end{aligned} \tag{5.25}$$

The steady state with $\omega_h\beta_h < \omega_c\beta_c$ will accumulate a population inversion and can be coupled to an external system which can then extract work by unitary evolution [180]. A state with a higher population inversion would allow for more work extraction.

This motivates the first model for work extraction from the collective ensemble. Here we assume no coupling on the ω_l transition and couple the system to a hot bath and cold bath. This

system reaches a steady state under the dynamics,

$$\dot{\rho} = g_u \mathcal{L}_U(\bar{n}_h)\rho + g_v \mathcal{L}_V(\bar{n}_c)\rho. \tag{5.26}$$

The amount of population inversion between the levels of any particular irrep can now be characterised by the maximum energy that can be extracted by unitary operations,

$$\mathcal{E}(\rho) := \max_U \left(\mathrm{tr}[H_S \rho] - \mathrm{tr}\left[H_S U \rho U^\dagger\right] \right). \tag{5.27}$$

This quantity is called the ergotropy [180]. The maximum energy is extracted by the unitary operation that maps the state $\rho = \sum_k r_k \left|\psi_k\right\rangle\!\left\langle\psi_k\right|$ onto the transformed state $\tilde{\rho} = \sum_k r_k^{\downarrow} \left|E_k\right\rangle\!\left\langle E_k\right|$ which is fully depleted of work. This state is called the passive state, it is diagonal in the energy eigenbasis and the populations $r_k^{\downarrow}$ are ordered in decreasing order. If ρ is diagonal in the energy eigenbasis, then the ergotropy is a measure of population inversion since U rearranges the populations in decreasing order with respect to increasing E_k. The operator U does not have to stay within the $SU(3)$ product representation, it must merely maintain permutation invariance. It is given in general by

$$U = \sum_k \left|E_k\right\rangle\!\left\langle\psi_k\right|. \tag{5.28}$$

In practice, it may only be possible to apply operations that are constrained to the $SU(3)$ product representation. For example the work extraction may be limited to work that is extracted from the lasing transition. This could be achieved by performing rotations only within this subspace with the operator $W_x = W_+ + W_-$. We call the energy extracted by these rotations the lasing ergotropy and it is defined by,

$$\mathcal{E}_l(\rho) := \max_\theta \left(\mathrm{tr}[H_S \rho] - \mathrm{tr}\left[H_S e^{-i\theta W_x} \rho e^{i\theta W_x}\right] \right). \tag{5.29}$$

Clearly, the full ergotropy will sometimes differ from the lasing ergotropy. How this happens is illustrated in Fig. 5.3. The populations of a typical steady state of the (4,0) irrep is illustrated here with the shaded circles, with the size of the circle proportional to the value of the steady state population. After rotation by the lasing ergotropy unitary, the populations would be mirrored along the y-axis, this is indicated on the diagram with dashed black circles. The values of the

Hamiltonian eigenvalues are indicated by the shading of the circles with values decreasing from red to blue. From this one can see that constructing the passive state generally requires transitions that go diagonally in the diagram and suggests that one may extract more energy from the state than just the energy that is available in the lasing transition.

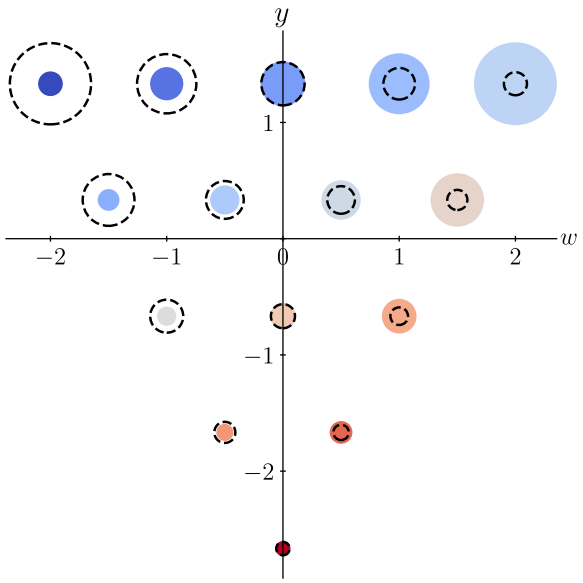


Figure 5.3: The size of each shaded circle at point (w_i, y_i) is proportional to the size of the steady state population of the state $|(4,0), W_i, w_i, y_i\rangle$ for Eq. (5.26) with the system prepared so that the hot and cold reservoirs have frequencies and inverse temperatures $\omega_c = 2/3, \omega_h = 5/3$, $\beta_c = 1.5$ and $\beta_h = 0.8$ with all frequencies in units of the lasing frequency ω_l and inverse temperatures in units ω_l^{-1}. The magnitude of the Hamiltonian eigenvalue at each point is indicated by the colour of the shading with the values decreasing from dark red to dark blue. The state rotated by W_x is shown by the dashed black circles. Adapted from [C].

In the steady state, it is possible to get a simplified expression for the lasing ergotropy. Notice first that, since Y commutes with the operators $W_\pm$, it is unchanged by rotation by W_x. Therefore, there is no contribution to the ergotropy from terms containing Y in the Hamiltonian and we are left only with the $\omega_l W_z$ term. Then using, $e^{i\theta W_x} W_z e^{-i\theta W_x} = -\cos\theta W_z - \sin\theta W_y$ we obtain,

$$\mathcal{E}_l(\rho_\infty^\lambda) = \omega_l \max_\theta \operatorname{tr}\left\{\left[(1-\cos\theta)W_z - \sin\theta W_y\right]\rho\right\}. \tag{5.30}$$

This is simplified further in the steady state since ρ_∞ is diagonal in w and therefore it is invariant under rotations generated by W_z so $\mathrm{tr}\{W_y\rho\} = \langle W_y\rangle = 0$. Then either $\langle W_z\rangle \leq 0$ and no work is extractable through the lasing transition or $\langle W_z\rangle > 0$. In this case, the maximum work is extracted when $\cos\theta = -1$ and the optimal rotation angle is thus $\theta = \pi$.

Numerical observations

In contrast to the description for a single-particle or for systems restricted only to the fully symmetric subspace, it is difficult to characterise the steady state for hot and cold baths with generic temperatures. We can make some observations numerically[4]. Numerical results for a collection of 4 particles are shown in Fig. 5.4. Here we assume that the state can be prepared with a particular symmetry type and then the dynamics are constrained to the subspace with that symmetry type. This figure shows how much energy could be extracted from each subspace. The first observation is that the single-particle population inversion condition $\omega_c\beta_c > \omega_h\beta_h$ is the same as the condition for the lasing ergotropy to be non-zero. Additionally, we observe that the fully symmetric subspace, $(n, 0)$, will always allow the maximum energy to be extracted. We also compare each subspace to four individually coupled particles. In Fig. 5.4(a) we see that when $\omega_c\beta_c \gg \omega_h\beta_h$, the energy extracted from independent particles (top of shaded of shaded region) is identical to the energy extracted from symmetric particles (red lines) and the lasing and full ergotropy are equal. In Fig. 5.4(b), a less extreme temperature difference is considered. Here, we see that it is possible to extract more energy from the state than just the work extracted from the lasing transition and that this additional work resource can extend beyond the temperature window of lasing. Additionally, individually coupled particles do not produce as much energy as collectively coupled particles in the fully symmetric state. In both figures, we see that The maximum energy will be extracted from the subspace with the highest value of p, and this will always be the fully symmetric subspace, i.e., $p = n$, $q = 0$.

[4] This was done for the most part using standard algebraic packages from the Julia programming language [203] to solve the Lindbladian in each irrep. Special functions for Young tableaux that allow extensions to higher dimensions eg. the four level system presented in [C] come from the JuLie package [204].

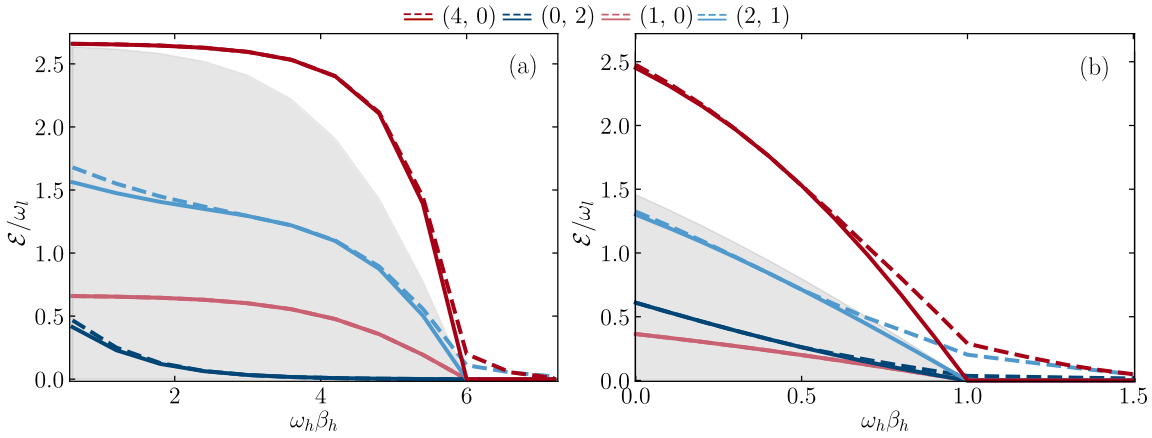


Figure 5.4: The ergotropy (solid lines) and lasing ergotropy (dashed lines) are compared here for the different irreps of four three level systems. The cold reservoir is fixed in (a) to $\omega_c = 2/3$, $\beta_c = 10.0$ and the hot reservoir temperature is changed while the frequency is fixed to $\omega_h = 5/3$. The coupling to both reservoirs is fixed to $g_u, g_v = 0.1$. The irreps are indicated in the key at the top of the figure using (p, q) notation. The lasing ergotropy of four independent particles is shown by the upper border of the grey shaded region. When the cold bath is set to this very low temperature, the full and lasing ergotropy are maximised for high hot bath temperatures.For intermediate temperatures, depicted in (b) where, $\omega_c = 2/3$, $\beta_c = 1.5$, neither collective nor independent particles reach this limit even for high hot bath temperatures. All frequencies and couplings are given in units of the lasing frequency ω_l and inverse temperatures in units ω_l^{-1}. Adapted from [C].

Cold temperature limit

In the limit of $\omega_c\beta_c \gg 1$ (for example the far left of Fig 5.4(a)), it is possible to derive an analytical result for the lasing ergotropy. In this limit the average occupation number of the cold bath is effectively zero. The master equation is then,

$$\dot{\rho} \stackrel{!}{=} g_u\mathcal{L}_U(n_h)\rho + g_v(V_+\rho V_- - \frac{1}{2}\{V_-V_+, \rho\}). \tag{5.31}$$

Since the steady state is unique and takes the form Eq. (5.23), we can compute the dynamics using the weight diagram as a guide. Define a family of lines in the space of basis vectors of the irrep that can be traversed by the $U_\pm$ operators. These lines are indexed by the numbers $L = y/2 + w$, where y and w are the values of the weights at a particular point in the space.

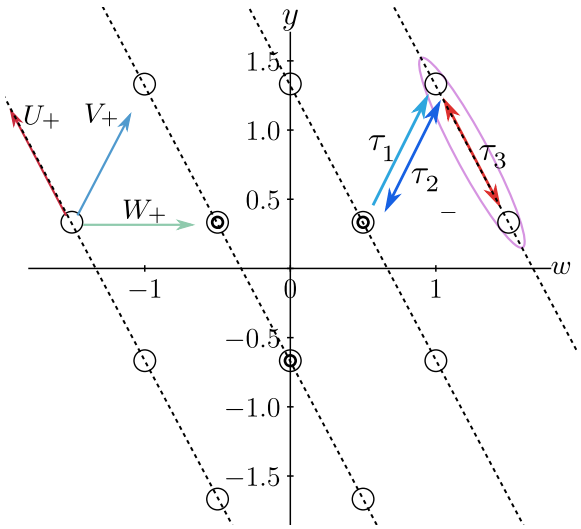


Figure 5.5: A diagram of the reduced basis states $|(p,q),W_i,w_i,y_i\rangle$ of one of the irreducible representations found for four three level systems, labelled $\lambda = (p,q) = (2,1)$. Dotted lines form the family of lines with equation $L = y/2 + w$. The arrows labelled by τ_i depict the action of each of the terms in Eq (5.33). The points inside the lilac ellipse form the line L_* on which the steady state lives for a low-temperature cold bath. The reflected steady state ρ_r is made up of the populations inside the purple dashed ellipse. Adapted from [C].

These are the dotted lines in Fig 5.5, which illustrates the weight space of the $(2,1)$ irrep. This irrep is used as an illustration but the argument to get to the steady state will work for any irrep. We will show that in the steady state all the populations are concentrated in the states found on the right-most line.

To do this, first define the sum of the time derivatives of the populations along each of these lines,

$$\sum_{i\in L} \dot{\rho}^i_L := \sum_{w_i+y_i/2=L} \langle (p,q),W_i,w_i,y_i|\,\dot{\rho}\,|(p,q),W_i,w_i,y_i\rangle \stackrel{!}{=} 0. \tag{5.32}$$

Here, the Schur basis states that lie on the line L for a given irrep (p,q) are indexed by i. To determine the expression for the steady state we break up this sum into three parts.

$$\sum_{i\in L} \dot{\rho}^i_L = \tau_1^{(L)} + \tau_2^{(L)} + \tau_3^{(L)} \tag{5.33}$$

$$
\begin{aligned}
&= g_v \sum_{i\in L} \langle (p,q), W_i, w_i, y_i |\, V_+ \rho V_- \,| (p,q), W_i, w_i, y_i\rangle \\
&- \frac{g_v}{2} \sum_{i\in L} \langle (p,q), W_i, w_i, y_i |\, \{V_- V_+, \rho\} \,| (p,q), W_i, w_i, y_i\rangle \\
&+ g_u \sum_{i\in L} \langle (p,q), W_i, w_i, y_i |\, \mathcal{L}_U \rho \,| (p,q), W_i, w_i, y_i\rangle \,.
\end{aligned}
$$

We will use the result [170] that the populations are constant along each line in the steady state with both $\tau_1^{(L)} + \tau_2^{(L)} = 0$ and $\tau_3^{(L)} = 0$ on every line. This is illustrated in Fig 5.5. The direction of transitions between basis elements generated by the three parts of the equation are shown on the right using the blue arrows for the two parts describing the interaction with the cold bath. The double headed red arrow describes the Lindblad term generated by $\mathcal{L}_U$.

Now we look at the other two terms for each line starting from the line, $L_0 = (p-2q)/3$, that is furtherest to the left in Fig 5.5. On this line $\tau_1^{(L_0)} = 0$ since there are no states on the line$L_{(-1)} = L_0 - 1$. since we consider only the steady state, this leaves $\tau_2^{(L_0)} = 0$. Then by using the the operator coefficients given in App D, we see that

$$
\begin{aligned}
V_- V_+ \,| (p,q), W, w, y\rangle = &(B^{p,q}_{W+1,w,y} A^{p,q}_{W,w,y} + A^{p,q}_{W-1,w,y} B^{p,q}_{W,w,y}) \,| (p,q), W, w, y\rangle \\
&+ (A^{p,q}_{W,w,y})^2 \,| (p,q), W+1, w, y\rangle \\
&+ (B^{p,q}_{W,w,y})^2 \,| (p,q), W-1, w, y\rangle \,.
\end{aligned} \tag{5.34}
$$

The coefficients are all non-negative and so the expectation value of each term in τ_2 is always positive. But since $\tau_2^{(L_0)} = 0$ the total population of each point on this line must be zero.

We then proceed through the remaining lines in the irrep. First, each term in $\tau_1^{(L+1)}$ vanishes since they all depends on the populations from the preceding line which are all zero. Then we are left with $\tau_2^{(L+1)} = 0$ and $\tau_3^{(L+1)} = 0$. By the same logic as before, the populations on the line must all be individually zero for the sum in $\tau_2^{(L+1)}$ to be zero. This logic is repeated until the last line in the state space $L_* = (2p+q)/3$. On this line the term $\tau_2^{(L_*)} = 0$ automatically, since no states are reached by transitions with V_+. Similarly, each term in $\tau_1^{(L_*)}$ vanishes because there are no non-zero populations on the previous line. Finally, $\tau_3^{(L_*)} = 0$, and the populations

on this line are determined only by the hot bath,

$$\dot{\rho}_{(2p+q)/3} = g_u \mathcal{L}_U \rho_{(2p+q)/3}. \tag{5.35}$$

The populations on this line are guaranteed to be non-degenerate since they are on the border of the weight diagram. The steady state has a Boltzmann distribution over the $q+1$ states on this line. The magnitude of the populations will decrease with decreasing y giving the k^{th} population

$$\rho_\infty^k = \frac{e^{-k\beta_h\omega_h}}{Z}, \tag{5.36}$$

with $Z = \sum_{k=0}^{q} e^{-k\beta_h\omega_h}$.

The lasing ergotropy can then be calculated for this steady state.

$$\mathcal{E}_l = 2\omega_l \operatorname{tr}\left[W_z \rho_\infty^\lambda\right] \tag{5.37}$$

The w coordinates of this line are $w_k = (p-k)/2$ with $k = 0, 1, ..., q$, which gives

$$\mathcal{E}_l = \omega_l \sum_{k=0}^{q} \frac{e^{-k\beta_h\omega_h}}{Z}(p+k) = \omega_l \left[p + \sum_{k=0}^{q} \frac{k e^{-k\beta_h\omega_h}}{Z} \right]. \tag{5.38}$$

By performing the sum, we obtain

$$\mathcal{E}_l(\rho_\infty^\lambda) = \omega_l \left[p + q + \frac{1}{1-e^{-\beta_h\omega_h}} - \frac{q+1}{1-e^{-(q+1)\beta_h\omega_h}} \right]$$
$$(\text{for } \omega_c\beta_c \gg 1). \tag{5.39}$$

We know that for a particular irrep, $\mathcal{E}_l \leq \mathcal{E}$, but in this very low temperature cold bath limit we can also find the condition for $\mathcal{E}_l = \mathcal{E}$. Equality happens when the state which is a reflection of the steady state along the y-axis, $\rho_r = e^{-i\pi W_x} \rho_\infty^\lambda e^{i\pi W_x}$ is a passive state. This state lies on the line L_r which is the reflection of L_*. The populations of the states decrease by the Boltzmann ratio $e^{-\beta_h\omega_h}$ with decreasing y-coordinate and the energies of the states on the line L_r also decrease with decreasing y. This means that ρ_r is passive whenever the states on L_r are

the $q+1$ lowest energy eigenstates.

Therefore, whether $\mathcal{E}_l = \mathcal{E}$ is true in the low temperature limit, depends only on the Hamiltonian. First, the ground state is always the upper left-most point in the diagram. The point on the line L_r with the next lowest energy is obtained by applying V_- which increases the energy of the state by ω_c. This continues down the line until the highest energy state of the line is reached with energy $q\omega_c$ relative to the ground state. However, from the ground state depending on the shape of the diagram, both U_- and W_+ may also be possible transitions which increase the energy by ω_h and ω_l respectively with $\omega_h > \omega_l$ by definition. The only way to get all $q+1$ lowest energy states on the line L_r is when $q\omega_c \leq \omega_l$ and this is the condition for $\mathcal{E}_l = \mathcal{E}$. The ergotropies in the limit of a very low temperature cold bath are illustrated in Fig. 5.4(b). In the example illustrated, the lasing and full ergotropies do not coincide for all of the irreps. However, one can see that the lasing ergotropy reaches the maximum given in Eq. (5.39) for very large hot bath temperatures.

Scaling of the ergotropy with n

Finally, we can look at the scaling behaviour of the ergotropy and how the choice to favour different irreps affects this behaviour. In Fig 5.6 we compare the performance of a collective system made up of n particles with the equivalent number of independent particles with their dynamics each individually governed by the single particle master equation (5.14). The ergotropy of all the individually coupled particles together is then the sum of single particle ergotropies. In both cases we individually prepare each of the n particles in a thermal state at some fixed temperature β_0. Then for the collectively coupled engine, we bring them together into the collective state that interacts with the hot and cold baths. The collective system then has a fixed probability distribution over the irreps as described in sec 5.4.3

$$p^\lambda = \frac{m_\lambda Z^\lambda_{\beta_0}}{Z_{\beta_0}}, \tag{5.40}$$

which is determined by this initial product state. This preparation will not affect the performance for individual particles, since each independent particle equilibrates individually and they do not have a block diagonal Hilbert space with different irreps λ. The optimal ergotropy is

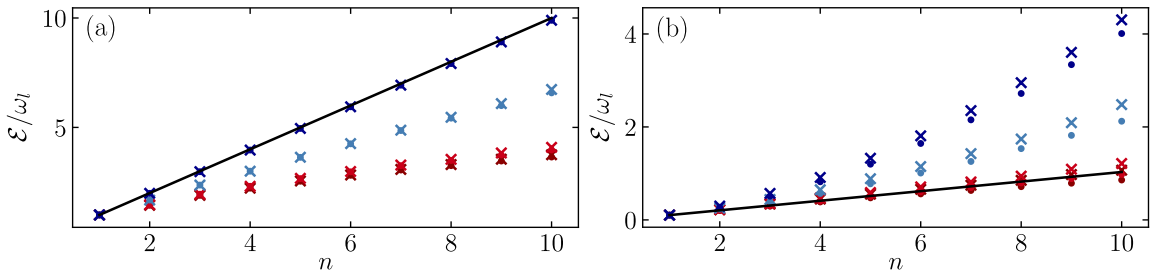


Figure 5.6: The scaling of the full ergotropy (crosses) and lasing ergotropy (points) with the number of particles n is compared here for states prepared in an initially thermal state. The initial temperature of the state $\beta_0 = \{0.09, 0.45, 1.5, 5.0\}$ is indicated by the colour of the markers from dark red to blue. The ergotropy obtained from n individual particles is indicated by the black line. Two scenarios are explored. First (a) the limit when $\beta_c \to \infty$ where it is only possible to extract as much work from collectively coupled systems as individually coupled systems when the initial state is prepared at a low enough temperature. Next, (b) an intermediate temperature is chosen for the cold bath $\beta_c = 1.5$. Here, collectively coupled systems are able to perform better than independent systems even when they are prepared at moderate $\beta_c = 0.45$ temperatures. This effect is enhanced for large n. In all scenarios, $\omega_c = 2/3$, $\omega_h = 5/3$, $\beta_h = 0.8$ and the coupling to reservoirs is $g_u, g_v = 0.1$. All frequencies and couplings are given in units of the lasing frequency ω_l and inverse temperatures in units ω_l^{-1}. Adapted from [C].

reached when the cold bath temperature is very small. In this case, independent particles or collective particles prepared in a very low temperature thermal state produce the optimal ergotropy. Therefore, there is no advantage to using a collectively coupled system, especially since any contributions from irreps other than the fully symmetric irrep only worsen the performance. On the other hand when the cold bath is set to an intermediate temperature with respect to the hot bath, preparing the state at a small initial temperature leads to collective performance that is better than individual particles. This is true even for moderate β_0 where irreps other than the fully symmetric subspace contribute to the total performance. However, it remains true that a distribution very sharply peaked about the fully symmetric subspace is optimal.

Although the block diagonal structure allows calculations for multiple particles without approximations beyond the Born, Markov and Secular approximations used to derive the Master equation, at some point even with high computational power, further approximations will have to be made in the large n limit. One way that this could be done is as follows. Suppose that initially each particle has the single-particle state ρ_1, with eigenvalues $\{r_1, r_2, r_3\}$ in decreasing

order. For the very large particle limit, we can assume that there is a typical irrep that contributes the most to the steady state with $(p, q) = (n[r_1 - r_2], n[r_2 - r_3])$ (see Sec. 5.4.3). This allows us to calculate the result only for this irrep and ignore the rest of the state space. Additionally, we can use the result (Eq. (5.39)) for the lasing ergotropy when $\beta_c \to \infty$, which depends only on ω_l, p and q. The ergotropy in this limit is then $\mathcal{E}_l \approx \omega_l p = \omega_l n(r_2 - r_3)$.

We could also prepare the particles in a single-particle steady state with $\omega_c\beta_c = \omega_h\beta_h$ before bringing them together into the collective ensemble which subsequently interacts with the baths. Again, this initial state preparation will determine the distribution over the different irreps of the collective system which persists even after the collective system reaches its steady state. For this single particle initial state, the levels $|g\rangle$ and $|e\rangle$ have equal populations and thus, even though it is not a thermal state, the single-particle state has zero ergotropy and is passive. For large n, the many copy state $\rho^{\otimes n}$ is not passive [205]. It has a typical irrep with parameters $p = 0$ and $q \approx n(1-e^{-\beta_c\omega_c})/(2+e^{-\beta_c\omega_c})$. This corresponds to a triangular weight diagram that is flipped vertically with respect to the fully symmetric weight diagram. The system is thus able to extract work from a system with symmetry that is far from bosonic exchange symmetry.

5.5.2 Power emitted into a dissipative load

The ergotropy can be used to determine if the steady state has population inversion. However, it would also be beneficial to be able to determine the efficiency of the engine under continuous operation in analogy to the SSDB engine. A simple model that admits a description of the rate of energy that is absorbed and emitted is studied in this section. Work can be extracted from the system by coupling it to a work reservoir that functions like a battery or flywheel [33, 36, 206–208]. The energy that would be captured can also be quantified by coupling the system to an external load [26, 209]. This is a simplistic analogy for doing work against friction [186]. We model this by coupling the lasing ($W_\pm$) transition to a zero temperature bath in addition to the coupling to the cold and hot baths with finite inverse temperatures β_c and β_h. This bath then functions as a dissipative load. The master equation describing this model is,

$$\dot{\rho} = g_u\mathcal{L}_U(\bar{n}_h)\rho + g_v\mathcal{L}_V(\bar{n}_c)\rho + g_w\mathcal{D}[W_-]\rho. \tag{5.41}$$

The total internal energy of the system is $\mathrm{tr}\{\rho H_S\}$. From the first law of thermodynamics, since the Hamiltonian is time independent, we have that the change in internal energy must equal the sum of the energy currents supplied by each of the baths [11,13]. The energy currents from the hot and cold baths are then

$$\mathcal{I}_h = \mathrm{tr}[g_u \mathcal{L}_U(\bar{n}_h)\rho H_S], \quad \mathcal{I}_c = \mathrm{tr}[g_c \mathcal{L}_V(\bar{n}_c)\rho H_S]. \tag{5.42}$$

These expressions can be simplified using the commutators of the operators $V_\pm$ and $U_\pm$ with the Hamiltonian to obtain

$$\mathcal{I}_h = \omega_h g_u \,\mathrm{tr}\{[(\bar{n}_h + 1)U_- U_+ - \bar{n}_h U_+ U_-]\rho\} \tag{5.43}$$

$$\mathcal{I}_c = \omega_c g_c \,\mathrm{tr}\{[(\bar{n}_c + 1)V_- V_+ - \bar{n}_c V_+ V_-]\rho\}. \tag{5.44}$$

The energy current flowing out of the system into the dissipative load is

$$\mathcal{P} = -\,\mathrm{tr}\{H_S g_w \mathcal{D}[W_-](\rho)\} = \omega_l g_w \,\mathrm{tr}\{W_+ W_- \rho\}. \tag{5.45}$$

This expression can be identified with the power [210]. The efficiency of the engine, $\eta = \mathcal{P}/\mathcal{I}_h$, is the ratio of these two quantities. For a single-particle it is already known [13, 179] that the efficiency depends only on the frequencies and not the temperatures of the hot and cold baths and that we get the expression, $\eta = 1 - \omega_c/\omega_h$. To determine the efficiency for the collective system, first define the operator that counts the number of particles in the ground state. For a single-particle this is the operator $N_g = |g\rangle\langle g|$ which, using the dynamical operators of our system is $N_g = n/3 + Y/2 - W_z$. Consider the time derivative of the expectation value of this operator calculated in the standard way with the adjoint master equation[5],

$$\partial_t \langle N_g \rangle = \langle g_u[(\bar{n}_h + 1)U_- U_+ - \bar{n}_h U_+ U_-] + g_w W_+ W_- \rangle$$

[5]The adjoint superoperator $\mathcal{L}^*$ acting on the Hamiltonian is

$$\begin{aligned}\mathcal{L}^*_U(H_S(t)) =& \bar{n}_h \mathcal{D}^*[U_-](H_S(t)) \\ &+ (\bar{n}_h + 1)\mathcal{D}^*[U_+](H_S(t)).\end{aligned} \tag{5.46}$$

Here, $\mathcal{D}^*[O](H_S(t)) = O^\dagger H_S(t) O - \frac{1}{2}\{O^\dagger O, H_S(t)\}$, with an equivalent expression for the cold reservoir.

$$= \frac{\mathcal{I}_h}{\omega_h} - \frac{\mathcal{P}}{\omega_l}, \tag{5.47}$$

which vanishes in the steady state. Therefore, $\mathcal{P}/\mathcal{I}_h = \omega_l/\omega_h = 1 - \omega_c/\omega_h$ and we see that a collective engine has the same efficiency as a single-particle engine.

Therefore, to determine how they differ, we compare the performance of these engines on other grounds. We consider first the total energy current emitted into the zero temperature reservoir. Since the steady state may have a complicated form for arbitrary irreps and finite hot and cold bath temperatures, we first make some numerical observations. The energy emitted by a collection of four particles is compared in Fig. 5.7. Here we can see that the fully symmetric irrep performs better than both the $(2,1)$ irrep as well as the $(0,2)$ irrep. However, independent particles sometimes emit more energy than even the fully symmetric irrep. From Eq. (5.45), we see that the emitted energy is proportional to the coupling to the zero temperature bath. This relationship is apparent in Fig 5.7, however, the aim of the engine is to perform work on the load and so it is necessary to analyse not just the magnitude of the emitted energy but also whether this energy resembles work more than it resembles heat.

Determining if the energy transmitted resembles work

The classification of the energy as work or heat can be approached in different ways and is a field of ongoing research. Much of the debate arises from how to partition the change in internal energy of the system into a work-like and heat-like part. In particular, some approaches may allow for the environments that the working medium couples to, to exchange both work and heat simultaneously [211]. In this section, we couple the system to a dissipative load. In this case, whether the energy exchanged with this reservoir resembles work, depends on if it behaves differently from thermal energy in the particular application. In terms of the dissipative load of our example, if the energy transferred to an optical medium is associated with additional fluctuations, it resembles heat more than work. Thus, work-like energy transfer should be associated with emission of coherent light that has very few intensity or phase fluctuations. That is, the emitted light would resemble the light from an ideal laser. Therefore, if we consider this a defining property of work-like energy in our system we must be able to quantify the coherence of the light. To do this, we couple the output mode of the system, which is the transition with

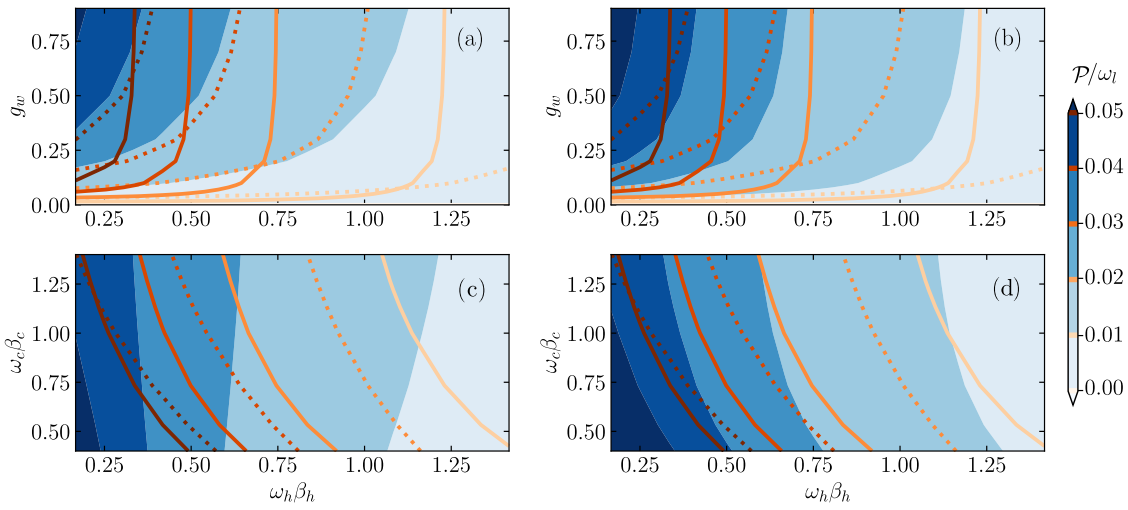


Figure 5.7: The energy current $\mathcal{P}$ emitted into the dissipative load is larger when the coupling to the load is stronger at a fixed cold bath temperature ($\beta_c = 1.1$) as shown in (a) and (b). Here, we compare the energy emitted by the different irreps of an ensemble of four particles the solid orange contours show the behaviour of the fully symmetric $(4, 0)$ irrep in all figures and the blue shaded regions show the $(2, 1)$ and $(0, 2)$ irreps in (a,c) and (b,d) respectively. The dotted contours show the energy emitted by four independently coupled particles. When the coupling is fixed to $g_w = 0.9$. In all cases, the frequencies of the hot and cold baths are set to $\omega_c = 2/3$ and $\omega_h = 5/3$ respectively and both baths couple to the system with a coupling strength of 0.1. All frequencies and couplings are given in units of the lasing frequency ω_l and inverse temperatures in units ω_l^{-1}. Adapted from [C].

frequency ω_l, to a 1-D waveguide [212]. This waveguide is a bosonic reservoir with modes that are indexed by ω. The description of the waveguide is done using the input-output formalism in quantum optics. The derivation of this given below for completeness and it closely follows the presentation found in Chapter 5 of [213].

The Hamiltonian of the system coupled to the waveguide within the rotating wave approximation is

$$H = H_S + \int_{-\infty}^{\infty} d\omega\ \omega b^\dagger(\omega)b(\omega) + i\int_{-\infty}^{\infty} d\omega\kappa(\omega)[b^\dagger(\omega)W_- - W_+b^\dagger(\omega)], \tag{5.48}$$

where, the rotating wave approximation allows us to drop rapidly oscillating terms like $W_-b(\omega)$ from the interaction term. This allows the integrals which should formally only be over the range $[0,\infty)$ to be extended to the range $(-\infty,\infty)$ since only terms that are close to resonant contribute significantly to the result. The Heisenberg equation of motion of the Bosonic mode is,

$$\dot{b}(\omega) = -i\omega b(\omega) + \kappa(\omega)W_- \tag{5.49}$$

We can then get a solution for the bosonic mode for $t_0 < t$ (at the input) and $t < t_1$(at the output),

$$b(\omega) = e^{-i\omega(t-t_0)}b_0(\omega) + \kappa(\omega)\int_{t_0}^{t} e^{-i\omega(t-t')}W_-(t')dt' \tag{5.50}$$

$$b(\omega) = e^{-i\omega(t-t_1)}b_1(\omega) + \kappa(\omega)\int_{t}^{t_1} e^{-i\omega(t-t')}W_-(t')dt' \tag{5.51}$$

where, $b_0(\omega)$ and $b_1(\omega)$ are the values of $b(\omega)$ at $t = t_0$ and $t = t_1$ respectively. The Heisenberg equations of motion for the system operators are,

$$\dot{W}_- = -i[W_-, H_S] + 2\int d\omega\kappa(\omega)W_z b(\omega) \tag{5.52}$$

$$\dot{W}_+ = -i[W_+, H_S] + 2\int d\omega\kappa(\omega)b(\omega)^\dagger W_z \tag{5.53}$$

$$\dot{W}_z = -\int d\omega\kappa(\omega)[b(\omega)^\dagger W_- + W_+b(\omega)] \tag{5.54}$$

Now, we apply the first Markov approximation [214], that assumes the coupling is constant $\kappa(\omega) = \sqrt{\gamma/2}$ and thus the operator evolution is Markovian. We use this to define the input and output fields,

$$b_{\text{in}}(t) = \frac{1}{\sqrt{2\pi}} \int d\omega e^{-i\omega(t-t_0)} b_0(\omega) \tag{5.55}$$

$$b_{\text{out}}(t) = \frac{1}{\sqrt{2\pi}} \int d\omega e^{-i\omega(t-t_1)} b_1(\omega) \tag{5.56}$$

The system operators can then be expressed in terms of these expressions for the input field and output field. This yields the quantum Langevin equations of the system which are formulated separately for the input and output fields. For the input fields they are,

$$\dot{W}_- = -i[W_-, H_S] - 2W_z \left[\frac{\gamma}{2} W_- + \sqrt{\gamma} b_{\text{in}}(t) \right] \tag{5.57}$$

$$\dot{W}_+ = -i[W_+, H_S] + 2 \left[\frac{\gamma}{2} W_+ + \sqrt{\gamma} b_{\text{in}}^\dagger(t) \right] W_z \tag{5.58}$$

$$\dot{W}_z = -W_+ \left[\frac{\gamma}{2} W_- + \sqrt{\gamma} b_{\text{in}}(t) \right] - \left[\frac{\gamma}{2} W_+ + \sqrt{\gamma} b_{\text{in}}^\dagger(t) \right] W_- \tag{5.59}$$

with equivalent expressions for the output field.Then integrating Eq. (5.50) and Eq. (5.51), we obtain the expressions,

$$\frac{1}{\sqrt{2\pi}} \int d\omega b(\omega) = b_{\text{in}}(t) + \frac{\sqrt{\gamma}}{2} W_-(t) \tag{5.60}$$

$$\frac{1}{\sqrt{2\pi}} \int d\omega b(\omega) = b_{\text{out}}(t) - \frac{\sqrt{\gamma}}{2} W_-(t) \tag{5.61}$$

and get the following relationship between the input and output field,

$$b_{\text{out}}(t) = b_{\text{in}}(t) + \sqrt{\gamma} W_-(t). \tag{5.62}$$

We use this expression to analyse the coherence properties of the emitted field using quantities based on the Glauber coherence functions [215]. In particular, we will use the first and second order coherence functions in what follows. The first order correlator, $G^{(1)}(t, t') =$

$\left\langle b_{\text{out}}(t)^\dagger b_{\text{out}}(t)\right\rangle$ can be calculated by noting that since the input field is a vacuum, $b_{\text{in}}(t)\rho = 0$ and so $G^{(1)}(t,t') \propto \langle W_+(t)W_-(t')\rangle$. We will now use the quantum regression theorem [213, 216] to determine the value of the correlator. This is derived here for convenience.

Let σ be the system and reservoir density matrix so that, $\text{tr}_E(\sigma) = \rho$. Then, since $W_\pm(0)$ only act on the system part of the Hilbert space, the first order correlator is

$$\langle W_+(t)W_-(t')\rangle = \text{tr}_S\{W_-(0)\,\text{tr}_E[e^{-iH(t'-t)}\sigma(0)W_+(0)e^{iH(t'-t)}]\}. \tag{5.63}$$

Then set $\sigma_T(t'-t) = e^{-iH(t'-t)}\sigma(0)W_+(0)e^{iH(t'-t)}$ and note that for system dynamics generically given by $\dot\rho = \mathcal{L}\rho$,

$$\text{tr}_E[\sigma_T(t'-t)] = e^{\mathcal{L}(t'-t)}(W_-(0)\rho(t)W_+(0)). \tag{5.64}$$

Therefore for the current system we can calculate the correlator for the steady state ρ_∞,

$$G^{(1)}(t,t') \propto \text{tr}\{W_-(0)e^{(g_u\mathcal{L}_U+g_v\mathcal{L}_V)(t'-t)}[\rho_\infty W_+(0)]\}. \tag{5.65}$$

Intuitively speaking, if the energy transferred by the medium is free of additional fluctuations it resembles work. This would be associated with the emission of coherent light, as would be the case for an ideal laser, which has the minimal possible simultaneous fluctuations in intensity and phase. We will therefore, identify the work-like energy that is emitted with the coherent part of the emitted field. To determine the coherence properties of the emitted field, first define the normalised intensity spectrum [217],

$$S(\omega) = \frac{1}{2\pi}\left[\int_0^\infty d\tau\, G^{(1)}(\tau,\tau)\right]^{-1}\int_0^\infty d\tau\int_0^\infty d\tau'\, e^{-i\omega(\tau-\tau')}G^{(1)}(\tau,\tau'). \tag{5.66}$$

In the steady state, this expression for the intensity spectrum can be simplified by defining, $G^{(1)}(\tau) = G^{(1)}(t+\tau,t)$ for any t,

$$S(\omega) = \frac{\frac{1}{2\pi}\int_{-\infty}^\infty d\tau\, e^{-i\omega\tau}G^{(1)}(\tau)}{G^{(1)}(0)} =: \frac{P(\omega)}{P_{\text{tot}}}. \tag{5.67}$$

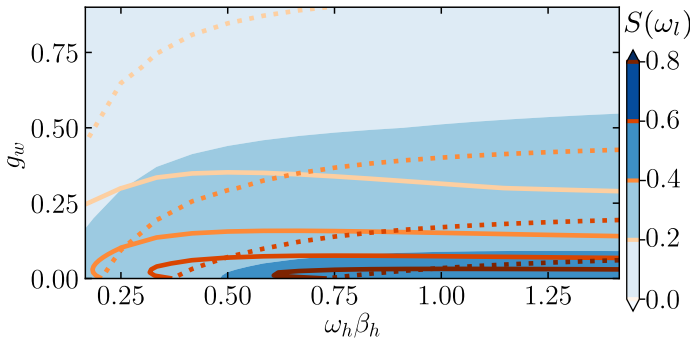


Figure 5.8: The peak value $S(\omega_l)$ of the intensity spectrum for the (4,0) and (2,1) irreps of a four particle collective ensemble are shown here with the solid orange contours and shaded blue areas respectively. We see that for a smaller coupling strength the quality of the energy that is emitted is more work like since the intensity spectrum is more sharply peaked. This is compared to individual particles (dotted orange contour) which show similar behaviour. Here, the frequencies of the hot and cold reservoirs are set to $\omega_c = 2/3$ and $\omega_h = 5/3$ and both baths have a coupling strength of $g_u, g_v = 0.1$. Finally, the temperature of the cold bath is fixed to $\beta_c = 1.1$. All frequencies and couplings are given in units of the lasing frequency ω_l and inverse temperatures in units ω_l^{-1}. Adapted from [C].

The numerator of this expression is the photon flux per unit frequency into mode ω of the waveguide. P_{tot} is the total photon flux [218]. In driven systems, $S(\omega)$ will have a delta peak and the weight of the delta peak of at the lasing frequency has been shown to be proportional to the power output [212,219]. In our example, since there is no driving, the intensity spectrum has no delta peak. Instead it is a continuous distribution over all frequencies ω with a peak at ω_l. If the integral over frequencies of $S(\omega)$ is taken over a very small frequency window about ω_l, this will give the proportion of the energy that is emitted into the mode with frequency ω_l. This works since the linewidth is much smaller than ω_l and so all the modes in the narrow window have approximately the same energy. A value of $S(\omega_l)$ closer to 1 is then characteristic of a more work-like quality of the output as this means the spectrum is very sharply peaked. The value of $S(\omega_l)$ is illustrated for a collection of four particles for different system parameters in Fig. 5.8. Here, we see that for collective as well as independent particles, the output is more work-like, when the coupling to the dissipative load is not too strong. Intuitively, a strong coupling to the zero temperature bath will lead to the dynamics being dominated by this bath which prevents a build up of population inversion by the other two baths. Therefore, even though (from Eq. (5.45)

and Fig. 5.7) more energy is emitted into the load for higher g_w, the emitted field will just be a diverted heat flow from the other baths with a broader intensity spectrum.

We also quantify the intensity fluctuations of the light since as we mentioned, work-like energy will have smaller intensity fluctuations than thermal light. This is done using the normalised second order correlator,

$$g^{(2)}(t) = \frac{G^{(2)}(t,0)}{\langle W_+ W_- \rangle^2}, \tag{5.68}$$

where, $G^{(2)}(t,t') \propto \langle W_+(t)W_+(t')W_-(t')W_-(t)\rangle$ and $g^{(2)}(\tau) \to 1$ for $\tau \to \infty$. We consider $g^{(2)}(0)$ in particular as a standard quantifier of coherence [216, 220, 221]. Perfect coherent emission will have photons emitted randomly in a Poisson distribution which results in $g^{(2)}(0) = 1$. Light that is characteristic of thermal emission will have $g^{(2)}(0) > 1$ as there is a tendency for photons to be bunched in the emission. Individually coupled systems can only have a $g^{(2)}(0) = 0$ since two photons can never be emitted simultaneously. In Fig 5.9 this is shown for weak coupling to the load, since we already established from the intensity spectrum that this is required for work-like output. Here we see that the range of coupling strengths with small intensity fluctuations is narrow. We also observe that the range of coupling strengths for a given hot and cold bath temperature where $g^{(2)}(0)$ is close to one is very different for different irreps. We see generally that a higher temperature of the hot bath leads to the emitted light being more characteristic of coherent light. Finally, taking both Fig. 5.8 and Fig. 5.9 into consideration, we see that for a fixed weak coupling to the dissipative load there is an optimal hot bath temperature that gives a large value for $S(\omega_l)$ and a value of $g^{(2)}(0)$ close to one. This optimal temperature may be very different for different irreps and it may be difficult to maximise both figures of merit simultaneously.

5.5.3 Work from stimulated emission

Finally, we consider an engine where the lasing transition is coupled to a driving field. This corresponds to a laser that amplifies the driving field to produce coherent light which is the work output. The Hamiltonian of the system with an interaction term that comes from resonant [6]

[6] the driving frequency is the same as the lasing frequency

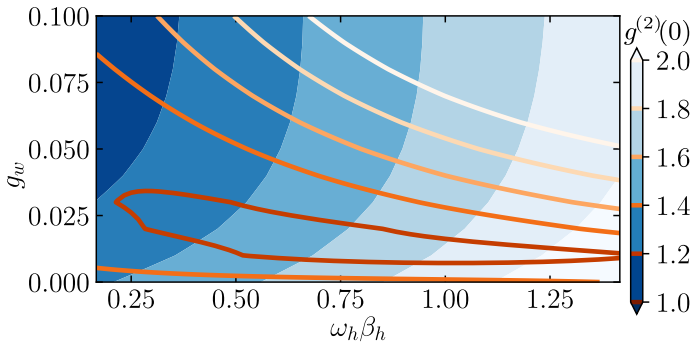


Figure 5.9: In this figure, we see that the coherence of the emitted field, quantified by $g^2(0)$ is quite different for the different irreps of a four particle ensemble. The (4,0) and (2,1) irreps of a four particle collective ensemble are shown here with the solid orange contours and shaded blue areas respectively. Work-like emission (closer to Poissonian emission statstics, $g^2(0)$ =1) occurs only when the coupling to the dissipative load is very small. Here, the cold bath parameters are set to $\omega_c = 2/3$ and $\beta_c = 1.1$. The hot bath has frequency $\omega_h = 5/3$ and both baths have a coupling strength of $g_u, g_v = 0.1$. All frequencies and couplings are given in units of the lasing frequency ω_l and inverse temperatures in units ω_l^{-1}. Adapted from [C].

driving is [216, 222]

$$H_S(t) = H_0 + \alpha(e^{-i\omega_l t}W_+ + e^{i\omega_l t}W_-). \tag{5.69}$$

All that is necessary to describe the system using the master equation from the previous sections (5.41), is to add the extra Hamiltonian term to the description. In the rotating frame[7], the Hamiltonian becomes $H_R = \alpha(W_+ + W_-)$ and the master equation is

$$\dot{\rho}_R = -i[H_R, \rho_R] + g_u\mathcal{L}_U(\bar{n}_h)\rho_R + g_v\mathcal{L}_V(\bar{n}_c)\rho_R. \tag{5.70}$$

Now, this master equation only gives a valid description under certain circumstances. First, the term obtained from the interaction Hamiltonian changes the Bohr frequencies of the system. This is not taken into account by the local master equation and this can lead to heat currents that give a total efficiency that exceeds the Carnot efficiency [187]. This is mitigated if only weak driving is allowed since the Bohr frequencies do not change by as much. Alternatively, the

[7] See (A.3), the rotating frame is a special case of the interaction picture which allows us to remove the time-dependence from the operator.

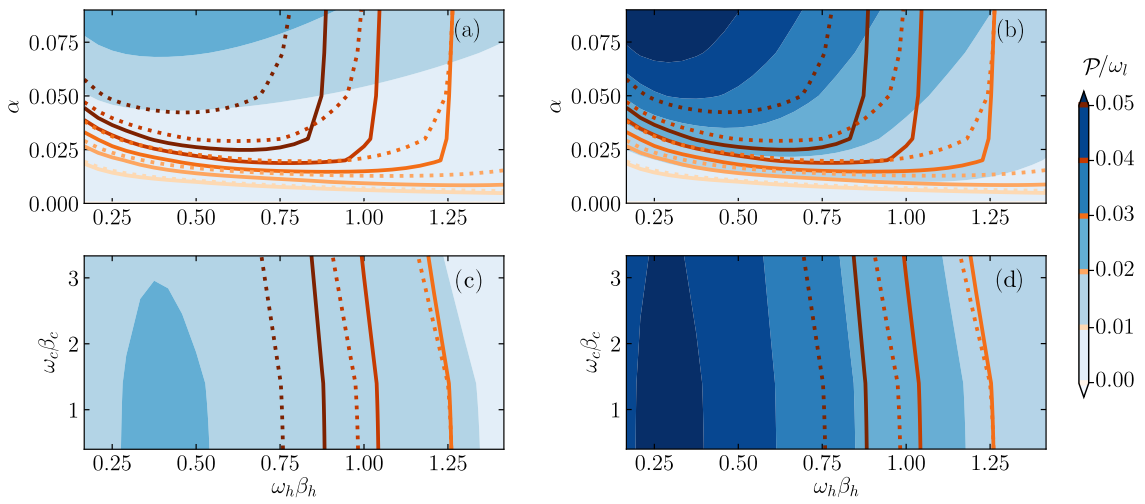


Figure 5.10: Here, the power emitted by four collectively coupled particles is compared to the the power emitted by four distinguishable particles (dotted orange contours). The solid orange contours show the behaviour of the fully symmetric $(4, 0)$ irrep in all figures and the blue shaded regions show the $(2, 1)$ and $(0, 2)$ irreps in (a,c) and (b,d) respectively. The systems are resonantly driven in the $W_{\pm}$ transition with driving strength α and emitted power increases with driving strength. In (a) and (b) this is shown for a cold reservoir set to $\beta_c = 1.1$. In (c) and (d), the reservoir frequencies are the same as in (a) and (b) but the driving strength is fixed to $\alpha = 0.07$ and in this case we see that the maximum power is not found for the minimum cold reservoir temperature. The coupling to the hot and cold reservoirs is fixed to 0.1 and the frequencies of the cold and hot baths are $\omega_c = 2/3$ and $\omega_h = 5/3$. All frequencies and couplings are given in units of the lasing frequency ω_l and inverse temperatures in units ω_l^{-1}. Adapted from [C].

changed Bohr frequencies can be taken into account by using the Floquet picture [175, 223].[8] However, we will consider only weak resonant driving.

Since the Hamiltonian is now time dependent, it is possible to break down the change in internal energy into two terms identified with the heat and power [11]. In the lab frame this is,

[8]The Floquet theorem states that the time evolution operator of a system with a periodic driving Hamiltonian can be decomposed into a periodic part and time independent part, $U(t, t_0) = e^{-iK(t)}e^{-i(H_F(t-t_0))}e^{iK(t)}$. Then, one can express all operators in the Floquet reference frame eg. $\rho_F = e^{iK(t)}\rho e^{-iK(t)}$. The master equation can then be derived in this reference frame and contributions from the shifted Bohr frequencies will be taken into account.

$$\dot{E} = \mathrm{tr}\{\dot{\rho}_S(t)H_S(t)\} + \mathrm{tr}\{\rho_S(t)\dot{H}_S(t)\}. \tag{5.71}$$

The total heat current into the system is given by the first term and the second term gives the power output of the system [212, 219]. In the rotating frame we obtain a steady state and this term can be simplified. The time dependence cancels due to the cyclic property of the trace and thus the power output is,

$$\mathcal{P} = -\alpha\omega_l \,\mathrm{tr}\{(W_- - W_+)\rho_R\} \tag{5.72}$$

The first term can be broken up into the heat currents from the cold and hot baths using the adjoint superoperators ((5.46)) acting on the Hamiltonian [13],

$$\begin{aligned}\mathcal{I}(t) &= \mathrm{tr}\{\rho_S(t)g_u\mathcal{L}_U^*(H_S(t))\} + \mathrm{tr}\{\rho_S(t)g_v\mathcal{L}_V^*(H_S(t))\} \\ &= \mathcal{I}_h(t) + \mathcal{I}_c(t).\end{aligned} \tag{5.73}$$

The steady state efficiency is once again, independent of the reservoir temperatures and is $\eta = 1 - \omega_c/\omega_h$. We can obtain this in the same way as we did in Sec. 5.5.2. Firstly in the rotating frame, the time derivative of the operator that counts the number of particles in the steady state (which again takes the form $N_g = n/3 + Y/2 - W_z$) must be zero.

$$\begin{aligned}\partial_t \langle N_g\rangle =& \langle g_u[(\bar{n}_h + 1)U_-U_+ - \bar{n}_h U_+U_-] + \alpha(W_- - W_+)\rangle \\ =& \frac{\mathcal{I}_h}{\omega_h} - \frac{\mathcal{P}}{\omega_l}\end{aligned} \tag{5.74}$$

Thus, $\mathcal{P}/\mathcal{I}_h = 1 - \omega_c/\omega_h$. Therefore, in a similar fashion to the previous section in order to quantify the differences between independent particles and the collective engine, the performance of the engine can be quantified by the absolute power output rather than the efficiency. Since $\mathcal{P}$ is proportional to the driving strength, the performance of the engine is expected to improve with increased driving strength. However, for large α, the driving is expected to dominate the dynamics and the performance should drop off. This does not occur for the weak driving case that we consider here. The numerical investigation of weak driving is illustrated in Fig. 5.10. Here, we see that stronger driving does lead to better engine performance and that the fully symmetric state performs better than the other irreps as we would expect from the res-

ults of the previous sections. We also show that independent systems perform worse than the fully symmetric subspace for a system of four particles. We see that for $\omega_h\beta_h \ll 1$ weak driving does not contribute significantly to the dynamics and therefore, the power output is small. We see a similar phenomenon with the cold bath temperatures, where a very large β_c leads to reduced performance. The results suggest there is an optimal driving strength for fixed hot and cold bath temperatures since we find that for the weak driving case we consider power output only increases with driving strength but for very strong driving the power output should drop off. Exploring this will require going beyond the current weak driving description. This would require the use of Floquet [175] or for large numbers of particles, Floquet-Redfield [223] master equations.

5.6 Conclusions

In this chapter, we went beyond the two-level models studied in the previous chapters and considered how multi-level systems interacting with an environment might present novel quantum effects. Although, these systems could be used for thermometry applications in analogy to the two level systems we discussed before, without a specific platform and measurement in mind, much of the advantages that could be gained from coupling collectively were studied previously in [151], since the quantum Fisher information is proportional to the heat capacity. However, the thermodynamics of these systems were not fully understood, especially for systems where different baths couple to different transitions in the system. We addressed this by studying the paradigmatic three-level continuous engine extended to a many-body working medium where the particles interact collectively with the baths. The formulation using tools from representation theory allowed us to study non-trivial symmetry types in the state going beyond the standard Bosonic or Fermionic exchange symmetries studied before [170, 174].

First, we derived the structure of the steady state. We used that the Hamiltonian is made up of only a linear combination of the diagonal generators of $\mathrm{SU}(3)$, and found that there is a convenient representation of the operators in the Schur basis simplifying the block diagonal structure found in [152]. This result was then used to study three ways that the work output of the collective system could be measured. The first was work that is extracted by unitary opera-

tions. This quantifies how much of a population inversion the system is able to produce, which corresponds to the maximum amount of work that the system could in theory transfer to a load. Here we found that a system that is prepared as close as possible to the fully symmetric state, will yield the highest amount of population inversion and therefore the most energy can be extracted. This was shown both numerically as well as analytically in the limit that $\omega_c\beta_c \gg \omega_h\beta_h$. We also saw that collectively coupled systems can have steady states with ergotropy outside of the range of parameters that would be expected from single-particles.

Next, the system was coupled to a dissipative load consisting of a zero temperature reservoir. This model is a simple theorist's analogy to "putting the wheels on the ground" [186] or in other words doing work against a dissipative force. In this example, we expected that some of the energy emitted into this bath will be in the form of heat and therefore the performance of the engine may differ from what would be found from population inversion alone. We found that although more energy was emitted when the coupling to the load was strong, the energy resembled heat because of its wide bandwidth and intensity fluctuations that were larger than was observed for weak coupling. Finally, we analysed the output power of a driven system. This model is the closest to an experimentally realisable lasing model. In this case all energy output can be considered to be work since the system is driven with a coherent field. In this case, the output energy was unambiguously be associated with power. Here we found that similarly to the other two cases, fully symmetric states generate the most power. For this model, we only considered weak coupling to ensure that the results are thermodynamically consistent.

There are a number of extensions to this work that would be immediate future directions. The mathematical framework presented here could be used to study higher dimensional systems. It would also be a natural next step to introduce interactions between particles of the working medium. This may allow the degeneracies in the weight spaces to be lifted which could lead to different thermodynamics than what was found here. Finally, one could extend these results to metrology platforms. For example, one could use the system to detect multiple unknown temperatures or coupling rates simultaneously.

Chapter 6

Summary and Outlook

This thesis theoretically explored two aspects of open quantum systems. Firstly, how Bayesian quantum parameter estimation can be used for the accurate estimation of environmental parameters and secondly, how an equilibrating quantum system can be used to perform thermodynamic tasks.

We began with an introduction to Bayesian methods. Here, we presented a general framework for performing Bayesian estimation in quantum systems starting from how to specify prior information in the case where the experimenter wants to make minimal prior assumptions. The methods that we discussed were based on specifying invariances in the problem either on the level of the physical parameter or on the likelihood. From this point, we discussed how optimal estimators and cost functions can be derived in particular, while taking into account the invariances of the problem. We also discussed various bounds on the error of an experiment from a Bayesian point of view as well as how the experiment can be optimised with respect to the probes and measurements made in a Bayesian framework. These techniques were then applied to various case studies. Here we found that when invariances in the problem are not properly taken into account it can lead to misleading estimates of the parameter. We also demonstrated that when the prior distribution is based only on invariances in the parameter, the strategy is the most uninformed and often leads to simpler expressions for the prior and for optimal global estimators derived from the prior. We showed that probe and measurement optimisation is also possible from a totally global and Bayesian perspective when this specification of the prior is used. These

results can be extended in future to multiparameter estimation as well as applied more extensively to adaptive estimation. The streamlined approach of Bayesian estimation where all figures of merit are determined by the prior distribution also provides a simplified approach so that the ambiguity of how to summarise the posterior distribution in situations with little data is avoided.

We primarily considered qubit thermometry as the test case for these techniques in both discrete and continuous measurement scenarios. This is because thermometry is the paradigmatic example of open quantum system parameter estimation. However, many of the results can be easily generalised to any parameter that depends on the interaction rate between the system and environment. For the discrete thermometry case study we used a global sensitivity measure, for both equilibrium and non-equilibrium probes, to find that partially thermalised probes are sensitive to a wider range of temperatures but this comes at the cost of their maximum achievable precision. This result inspired the study of thermometry using continuous measurements. Here, we derived an expression for the Fisher information of the probes which for long times, which allowed us to define the Bayesian Cramér-Rao bound as a benchmark for our estimation strategy. Here we also studied an asymptotically optimal adaptive strategy and a scenario with and without measurement noise, and a finite detector bandwidth. The thermometry case studies can be extended in several ways. We consider only probes with no quantum coherence. For discrete measurements this is known to be optimal. However, in the continuous case, it would be interesting to determine if including coherences in the description improves the performance of the thermometer. Additionally, the Bayesian techniques and continuous monitoring could be applied to multiple interacting probes which may allow quantum coherences to improve the measurement accuracy. Finally, different adaptive strategies could be explored that go beyond the asymptotically optimal one we consider here but that are not as computationally expensive as optimising the Bayesian error at every step.

Finally, we studied an example of how collective effects in an open quantum system can be harnessed to perform a thermodynamic task. Using representation theoretic tools we were able to go beyond a description of states that have purely Bosonic exchange symmetry and characterised the steady state thermodynamics of a collective system coupled to multiple thermal baths. Using the representation theoretic tools we were able to find analytical results in a limiting case but also more generally study the system numerically. This allowed us to determine that the maximum energy that is extractable by unitary operations occurs when the system is prepared

as close as possible to a fully symmetric state. We also found that work can be extracted from a collective system outside of the lasing regime of a single particle but that system parameters should be chosen carefully to ensure that a collective system outperforms individual systems. In addition we studied the power output when the system is coupled to a dissipative load as well as when it is coupled to a periodic driving field. The latter case corresponds to the paradigmatic case of a lasing engine. Here we compared the performance between symmetry types and also between individually coupled and collectively coupled working systems. A number of future directions can be explored. The method developed here could be adapted to explore a working medium with interactions between particles. This may allow degeneracies in the state space to be lifted which could lead to markedly different dynamics. The method is also easily extended to higher dimensional particles. Finally, applications could also be found in open system metrology. Here, the Bayesian methods developed in earlier chapters could be extended to multi parameter estimation and used to estimate multiple environmental parameters simultaneously since the collective system we consider here is coupled to two baths.

Appendix A

Modelling the interaction with the environment

In this thesis we are concerned with systems that interact with an environment. This means that in contrast to a closed system, the dynamics cannot be fully described by the Hamiltonian evolution of the system but rather the Hamiltonian includes the interaction with an environment. The following chapters we will be concerned with n-level systems with Hamiltonian H_S interacting with an environment that consists of a continuum of bosonic or fermionic modes described by H_B and a possibly time dependent interaction Hamiltonian $H_{SB}(t)$. The total Hamiltonian is then given by,

$$H(t) = H_S + H_B + H_{SB}(t). \tag{A.1}$$

We will assume that the environment is a thermal reservoir with Canonical distribution,

$$\rho_B = \frac{e^{-\beta H_B}}{\mathrm{tr}\{e^{-\beta H_B}\}}, \tag{A.2}$$

where, β is the inverse temperature of the environment. The state of the total system at time t is described by the density operator $\rho(t) = \sum_i p_i \left|\psi_i\right\rangle\!\left\langle\psi_i\right|$, where p_i is the probability that the total system is in the state$|\psi_i\rangle$. Practically speaking, the bath degrees of freedom are inaccessible and the observables of the system are what we will be interested in. Therefore, we need the

dynamics of the system density matrix. What follows is a sketch of the microscopic derivation of the standard Lindblad master equation that defines the evolution of the reduced density matrix $\rho_S(t) = \mathrm{tr}_B[\rho(t)]$. This is a standard textbook result, see for example, [67, 68].

First, we write down the state and Hamiltonian in the interaction picture,

$$\rho_I(t) = e^{i(H_S+H_B)t}\rho(t)e^{-i(H_S+H_B)t} \quad H_I(t) = e^{i(H_S+H_B)t}H(t)e^{-i(H_S+H_B)t}. \tag{A.3}$$

In this picture, the evolution of the system and environment is given by the von Neumann equation

$$\dot{\rho}_I(t) = -i[H_I(t), \rho_I(t)]. \tag{A.4}$$

This is integrated to give,

$$\rho(t) = \rho(0) - i\int_0^t ds[H_I(t), \rho(s)]. \tag{A.5}$$

This result is substituted back into Eq A.4 and we perform the partial trace over the environment degrees of freedom to get,

$$\dot{\rho}_S(t) = -i\,\mathrm{tr}_B[H_I(t), \rho(0)] - \int_0^t ds\,\mathrm{tr}_B[H_I(t), [H_I(t), \rho(s)]]. \tag{A.6}$$

In cases where the environment is a standard thermal reservoir and there is no external driving we can assume that $\mathrm{tr}_B[H_I(t), \rho(0)] = 0$.

Now we make the first approximation called the Born approximation which assumes that the coupling between the system and environment is weak and that the environment is much larger than the system so that the effect of $H_I(t)$ on the state of the bath is negligible. This means that if the system begins in an initial state uncorrelated product state, $\rho(0) = \rho_S(0) \otimes \rho_B(0)$ the total state at time t can be approximated to, $\rho(t) = \rho_S(t) \otimes \rho_B(0) + \mathcal{O}(H_I)$. Substituting this into Eq. A.6 yields,

$$\dot{\rho}_S(t) = -\int_0^t ds\,\mathrm{tr}_B[H_I(t), [H_I(t), \rho_S(s) \otimes \rho_B(0)]] + \mathcal{O}(H_I^3). \tag{A.7}$$

The next assumption we make is that the system has a Markovian evolution, that is, knowledge of the state $\rho(t)$ allows us to determine the state $\rho(t_1)$ for any time $t_1 > t$. We do this by setting $\rho(s) \to \rho(t)$ and letting the upper limit of the integral go to infinity in Eq. A.7. This is justified if the change in the system state happens on time scales τ_S which are much larger than the correlation time of the bath τ_B, in other words, on the time scales we consider the bath has no memory of its interaction with the system. This gives the Markovian Redfield equation,

$$\dot{\rho}(t) = -i\,\mathrm{tr}_B[H_I(t), \rho(0)] - \int_0^\infty ds\,\mathrm{tr}_B[H_I(t), [H_I(t), \rho_S(s) \otimes \rho_B(0)]]. \tag{A.8}$$

To further simplify this equation and to guarantee that the density matrix will be positive under the evolution by the master equation, we make an additional approximation known as the Secular approximation. This requires a few additional definitions. First define the operators,

$$A_\alpha(\omega) = \sum_{\varepsilon' - \varepsilon = \omega} \Pi(\varepsilon) A_\alpha \Pi(\varepsilon') \tag{A.9}$$

where ε are the eigenvalues of the system Hamiltonian and $\Pi(\varepsilon)$ is the projector onto the eigenvector of H_S with eigenvalue ε. The operators A_α act as raising and lowering operators of the system Hamiltonian,

$$[H_s, A_\alpha(\omega)] = \omega A_\alpha(\omega) \tag{A.10}$$

And we can write down H_I in the interaction picture as,

$$H_I(t) = \sum_{\alpha,\omega} A_\alpha(\omega) \otimes B_\alpha \tag{A.11}$$

Here we note that the condition $\mathrm{tr}_B[H_I(t), \rho(0)] \stackrel{!}{=} 0$ becomes $\mathrm{tr}(B_\alpha(t)\rho_B) = 0$ so we see that it implies that the bath average of the interaction Hamiltonian vanishes. Next, define the Fourier transform of the bath correlation functions,

$$\Gamma(\omega) = \int_0^\infty ds e^{i\omega s}\,\mathrm{tr}_B\{B^\dagger(t)B(t-s)\rho_B\} \tag{A.12}$$

which, if we assume that the bath is in a stationary state, simplify to the time independent expression

$$\Gamma(\omega) = \int_0^\infty ds e^{i\omega s} \operatorname{tr}_B\{B^\dagger(s)B(0)\rho_B\}. \tag{A.13}$$

These definitions allow us to write Eq. (A.8) in the following form,

$$\dot{\rho}(t) = \sum_{\omega,\omega'}\sum_{\alpha} e^{i(\omega'-\omega)t}\Gamma_\alpha(\omega)(A_\alpha(\omega)\rho_S(t)A_\alpha^\dagger(\omega') - A_\alpha^\dagger(\omega')A_\alpha(\omega)\rho_S(t)) + \text{h.c.} \tag{A.14}$$

We now make the secular approximation, which assumes that the terms where $\omega \neq \omega'$ oscillate much faster than the typical time scale of the system evolution and so they may be neglected in the summation.

$$\dot{\rho}(t) = \sum_{\alpha,\omega} \Gamma_\alpha(\omega)(A_\alpha(\omega)\rho_S(t)A_\alpha^\dagger(\omega) - A_\alpha^\dagger(\omega)A_\alpha(\omega)\rho_S(t)) + \text{h.c.} \tag{A.15}$$

We can simplify this further by writing $\Gamma_\alpha(\omega) = 1/2\gamma(\omega) + iS(\omega)$. The imaginary part of this expression is typically ignored as it contributes only to a small shift in the Hamiltonian called the Lamb shift. The real part can be written as,

$$\gamma(\omega) = \int_{-\infty}^\infty ds e^{i\omega s} \operatorname{tr}_B\{B^\dagger(s)B(0)\rho_B\} \tag{A.16}$$

which for a thermal reservoir are related by,

$$\gamma(-\omega) = e^{-\beta\omega}\gamma(\omega). \tag{A.17}$$

This results in the standard Lindblad master equation.

$$\dot{\rho}(t) = \sum_{\alpha,\omega} \gamma_\alpha(\omega)\left(A_\alpha(\omega)\rho_S(t)A_\alpha^\dagger(\omega) - \frac{1}{2}\{A_\alpha^\dagger(\omega)A_\alpha(\omega), \rho_S(t)\}\right). \tag{A.18}$$

A.1 Lindblad master equation for a two-level system

Several examples in this thesis consider the interaction of a thermal environment with a two level system with ground and excited states $|0\rangle$ and $|1\rangle$ and non-interacting Hamiltonian,

$$H_S = \frac{E}{2}(|1\rangle\langle 1| - |0\rangle\langle 0|). \tag{A.19}$$

In this case, the jump operators $A_\alpha(\omega)$ are simply, $|0\rangle\langle 1|$ and $|1\rangle\langle 0|$. Then we will consider two commonly used expressions for the coupling rates. First, a bath of resonantly coupled Bosonic modes with rate,

$$\gamma(\omega) = \bar{n}_B(\omega)J(\omega) \tag{A.20}$$

where $J(\omega)$ is the bath spectral density (we consider both Ohmic and flat spectral densities) and $\bar{n}_B(\omega)$, the average occupation of mode with frequency ω,

$$\bar{n}_B(\omega) = \frac{1}{e^{\beta\omega} - 1}. \tag{A.21}$$

Second, a Fermionic bath with flat (constant) spectral density,

$$\gamma(\omega) = \bar{n}_F(\omega) \tag{A.22}$$

where $\bar{n}_F(\omega)$ is the average occupation of mode with frequency ω,

$$\bar{n}_F(\omega) = \frac{1}{e^{\beta\omega} + 1}. \tag{A.23}$$

This yields the Master equation for a Bosonic reservoir,

$$\dot{\rho} = J(\omega)\bar{n}_B(E)\left(\mathcal{D}[|0\rangle\langle 1|]\rho + e^{-E/k_BT}\mathcal{D}[|1\rangle\langle 0|]\rho\right), \tag{A.24}$$

with $\mathcal{D}[\hat{A}]\rho = \hat{A}\rho\hat{A}^\dagger - \{\hat{A}^\dagger\hat{A}, \rho\}/2$. A similar expression is obtained for a Fermionic bath with the Bosonic rates replaced by Fermionic rates.

Appendix B

Optimal Bayesian estimator for a relative cost function

In this appendix, the optimal Bayesian estimator is derived for the cost function $c^{(2r)}(\vartheta_x, \theta) = (\vartheta_x/\theta - 1)$ introduced in Sec. 2.4.1. Other Bayesian estimators can be obtained in a similar way from other cost functions. See for example [64, 71, 101].

The optimal estimator will minimise the average loss,

$$\bar{\varepsilon}^{(2r)}(\vartheta_x) = \int dx P(x) \int d\theta P(\theta|x) \left(\frac{\vartheta_x}{\theta} - 1 \right), \tag{B.1}$$

This is done variationally, by finding the estimator such that

$$\frac{d}{dy} \bar{\varepsilon}^{(2r)}(\vartheta_x + y\eta_x)\Big|_{\eta_x=0} \stackrel{!}{=} 0 \tag{B.2}$$

If we evaluate the derivative on the LHS we obtain,

$$2 \int dx P(x) \int d\theta P(\theta|x) \eta_x \left(\frac{\vartheta_x}{\theta^2} - \frac{1}{\theta} \right), \tag{B.3}$$

which is zero whenever, $\int d\theta P(\theta|x) \left(\frac{\vartheta_x}{\theta^2} - \frac{1}{\theta} \right) \stackrel{!}{=} 0$.

This is equivalent to,

$$\vartheta_x \int d\theta P(\theta|x) \left(\frac{1}{\theta^2}\right) - \int d\theta P(\theta|x) \left(\frac{1}{\theta}\right) = 0. \tag{B.4}$$

Therefore, we solve for ϑ_x to recover the optimal relative estimator

$$\vartheta_x = \frac{\int d\theta P(\theta|x) \left(\frac{1}{\theta}\right)}{\int d\theta P(\theta|x) \left(\frac{1}{\theta^2}\right)}. \tag{B.5}$$

Appendix C

Van Trees inequality for an average relative error

In this appendix we will derive the van Trees bound for a relative root mean squared error which is applicable to estimation of a scale invariant parameter defined in the range $\theta \in (0, \infty)$. This derivation closely follows the derivation for the usual RMS error derived in [127].

Here we assume that we have a data set of measurement outcomes $\boldsymbol{x}$ and a likelihood function $P(\boldsymbol{x}|\theta)$ which is normalised, $\int \mathrm{d}\boldsymbol{x}\, P(\boldsymbol{x}|\theta) = 1$ for all values of θ. Furthermore, we assume that the Fisher information is continuous on the range of θ and that the prior distribution is normalised and decays on the end points of the range of θ, that is, $P(\theta) \xrightarrow{\theta\to 0} 0$ and $\theta P(\theta) \xrightarrow{\theta\to\infty} 0$.

Now define the scalar product: $\langle f, g\rangle = \int_0^\infty \mathrm{d}T \int \mathrm{d}\boldsymbol{x}\, f(\boldsymbol{x}, T) g(\boldsymbol{x}, T)$ for real-valued functions over the combined $(\boldsymbol{x}, \theta)$ space.

Then define the functions,

$$
\begin{aligned}
f(\boldsymbol{x}, \theta) &= \sqrt{P(\boldsymbol{x}|\theta) P(\theta)}\, \frac{\vartheta_{\boldsymbol{x}} - \theta}{\theta}, \\
g(\boldsymbol{x}, \theta) &= \theta \sqrt{P(\boldsymbol{x}|\theta) P^{(0)}(T)} \frac{\mathrm{d}}{\mathrm{d}T} \ln\left[P(\boldsymbol{x}|T) P^{(0)}(T)\right].
\end{aligned}
\tag{C.1}
$$

and assuming that both functions have finite norm, we can write the norms $\langle f, g\rangle$,$\langle f, f\rangle$ and $\langle g, g\rangle$,

$$
\langle f, g\rangle = \int_0^\infty \mathrm{d}\theta \int \mathrm{d}\boldsymbol{x}\, [\vartheta_{\boldsymbol{x}} - \theta] \frac{\mathrm{d}}{\mathrm{d}\theta} \left[P(\boldsymbol{x}|\theta) P(\theta)\right], \tag{C.2}
$$

$$\begin{aligned}\langle f, f\rangle &= \int_0^\infty \mathrm{d}\theta \int \mathrm{d}\boldsymbol{x} \left[\frac{\vartheta_{\boldsymbol{x}} - \theta}{\theta}\right]^2 P(\boldsymbol{x}|\theta)P(\theta) \\ &= \mathcal{C}^{(2r)}(\vartheta),\end{aligned} \tag{C.3}$$

and,

$$\begin{aligned}\langle g, g\rangle &= \int_0^\infty \mathrm{d}\theta \int \mathrm{d}\boldsymbol{x}\, P(\boldsymbol{x}|\theta)P(\theta)\theta^2 \left[\frac{\mathrm{d}}{\mathrm{d}\theta}\ln P(\boldsymbol{x}|\theta) + \frac{\mathrm{d}}{\mathrm{d}\theta}\ln P(\theta)\right]^2 \\ &= \int_0^\infty \mathrm{d}\theta\, P(\theta)\theta^2 \int \mathrm{d}\boldsymbol{x}\, P(\boldsymbol{x}|\theta) \left[\frac{\mathrm{d}\ln P(\boldsymbol{x}|\theta)}{\mathrm{d}\theta}\right]^2 + \int_0^\infty \mathrm{d}\theta\, P(\theta)\theta^2 \left[\frac{\mathrm{d}\ln P(\theta)}{\mathrm{d}\theta}\right]^2 \\ &\quad + 2\int_0^\infty \mathrm{d}\theta\, P(\theta)\theta^2 \frac{\mathrm{d}\ln P(\theta)}{\mathrm{d}\theta} \int \mathrm{d}\boldsymbol{x} \frac{\mathrm{d}}{\mathrm{d}\theta} P(\boldsymbol{x}|\theta) \\ &= \int_0^\infty \mathrm{d}\theta\, P(\theta)\theta^2 I(\theta) + I_0 + 0.\end{aligned} \tag{C.4}$$

In the last line we define $I_0 = \int_0^\infty \mathrm{d}\theta\, P(\theta)\theta^2 \left[\partial_\theta \ln P(\theta)\right]^2$ and by exchanging the θ derivative and the integral over $\boldsymbol{x}$ in the last term of the second last line we see that this term vanishes. We can relate these three norms with the Cauchy-Schwarz inequality, to get

$$\langle f, f\rangle = \mathcal{C}^{(2r)}(\vartheta) \leq \frac{\langle f, g\rangle}{\int_0^\infty \mathrm{d}\theta\, P(\theta)\theta^2 I(\theta) + I_0}. \tag{C.5}$$

Finally, we simplify the expression (C.2) for $\langle f, g\rangle$ by means of partial integration,

$$\langle f, g\rangle = \int \mathrm{d}\boldsymbol{x} \left\{ \left[P(\theta)P(\boldsymbol{x}|\theta)(\vartheta_{\boldsymbol{x}} - \theta)\right]_0^\infty + \int_0^\infty \mathrm{d}\theta\, P(\boldsymbol{x}|\theta)P(\theta) \right\} = 1, \tag{C.6}$$

where we have used the assumptions made about the behaviour of the prior at the boundary so that the first term vanishes. thus we achieve the desired bound,

$$\sqrt{\mathcal{C}^{(2r)}(\vartheta)} \geq \frac{1}{\sqrt{\int d\theta P(\theta)\theta^2 \mathcal{I}(\theta) + I_0}}. \tag{C.7}$$

Appendix D

$\mathrm{SU}(3)$ commutators

We give all commutation relations of the $\mathrm{SU}(3)$ generators for reference [189]:

$$3[W_z, W_\pm] = \pm W_\pm \qquad [W_+, W_-] = 2W_z \tag{D.1}$$

$$[W_z, U_\pm] = \mp 1/2 U_\pm \qquad [U_+, U_-] = \frac{3}{2}Y - W_z \tag{D.2}$$

$$[W_z, V_\pm] = \pm 1/2 V_\pm \qquad [V_+, V_-] = \frac{3}{2}Y + W_z \tag{D.3}$$

$$[Y, U_\pm] = \pm U_\pm \qquad [Y, V_\pm] = \pm V_\pm \tag{D.4}$$

$$[W_+, V_-] = -U_- \qquad [W_+, U_+] = V_+ \tag{D.5}$$

$$[U_+, V_-] = W_-, \tag{D.6}$$

with the remaining independent ones all vanishing.

The generators act on the basis in a $\lambda = (p, q)$ irrep in the following way [194]:

$$\begin{aligned}
W_z \left|\lambda, W, w, y\right\rangle &= t \left|\lambda, W, w, y\right\rangle \\
W_+ \left|\lambda, W, w, y\right\rangle &= \sqrt{(W-w)(W+w+1)} \left|\lambda, W, w+1, y\right\rangle \\
W_- \left|\lambda, W, w, y\right\rangle &= \sqrt{(W+w)(W-w+1)} \left|\lambda, W, w-1, y\right\rangle \\
Y \left|\lambda, W, w, y\right\rangle &= y \left|\lambda, W, w, y\right\rangle \\
V_+ \left|\lambda, W, w, y\right\rangle &= A^{p,q}_{W,w,y} \left|\lambda, W+1/2, w+1/2, y+1\right\rangle
\end{aligned} \tag{D.7}$$

$$
\begin{aligned}
&+ B^{p,q}_{W,w,y} \left|\lambda, W-1/2, w+1/2, y+1\right\rangle \\
V_- \left|\lambda, W, w, y\right\rangle =& A^{p,q}_{W-1/2,w-1/2,y-1} \left|\lambda, W+1/2, w-1/2, y-1\right\rangle \\
&+ B^{p,q}_{W+1/2,w-1/2,y-1} \left|\lambda, W-1/2, w-1/2, y-1\right\rangle \\
U_+ \left|\lambda, W, w, y\right\rangle =& \sqrt{(W-w+1)}\left(A^{p,q}_{W,w,y}\sqrt{(W+w+1)} - A^{p,q}_{W,w-1,y}\sqrt{(W+w)}\right) \\
&\left|\lambda, W+1/2, w-1/2, y+1\right\rangle \\
&+ \sqrt{(W+w)}\left(B^{p,q}_{W,w,y}\sqrt{(W-w)} - B^{p,q}_{W,w-1,y}\sqrt{(W-w+1)}\right) \\
&\left|\lambda, W-1/2, w-1/2, y+1\right\rangle \\
U_- \left|\lambda, W, w, y\right\rangle =& \left(A^{p,q}_{W-1/2,w-1/2,y-1}\sqrt{(W+w)} + A^{p,q}_{W-1/2,w+1/2,y-1}\sqrt{(W+w+1)}\right) \\
&\sqrt{(W-w)}\left|\lambda, W+1/2, w+1/2, y-1\right\rangle \\
&+\left(B^{p,q}_{W+1/2,w+1/2,y-1}\sqrt{(W-w)} - B^{p,q}_{W+1/2,w-1/2,y-1}\sqrt{(W-w+1)}\right) \\
&\sqrt{(W+w+1)}\left|\lambda, W-1/2, w+1/2, y-1\right\rangle
\end{aligned}
$$

Where the coefficients $A^{p,q}_{W,w,y}$ and $B^{p,q}_{W,w,y}$ are positive and given by,

$$
A^{p,q}_{W,w,y} = \left[\frac{(W+w+1)}{2(W+1)(2W+1)}\left(\frac{W(p-q)}{3}+\frac{Y}{2}+1\right)\left(\frac{p+2q}{3}+W+\frac{Y}{2}+2\right)\left(\frac{2p+q}{3}-W-\frac{Y}{2}\right)\right]^{1/2} \tag{D.8}
$$

$$
B^{p,q}_{W,w,y} = \left[\frac{(W-w)}{2W(2W+1)}\left(\frac{W(q-p)}{3}+W-\frac{Y}{2}\right)\left(\frac{p+2q}{3}-W+\frac{Y}{2}+1\right)\left(\frac{2p+q}{3}+W-\frac{Y}{2}+1\right)\right]^{1/2}. \tag{D.9}
$$

Bibliography

[1] E. R. Bittner and C. Silva, *Noise-induced quantum coherence drives photo-carrier generation dynamics at polymeric semiconductor heterojunctions*, *Nat Commun* **5** (Jan., 2014) 3119. Publisher: Nature Publishing Group.

[2] K. E. Dorfman, D. V. Voronine, S. Mukamel, and M. O. Scully, *Photosynthetic reaction center as a quantum heat engine*, *Proceedings of the National Academy of Sciences* **110** (Feb., 2013) 2746–2751. Publisher: Proceedings of the National Academy of Sciences.

[3] P. Nalbach and M. Thorwart, *Enhanced quantum efficiency of light-harvesting in a biomolecular quantum "steam engine"*, *Proceedings of the National Academy of Sciences* **110** (Feb., 2013) 2693–2694. Publisher: Proceedings of the National Academy of Sciences.

[4] R. Alicki, D. Gelbwaser-Klimovsky, and K. Szczygielski, *Solar cell as self-oscillating heat engine*, *J. Phys. A: Math. Theor.* **49** (Jan., 2016) 015002. arXiv:1501.00701 [cond-mat, physics:quant-ph].

[5] T. Keller, T. Fogarty, J. Li, and T. Busch, *Feshbach engine in the Thomas-Fermi regime*, *Phys. Rev. Research* **2** (Aug., 2020) 033335.

[6] S. Deffner, *Quantum refrigerators – the quantum thermodynamics of cooling Bose gases*, *Quantum Views* **3** (Aug., 2019) 20. Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften.

[7] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics: Emergence of Thermodynamic Behavior Within Composite Quantum Systems*, vol. 784 of *Lecture Notes in Physics*. Springer, Berlin, Heidelberg, 2009.

[8] F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, eds., *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*, vol. 195 of *Fundamental Theories of Physics*. Springer International Publishing, Cham, 2018.

[9] H. E. D. Scovil and E. O. Schulz-DuBois, *Three-Level Masers as Heat Engines*, *Phys. Rev. Lett.* **2** (Mar., 1959) 262–263.

[10] J. E. Geusic, E. O. Schulz-DuBios, and H. E. D. Scovil, *Quantum Equivalent of the Carnot Cycle*, *Phys. Rev.* **156** (Apr., 1967) 343–351. Publisher: American Physical Society.

[11] R. Alicki, *The quantum open system as a model of the heat engine*, *J. Phys. A: Math. Gen.* **12** (May, 1979) L103–L107.

[12] R. Kosloff, *A quantum mechanical open system as a model of a heat engine*, *The Journal of Chemical Physics* **80** (Feb., 1984) 1625–1631.

[13] R. Kosloff and A. Levy, *Quantum heat engines and refrigerators: Continuous devices*, *Annu. Rev. Phys. Chem.* **65** (Apr., 2014) 365–393. arXiv:1310.0683 [quant-ph].

[14] R. Uzdin, A. Levy, and R. Kosloff, *Quantum Heat Machines Equivalence, Work Extraction beyond Markovianity, and Strong Coupling via Heat Exchangers*, *Entropy* **18** (Apr., 2016) 124. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[15] A. Ghosh, W. Niedenzu, V. Mukherjee, and G. Kurizki, *Thermodynamic Principles and Implementations of Quantum Machines*, in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions* (F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, eds.), pp. 37–66. Springer International Publishing, Cham, 2018.

[16] M. T. Mitchison and P. P. Potts, *Physical Implementations of Quantum Absorption Refrigerators*, in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions* (F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, eds.), pp. 149–174. Springer International Publishing, Cham, 2018.

[17] L. M. Cangemi, C. Bhadra, and A. Levy, *Quantum Engines and Refrigerators*, Feb., 2023. arXiv:2302.00726 [quant-ph].

[18] R. Kosloff, *Quantum Thermodynamics: A Dynamical Viewpoint*, *Entropy* **15** (June, 2013) 2100–2128. Number: 6 Publisher: Multidisciplinary Digital Publishing Institute.

[19] M. Horodecki and J. Oppenheim, *Fundamental limitations for quantum and nanoscale thermodynamics*, *Nat Commun* **4** (June, 2013) 2059. Publisher: Nature Publishing Group.

[20] P. Strasberg and A. Winter, *First and Second Law of Quantum Thermodynamics: A Consistent Derivation Based on a Microscopic Definition of Entropy*, *PRX Quantum* **2** (Aug., 2021) 030202.

[21] M. Campisi, P. Hänggi, and P. Talkner, *Colloquium. Quantum Fluctuation Relations: Foundations and Applications*, *Rev. Mod. Phys.* **83** (July, 2011) 771–791. arXiv:1012.2268 [cond-mat, physics:quant-ph].

[22] T. Denzler and E. Lutz, *Efficiency fluctuations of a quantum heat engine*, *Phys. Rev. Research* **2** (Sept., 2020) 032062.

[23] V. Holubec and A. Ryabov, *Work and power fluctuations in a critical heat engine*, *Phys. Rev. E* **96** (Sept., 2017) 030102.

[24] G. Verley, M. Esposito, T. Willaert, and C. Van den Broeck, *The unlikely Carnot efficiency*, *Nat Commun* **5** (Sept., 2014) 4721. Publisher: Nature Publishing Group.

[25] G. Benenti, G. Casati, K. Saito, and R. S. Whitney, *Fundamental aspects of steady-state conversion of heat to work at the nanoscale*, *Physics Reports* **694** (June, 2017) 1–124.

[26] W. Niedenzu, M. Huber, and E. Boukobza, *Concepts of work in autonomous quantum heat engines*, *Quantum* **3** (Oct., 2019) 195. arXiv:1907.01353 [quant-ph].

[27] M. L. Bera, S. Julià-Farré, M. Lewenstein, and M. N. Bera, *Quantum heat engines with Carnot efficiency at maximum power*, *Phys. Rev. Res.* **4** (Feb., 2022) 013157. Publisher: American Physical Society.

[28] L. A. Correa, J. P. Palao, D. Alonso, and G. Adesso, *Quantum-enhanced absorption refrigerators*, *Sci Rep* **4** (Feb., 2014) 3949. Publisher: Nature Publishing Group.

[29] M. Campisi and R. Fazio, *The power of a critical heat engine*, *Nat Commun* **7** (June, 2016) 11895. Publisher: Nature Publishing Group.

[30] G. Piccitto, M. Campisi, and D. Rossini, *The Ising critical quantum Otto engine*, *New J. Phys.* **24** (Oct., 2022) 103023. Publisher: IOP Publishing.

[31] M. Kloc, K. Meier, K. Hadjikyriakos, and G. Schaller, *Superradiant Many-Qubit Absorption Refrigerator*, *Phys. Rev. Applied* **16** (Oct., 2021) 044061.

[32] R. Kosloff and Y. Rezek, *The Quantum Harmonic Otto Cycle*, *Entropy* **19** (Apr., 2017) 136. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[33] D. von Lindenfels, O. Gräb, C. Schmiegelow, V. Kaushal, J. Schulz, M. T. Mitchison, J. Goold, F. Schmidt-Kaler, and U. Poschinger, *Spin Heat Engine Coupled to a Harmonic-Oscillator Flywheel*, *Phys. Rev. Lett.* **123** (Aug., 2019) 080602. Publisher: American Physical Society.

[34] J. Klaers, S. Faelt, A. Imamoglu, and E. Togan, *Squeezed thermal reservoirs as a resource for a nano-mechanical engine beyond the Carnot limit*, *Phys. Rev. X* **7** (Sept., 2017) 031044. arXiv:1703.10024 [cond-mat].

[35] J. Klatzow, J. N. Becker, P. M. Ledingham, C. Weinzetl, K. T. Kaczmarek, D. J. Saunders, J. Nunn, I. A. Walmsley, R. Uzdin, and E. Poem, *Experimental Demonstration of Quantum Effects in the Operation of Microscopic Heat Engines*, *Phys. Rev. Lett.* **122** (Mar., 2019) 110601.

[36] J. Roßnagel, S. T. Dawkins, K. N. Tolazzi, O. Abah, E. Lutz, F. Schmidt-Kaler, and K. Singer, *A single-atom heat engine*, *Science* **352** (Apr., 2016) 325–329. arXiv:1510.03681 [cond-mat, physics:physics, physics:quant-ph].

[37] K. Ono, S. Shevchenko, T. Mori, S. Moriyama, and F. Nori, *Analog of a Quantum Heat Engine Using a Single-Spin Qubit*, *Phys. Rev. Lett.* **125** (Oct., 2020) 166802. Publisher: American Physical Society.

[38] J. P. Peterson, T. B. Batalhão, M. Herrera, A. M. Souza, R. S. Sarthour, I. S. Oliveira, and R. M. Serra, *Experimental Characterization of a Spin Quantum Heat Engine*, *Phys. Rev. Lett.* **123** (Dec., 2019) 240601.

[39] N. Lörch, C. Bruder, N. Brunner, and P. P. Hofer, *Optimal work extraction from quantum states by photo-assisted Cooper pair tunneling*, *Quantum Sci. Technol.* **3** (June, 2018) 035014. Publisher: IOP Publishing.

[40] R. Schnabel, N. Mavalvala, D. E. McClelland, and P. K. Lam, *Quantum metrology for gravitational wave astronomy*, *Nat Commun* **1** (Nov., 2010) 121. Publisher: Nature Publishing Group.

[41] M. Tse, *et.al., Quantum-Enhanced Advanced LIGO Detectors in the Era of Gravitational-Wave Astronomy*, *Phys. Rev. Lett.* **123** (Dec., 2019) 231107. Publisher: American Physical Society.

[42] I. Baumgart, J.-M. Cai, A. Retzker, M. Plenio, and C. Wunderlich, *Ultrasensitive Magnetometer using a Single Atom*, *Phys. Rev. Lett.* **116** (June, 2016) 240801. Publisher: American Physical Society.

[43] D. Le Sage, K. Arai, D. R. Glenn, S. J. DeVience, L. M. Pham, L. Rahn-Lee, M. D. Lukin, A. Yacoby, A. Komeili, and R. L. Walsworth, *Optical magnetic imaging of living cells*, *Nature* **496** (Apr., 2013) 486–489. Publisher: Nature Publishing Group.

[44] A. Bhattacharyya, A. Ghoshal, and U. Sen, *Enhancing precision of atomic clocks by tuning disorder in accessories*, *Phys. Rev. A* **110** (July, 2024) 012620. Publisher: American Physical Society.

[45] M. A. Norcia, M. N. Winchester, J. R. K. Cline, and J. K. Thompson, *Superradiance on the millihertz linewidth strontium clock transition*, *Science Advances* **2** (Oct., 2016) e1601231. Publisher: American Association for the Advancement of Science.

[46] P. Erker, M. T. Mitchison, R. Silva, M. P. Woods, N. Brunner, and M. Huber, *Autonomous Quantum Clocks: Does Thermodynamics Limit Our Ability to Measure Time?*, *Phys. Rev. X* **7** (Aug., 2017) 031022. Publisher: American Physical Society.

[47] M. Mehboudi, M. R. Jørgensen, S. Seah, J. B. Brask, J. Kołodyński, and M. Perarnau-Llobet, *Fundamental Limits in Bayesian Thermometry and Attainability via Adaptive Strategies*, *Phys. Rev. Lett.* **128** (Apr., 2022) 130502. Publisher: American Physical Society.

[48] P. P. Potts, J. B. Brask, and N. Brunner, *Fundamental limits on low-temperature quantum thermometry with finite resolution*, *Quantum* **3** (July, 2019) 161. Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften.

[49] L. A. Correa, M. Mehboudi, G. Adesso, and A. Sanpera, *Individual Quantum Probes for Optimal Thermometry*, *Phys. Rev. Lett.* **114** (June, 2015) 220405.

[50] A. De Pasquale and T. M. Stace, *Quantum Thermometry*, in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions* (F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, eds.), pp. 503–527. Springer International Publishing, Cham, 2018.

[51] J. Glatthard, K. V. Hovhannisyan, M. Perarnau-Llobet, L. A. Correa, and H. J. D. Miller, *Energy measurements remain thermometrically optimal beyond weak coupling*, *Quantum* **7** (Nov., 2023) 1190. arXiv:2302.03061 [cond-mat, physics:quant-ph].

[52] M. R. Jørgensen, P. P. Potts, M. G. A. Paris, and J. B. Brask, *Tight bound on finite-resolution quantum thermometry at low temperatures*, *Phys. Rev. Res.* **2** (Sept., 2020) 033394. Publisher: American Physical Society.

[53] I. Henao, K. V. Hovhannisyan, and R. Uzdin, *Thermometric machine for ultraprecise thermometry of low temperatures*, Oct., 2021. arXiv:2108.10469.

[54] V. Cavina, L. Mancino, A. De Pasquale, I. Gianani, M. Sbroscia, R. I. Booth, E. Roccia, R. Raimondi, V. Giovannetti, and M. Barbieri, *Bridging thermodynamics and metrology in nonequilibrium quantum thermometry*, *Phys. Rev. A* **98** (Nov., 2018) 050101. Publisher: American Physical Society.

[55] F. Albarelli, M. G. A. Paris, B. Vacchini, and A. Smirne, *Invasiveness of nonequilibrium pure-dephasing quantum thermometry*, *Phys. Rev. A* **108** (Dec., 2023) 062421. Publisher: American Physical Society.

[56] G. Planella, M. F. Cenni, A. Acín, and M. Mehboudi, *Bath-Induced Correlations Enhance Thermometry Precision at Low Temperatures*, *Phys. Rev. Lett.* **128** (Jan., 2022) 040502. Publisher: American Physical Society.

[57] T. H. Johnson, F. Cosco, M. T. Mitchison, D. Jaksch, and S. R. Clark, *Thermometry of ultracold atoms via nonequilibrium work distributions*, *Phys. Rev. A* **93** (May, 2016) 053619. Publisher: American Physical Society.

[58] M. Yu, H. C. Nguyen, and S. Nimmrichter, *Criticality-Enhanced Precision in Phase Thermometry*, Dec., 2023. arXiv:2311.14578 [quant-ph].

[59] P. Abiuso, P. A. Erdman, M. Ronen, F. Noé, G. Haack, and M. Perarnau-Llobet, *Optimal thermometers with spin networks*, *Quantum Sci. Technol.* **9** (Apr., 2024) 035008. Publisher: IOP Publishing.

[60] Y. Yang, V. Montenegro, and A. Bayat, *Sequential measurements thermometry with quantum many-body probes*, Mar., 2024. arXiv:2403.10084 [cond-mat, physics:quant-ph].

[61] I. Vybornyi, L. S. Dreissen, D. Kiesenhofer, H. Hainzer, M. Bock, T. Ollikainen, D. Vadlejch, C. F. Roos, T. E. Mehlstäubler, and K. Hammerer, *Sideband thermometry of ion crystals*, Sept., 2023. arXiv:2306.07880 [physics, physics:quant-ph].

[62] G. Mihailescu, S. Campbell, and A. K. Mitchell, *Thermometry of Strongly Correlated Fermionic Quantum Systems using Impurity Probes*, *Phys. Rev. A* **107** (Apr., 2023) 042614. arXiv:2212.09618 [cond-mat, physics:quant-ph].

[63] M. R. Jørgensen, J. Kołodyński, M. Mehboudi, M. Perarnau-Llobet, and J. B. Brask, *Bayesian quantum thermometry based on thermodynamic length*, *Phys. Rev. A* **105** (Apr., 2022) 042601. arXiv:2108.05901 [quant-ph].

[64] J. Rubio, J. Anders, and L. A. Correa, *Global Quantum Thermometry*, *Phys. Rev. Lett.* **127** (Nov., 2021) 190402. arXiv:2011.13018 [cond-mat, physics:physics, physics:quant-ph].

[65] G. O. Alves and G. T. Landi, *Bayesian estimation for collisional thermometry*, *Phys. Rev. A* **105** (Jan., 2022) 012212. Publisher: American Physical Society.

[66] J. Glatthard, J. Rubio, R. Sawant, T. Hewitt, G. Barontini, and L. A. Correa, *Optimal Cold Atom Thermometry Using Adaptive Bayesian Strategies*, *PRX Quantum* **3** (Dec., 2022) 040330.

[67] H. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*. Oxford University Press, 2002.

[68] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control*. Cambridge University Press, Nov., 2009. Google-Books-ID: 8H8hAwAAQBAJ.

[69] R. A. Fisher and E. J. Russell, *On the mathematical foundations of theoretical statistics*, *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **222** (Jan., 1997) 309–368. Publisher: Royal Society.

[70] R. W. Keener, *Theoretical Statistics: Topics for a Core Course*. Springer Texts in Statistics. Springer, New York, NY, 2010.

[71] W. v. d. Linden, V. Dose, and U. v. Toussaint, *Bayesian Probability Theory: Applications in the Physical Sciences*. Cambridge University Press, Cambridge, 2014.

[72] C. R. Rao, *Information and accuracy attainable in the estimation of statistical parameters*, *Bulletin of the Calcutta Mathematical Society* **37** (1945), no. 3 81–91.

[73] S. L. Braunstein and C. M. Caves, *Statistical distance and the geometry of quantum states*, *Phys. Rev. Lett.* **72** (May, 1994) 3439–3443.

[74] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum Metrology*, *Phys. Rev. Lett.* **96** (Jan., 2006) 010401. Publisher: American Physical Society.

[75] Z. Y. Ou, *Fundamental quantum limit in precision phase measurement*, *Phys. Rev. A* **55** (Apr., 1997) 2598–2609. Publisher: American Physical Society.

[76] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum-Enhanced Measurements: Beating the Standard Quantum Limit*, *Science* **306** (2004), no. 5700 1330–1336. _eprint: https://www.science.org/doi/pdf/10.1126/science.1104149.

[77] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, *Entanglement-free Heisenberg-limited phase estimation*, *Nature* **450** (Nov., 2007) 393–396. Publisher: Nature Publishing Group.

[78] J. J. . Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, *Optimal frequency measurements with maximally correlated states*, *Phys. Rev. A* **54** (Dec., 1996) R4649–R4652. Publisher: American Physical Society.

[79] U. Dorner, R. Demkowicz-Dobrzanski, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley, *Optimal Quantum Phase Estimation*, *Phys. Rev. Lett.* **102** (Jan., 2009) 040403. Publisher: American Physical Society.

[80] M. Valeri, E. Polino, D. Poderini, I. Gianani, G. Corrielli, A. Crespi, R. Osellame, N. Spagnolo, and F. Sciarrino, *Experimental adaptive Bayesian estimation of multiple phases with limited data*, *npj Quantum Inf* **6** (Dec., 2020) 1–11. Publisher: Nature Publishing Group.

[81] G. Y. Xiang, B. L. Higgins, D. W. Berry, H. M. Wiseman, and G. J. Pryde, *Entanglement-enhanced measurement of a completely unknown optical phase*, *Nature Photon* **5** (Jan., 2011) 43–47. Publisher: Nature Publishing Group.

[82] N. Wiebe and C. E. Granade, *Efficient Bayesian Phase Estimation*, *Phys. Rev. Lett.* **117** (June, 2016) 010503. arXiv:1508.00869 [quant-ph].

[83] M. A. Rodríguez-García and F. E. Becerra, *Adaptive Phase Estimation with Squeezed Vacuum Approaching the Quantum Limit*, *Quantum* **8** (Sept., 2024) 1480. arXiv:2312.07686 [quant-ph].

[84] F. Martínez-García, D. Vodola, and M. Müller, *Adaptive Bayesian phase estimation for quantum error correcting codes*, *New J. Phys.* **21** (Dec., 2019) 123027. Publisher: IOP Publishing.

[85] E. Martínez-Vargas, C. Pineda, F. Leyvraz, and P. Barberis-Blostein, *Quantum estimation of unknown parameters*, *Phys. Rev. A* **95** (Jan., 2017) 012136. Publisher: American Physical Society.

[86] A. A. Berni, T. Gehring, B. M. Nielsen, V. Händchen, M. G. A. Paris, and U. L. Andersen, *Ab initio quantum-enhanced optical phase estimation using real-time feedback control*, *Nature Photon* **9** (Sept., 2015) 577–581. Publisher: Nature Publishing Group.

[87] C. L. Degen, F. Reinhard, and P. Cappellaro, *Quantum sensing*, *Rev. Mod. Phys.* **89** (July, 2017) 035002. arXiv:1611.02427 [quant-ph].

[88] J. Amorós-Binefa and J. Kołodyński, *Noisy atomic magnetometry in real time*, *New J. Phys.* **23** (Dec., 2021) 123030.

[89] M. Mehboudi, A. Sanpera, and L. A. Correa, *Thermometry in the quantum regime: recent theoretical progress*, *J. Phys. A: Math. Theor.* **52** (July, 2019) 303001. Publisher: IOP Publishing.

[90] M. G. A. Paris, *Achieving the Landau bound to precision of quantum thermometry in systems with vanishing gap*, *J. Phys. A: Math. Theor.* **49** (Dec., 2015) 03LT02. Publisher: IOP Publishing.

[91] R. R. Rodríguez, M. Mehboudi, M. Horodecki, and M. Perarnau-Llobet, *Strongly coupled fermionic probe for nonequilibrium thermometry*, *New J. Phys.* **26** (Jan., 2024) 013046. Publisher: IOP Publishing.

[92] L. Mancino, M. Sbroscia, I. Gianani, E. Roccia, and M. Barbieri, *Quantum Simulation of Single-Qubit Thermometry Using Linear Optics*, *Phys. Rev. Lett.* **118** (Mar., 2017) 130502. Publisher: American Physical Society.

[93] A. H. Kiilerich, A. De Pasquale, and V. Giovannetti, *Dynamical approach to ancilla-assisted quantum thermometry*, *Phys. Rev. A* **98** (Oct., 2018) 042124. Publisher: American Physical Society.

[94] S. Seah, S. Nimmrichter, D. Grimmer, J. P. Santos, V. Scarani, and G. T. Landi, *Collisional Quantum Thermometry*, *Phys. Rev. Lett.* **123** (Oct., 2019) 180602. Publisher: American Physical Society.

[95] A. Shu, S. Seah, and V. Scarani, *Surpassing the thermal Cram\'er-Rao bound with collisional thermometry*, *Phys. Rev. A* **102** (Oct., 2020) 042417. Publisher: American Physical Society.

[96] W.-K. Mok, K. Bharti, L.-C. Kwek, and A. Bayat, *Optimal probes for global quantum thermometry*, *Commun Phys* **4** (Mar., 2021) 1–8. Publisher: Nature Publishing Group.

[97] F. S. Luiz, A. d. O. Junior, F. F. Fanchini, and G. T. Landi, *Machine classification for probe-based quantum thermometry*, *Phys. Rev. A* **105** (Feb., 2022) 022413. Publisher: American Physical Society.

[98] M. Mehboudi, F. Meier, M. Huber, and H. J. D. Miller, *Optimal limits of continuously monitored thermometers and their Hamiltonian structure*, Aug., 2024. arXiv:2408.01313 [quant-ph].

[99] E. T. Jaynes, *Information Theory and Statistical Mechanics*, *Phys. Rev.* **106** (May, 1957) 620–630.

[100] A. W. v. d. Vaart, *Chapter 10 Bayes Procedures*, in *Asymptotic Statistics*. Cambridge University Press, June, 2000. Google-Books-ID: SYlmEAAAQBAJ.

[101] E. L. Lehmann and G. Casella, *Theory of Point Estimation*. Springer Science & Business Media, May, 2006. Google-Books-ID: 4f24CgAAQBAJ.

[102] T. Seidenfeld, *Why I am not an objective Bayesian; some reflections prompted by Rosenkrantz*, *Theor Decis* **11** (Dec., 1979) 413–440.

[103] R. E. Kass and L. Wasserman, *The Selection of Prior Distributions by Formal Rules*, *Journal of the American Statistical Association* **91** (Sept., 1996) 1343–1370. Publisher: ASA Website _eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1996.10477003.

[104] J. Rubio, *First-principles construction of symmetry-informed quantum metrologies*, Mar., 2024. arXiv:2402.16410 [math-ph, physics:physics, physics:quant-ph].

[105] J. Rubio, *Quantum scale estimation*, *Quantum Sci. Technol.* **8** (Jan., 2023) 015009. arXiv:2111.11921 [math-ph, physics:physics, physics:quant-ph].

[106] E. T. Jaynes, *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge, 2003.

[107] E. Jaynes, *Prior Probabilities*, *IEEE Trans. Syst. Sci. Cyber.* **4** (1968), no. 3 227–241.

[108] R. Demkowicz-Dobrzański, *Optimal phase estimation with arbitrary a priori knowledge*, *Phys. Rev. A* **83** (June, 2011) 061802.

[109] H. L. Harney, *Bayesian Inference*. Springer, Berlin, Heidelberg, 2003.

[110] O.-A. Al-Hujaj and H. L. Harney, *Objective Bayesian Statistics*, June, 1997. arXiv:physics/9706025.

[111] M. L. Eaton and W. D. Sudderth, *Properties of Right Haar Predictive Inference*, .

[112] H. Snoussi, *Bayesian Information Geometry: Application to Prior Selection on Statistical Manifolds*, in *Advances in Imaging and Electron Physics*, vol. 146, pp. 163–207. Elsevier, 2007.

[113] C. C. Rodriguez, *Are We Cruising a Hypothesis Space?*, Aug., 1998. arXiv:physics/9808009.

[114] C. W. Helstrom, *Quantum detection and estimation theory*, *J Stat Phys* **1** (June, 1969) 231–252.

[115] A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*. Edizioni della Normale, Pisa, 2011.

[116] M. G. A. Paris, *Quantum estimation for quantum technology*, *Int. J. Quantum Inform.* **07** (Jan., 2009) 125–137. Publisher: World Scientific Publishing Co.

[117] C. E. Shannon, *A mathematical theory of communication*, *The Bell System Technical Journal* **27** (July, 1948) 379–423. Conference Name: The Bell System Technical Journal.

[118] H. Snoussi and A. Mohammad-Djafari, *Information geometry and prior selection*, . Conference Name: Bayesian Inference and Maximum Entropy Methods in Science and Engineering Publisher: AIP ADS Bibcode: 2003AIPC..659..307S.

[119] F. Nielsen, *The Many Faces of Information Geometry*, *Notices Amer. Math. Soc.* **69** (Jan., 2022) 1.

[120] M. L. Eaton and W. D. Sudderth, *Invariance of posterior distributions under reparametrization*, *Sankhya* **72** (Feb., 2010) 101–118.

[121] E. T. Jaynes, *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge, 2003.

[122] S. Barnett, *Quantum Information*. OUP Oxford, May, 2009. Google-Books-ID: A2k4HH2tFR8C.

[123] P. G. Hoel, *Introduction to Mathematical Statistics*. Wiley, 1954. Google-Books-ID: _xPO875cT4gC.

[124] S. Personick, *Application of quantum estimation theory to analog communication over quantum channels*, *IEEE Transactions on Information Theory* **17** (May, 1971) 240–246. Conference Name: IEEE Transactions on Information Theory.

[125] L. Bacharach, C. Fritsche, U. Orguner, and E. Chaumette, *A Tighter Bayesian CramÉR-rao Bound*, in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5277–5281, May, 2019. ISSN: 2379-190X.

[126] Y. Li, L. Pezzè, M. Gessner, W. Li, and A. Smerzi, *Frequentist and Bayesian Quantum Phase Estimation*, *Entropy* **20** (Aug., 2018) 628. arXiv:1804.10048 [quant-ph].

[127] R. D. Gill and B. Y. Levit, *Applications of the van Trees Inequality: A Bayesian Cramér-Rao Bound*, *Bernoulli* **1** (1995), no. 1/2 59–79. Publisher: International Statistical Institute (ISI) and Bernoulli Society for Mathematical Statistics and Probability.

[128] R. H. Bartels and G. W. Stewart, *Algorithm 432 [C2]: Solution of the matrix equation AX + XB = C [F4]*, *Commun. ACM* **15** (Sept., 1972) 820–826.

[129] C. Sayrin, I. Dotsenko, X. Zhou, B. Peaudecerf, T. Rybarczyk, S. Gleyzes, P. Rouchon, M. Mirrahimi, H. Amini, M. Brune, and others, *Real-time quantum feedback prepares and stabilizes photon number states*, *Nature* **477** (2011), no. 7362 73–77. Publisher: Nature Publishing Group UK London.

[130] R. Vijay, C. Macklin, D. Slichter, S. Weber, K. Murch, R. Naik, A. N. Korotkov, and I. Siddiqi, *Stabilizing Rabi oscillations in a superconducting qubit using quantum feedback*, *Nature* **490** (2012), no. 7418 77–80. Publisher: Nature Publishing Group UK London.

[131] V. Shah, G. Vasilakis, and M. V. Romalis, *High Bandwidth Atomic Magnetometery with Continuous Quantum Nondemolition Measurements*, *Phys. Rev. Lett.* **104** (Jan., 2010) 013601. Publisher: American Physical Society.

[132] R. Jiménez-Martínez, J. Kołodyıfmmode \acuten\else ń\fiski, C. Troullinou, V. G. Lucivero, J. Kong, and M. W. Mitchell, *Signal Tracking Beyond the Time Resolution of an Atomic Sensor by Kalman Filtering*, *Phys. Rev. Lett.* **120** (Jan., 2018) 040503. Publisher: American Physical Society.

[133] M. Swar, D. Roy, D. D, S. Chaudhuri, S. Roy, and H. Ramachandran, *Measurements of spin properties of atomic systems in and out of equilibrium via noise spectroscopy*, *Opt. Express, OE* **26** (Nov., 2018) 32168–32183. Publisher: Optica Publishing Group.

[134] J. K. Stockton, J. M. Geremia, A. C. Doherty, and H. Mabuchi, *Robust quantum parameter estimation: Coherent magnetometry with feedback*, *Phys. Rev. A* **69** (Mar., 2004) 032109. Publisher: American Physical Society.

[135] J. Geremia, J. K. Stockton, A. C. Doherty, and H. Mabuchi, *Quantum Kalman Filtering and the Heisenberg Limit in Atomic Magnetometry*, *Phys. Rev. Lett.* **91** (Dec., 2003) 250801. Publisher: American Physical Society.

[136] S. Gammelmark and K. Mølmer, *Bayesian parameter inference from continuously monitored quantum systems*, *Phys. Rev. A* **87** (Mar., 2013) 032115. Publisher: American Physical Society.

[137] F. Albarelli, M. A. C. Rossi, D. Tamascelli, and M. G. Genoni, *Restoring Heisenberg scaling in noisy quantum metrology by monitoring the environment*, *Quantum* **2** (Dec., 2018) 110. Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften.

[138] A. H. Kiilerich and K. Mølmer, *Bayesian parameter estimation by continuous homodyne detection*, *Phys. Rev. A* **94** (Sept., 2016) 032103. Publisher: American Physical Society.

[139] K. Jacobs, *Continuous measurement*, in *Quantum Measurement Theory and its Applications*, pp. 90–159. Cambridge University Press, 2014.

[140] C. Benedetti, F. Salari Sehdaran, M. H. Zandi, and M. G. A. Paris, *Quantum probes for the cutoff frequency of Ohmic environments*, *Phys. Rev. A* **97** (Jan., 2018) 012126. Publisher: American Physical Society.

[141] C. Arnold, V. Loo, A. Lemaître, I. Sagnes, O. Krebs, P. Voisin, P. Senellart, and L. Lanco, *Cavity-Enhanced Real-Time Monitoring of Single-Charge Jumps at the Microsecond Time Scale*, *Phys. Rev. X* **4** (Apr., 2014) 021004. Publisher: American Physical Society.

[142] J. Borwanker, G. Kallianpur, and B. P. Rao, *The Bernstein-von Mises theorem for Markov processes*, *The Annals of Mathematical Statistics* (1971) 1241–1253. Publisher: JSTOR.

[143] D. Barker, M. Scandi, S. Lehmann, C. Thelander, K. A. Dick, M. Perarnau-Llobet, and V. F. Maisi, *Experimental Verification of the Work Fluctuation-Dissipation Relation for Information-to-Work Conversion*, *Phys. Rev. Lett.* **128** (Jan., 2022) 040602. Publisher: American Physical Society.

[144] B. Annby-Andersson, F. Bakhshinezhad, D. Bhattacharyya, G. De Sousa, C. Jarzynski, P. Samuelsson, and P. P. Potts, *Quantum Fokker-Planck Master Equation for Continuous Feedback Control*, *Phys. Rev. Lett.* **129** (July, 2022) 050401. Publisher: American Physical Society.

[145] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin, Heidelberg, 1992.

[146] D. Walls and G. J. Milburn, *Stochastic Methods*, in *Quantum Optics* (D. Walls and G. J. Milburn, eds.), pp. 93–126. Springer, Berlin, Heidelberg, 2008.

[147] A. Kutschireiter, S. C. Surace, and J.-P. Pfister, *The Hitchhiker's guide to nonlinear filtering*, *Journal of Mathematical Psychology* **94** (Feb., 2020) 102307.

[148] *5 Introduction to Filtering Theory*, in *Stochastic Processes and Filtering Theory* (A. H. Jazwinski, ed.), vol. 64 of *Mathematics in Science and Engineering*, pp. 142–161. Elsevier, 1970. ISSN: 0076-5392.

[149] M. Brenes and D. Segal, *Multispin probes for thermometry in the strong-coupling regime*, *Phys. Rev. A* **108** (Sept., 2023) 032220.

[150] E. Aybar, A. Niezgoda, S. S. Mirkhalaf, M. W. Mitchell, D. B. Orenes, and E. Witkowska, *Critical quantum thermometry and its feasibility in spin systems*, *Quantum* **6** (Sept., 2022) 808. Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften.

[151] C. L. Latune, I. Sinayskiy, and F. Petruccione, *Collective heat capacity for quantum thermometry and quantum engine enhancements*, *New J. Phys.* **22** (Aug., 2020) 083049. Publisher: IOP Publishing.

[152] B. Yadin, B. Morris, and K. Brandner, *Thermodynamics of Permutation-Invariant Quantum Many-Body Systems: A Group-Theoretical Framework*, *Phys. Rev. Research* **5** (July, 2023) 033018. arXiv:2206.12639 [cond-mat, physics:quant-ph].

[153] A. M. L. Messiah and O. W. Greenberg, *Symmetrization Postulate and Its Experimental Foundation*, *Phys. Rev.* **136** (Oct., 1964) B248–B267. Publisher: American Physical Society.

[154] M. C. Tichy and K. Mølmer, *Extending exchange symmetry beyond bosons and fermions*, *Phys. Rev. A* **96** (Aug., 2017) 022119. Publisher: American Physical Society.

[155] J. Jaramillo, M. Beau, and A. D. Campo, *Quantum supremacy of many-particle thermal machines*, *New J. Phys.* **18** (July, 2016) 075019.

[156] F. Carollo, K. Brandner, and I. Lesanovsky, *Nonequilibrium Many-Body Quantum Engine Driven by Time-Translation Symmetry Breaking*, *Phys. Rev. Lett.* **125** (Dec., 2020) 240602.

[157] L. A. Williamson and M. J. Davis, *Many-body enhancement in a spin-chain quantum heat engine*, *Phys. Rev. B* **109** (Jan., 2024) 024310.

[158] J. Li, T. Fogarty, S. Campbell, X. Chen, and T. Busch, *An efficient nonlinear Feshbach engine*, *New J. Phys.* **20** (Jan., 2018) 015005.

[159] Y.-Y. Chen, G. Watanabe, Y.-C. Yu, X.-W. Guan, and A. del Campo, *An interaction-driven many-particle quantum heat engine and its universal behavior*, *npj Quantum Inf* **5** (Oct., 2019) 1–6. Publisher: Nature Publishing Group.

[160] L. Fusco, M. Paternostro, and G. De Chiara, *Work extraction and energy storage in the Dicke model*, *Phys. Rev. E* **94** (Nov., 2016) 052122. Publisher: American Physical Society.

[161] T. Fogarty and T. Busch, *A many-body heat engine at criticality*, *Quantum Sci. Technol.* **6** (Nov., 2020) 015003. Publisher: IOP Publishing.

[162] N. Yunger Halpern, C. D. White, S. Gopalakrishnan, and G. Refael, *Quantum engine based on many-body localization*, *Phys. Rev. B* **99** (Jan., 2019) 024203.

[163] M. Gross and S. Haroche, *Superradiance: An essay on the theory of collective spontaneous emission*, *Physics Reports* **93** (Dec., 1982) 301–396.

[164] W. Niedenzu and G. Kurizki, *Cooperative many-body enhancement of quantum thermal machine power*, *New J. Phys.* **20** (Nov., 2018) 113038.

[165] V. Mukherjee and U. Divakaran, *Many-body quantum thermal machines*, *J. Phys.: Condens. Matter* **33** (Nov., 2021) 454001.

[166] L. D. S. Souza, G. Manzano, R. Fazio, and F. Iemini, *Collective effects on the performance and stability of quantum heat engines*, *Phys. Rev. E* **106** (July, 2022) 014143.

[167] S. Kamimura, H. Hakoshima, Y. Matsuzaki, K. Yoshida, and Y. Tokura, *Quantum-Enhanced Heat Engine Based on Superabsorption*, *Phys. Rev. Lett.* **128** (May, 2022) 180602.

[168] F. Mayo and A. J. Roncaglia, *Collective effects and quantum coherence in dissipative charging of quantum batteries*, *Phys. Rev. A* **105** (June, 2022) 062203.

[169] J. Kim, S.-h. Oh, D. Yang, J. Kim, M. Lee, and K. An, *A photonic quantum engine driven by superradiance*, *Nat. Photon.* **16** (Oct., 2022) 707–711. Publisher: Nature Publishing Group.

[170] C. L. Latune, I. Sinayskiy, and F. Petruccione, *Thermodynamics from indistinguishability: Mitigating and amplifying the effects of the bath*, *Phys. Rev. Res.* **1** (Dec., 2019) 033192. Publisher: American Physical Society.

[171] N. M. Myers and S. Deffner, *Bosons outperform fermions: The thermodynamic advantage of symmetry*, *Phys. Rev. E* **101** (Jan., 2020) 012110.

[172] N. M. Myers, F. J. Peña, O. Negrete, P. Vargas, G. D. Chiara, and S. Deffner, *Boosting engine performance with Bose–Einstein condensation*, *New J. Phys.* **24** (Feb., 2022) 025001. Publisher: IOP Publishing.

[173] M. A. Macovei, *Performance of the collective three-level quantum thermal engine*, *Phys. Rev. A* **105** (Apr., 2022) 043708.

[174] D. Kolisnyk and G. Schaller, *Performance boost of a collective qutrit refrigerator*, *Phys. Rev. Applied* **19** (Mar., 2023) 034023. arXiv:2210.07844 [cond-mat, physics:quant-ph].

[175] D. Kolisnyk, F. Queisser, G. Schaller, and R. Schützhold, *Floquet analysis of a superradiant many-qutrit refrigerator*, Apr., 2024. arXiv:2310.18126 [cond-mat, physics:physics, physics:quant-ph].

[176] H. T. Quan, Y.-x. Liu, C. P. Sun, and F. Nori, *Quantum thermodynamic cycles and quantum heat engines*, *Phys. Rev. E* **76** (Sept., 2007) 031105.

[177] S. Deffner, *Efficiency of harmonic quantum Otto engines at maximal power*, *Entropy* **20** (Nov., 2018) 875. arXiv:1810.09371 [cond-mat, physics:quant-ph].

[178] M. T. Mitchison, *Quantum thermal absorption machines: refrigerators, engines and clocks*, *Contemporary Physics* **60** (Apr., 2019) 164–187. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00107514.2019.1631555.

[179] V. Singh, *Optimal operation of a three-level quantum heat engine and universal nature of efficiency*, *Phys. Rev. Research* **2** (Nov., 2020) 043187.

[180] A. E. Allahverdyan, R. Balian, and T. M. Nieuwenhuizen, *Maximal work extraction from finite quantum systems*, *Europhys. Lett.* **67** (Aug., 2004) 565–571.

[181] O. Astafiev, A. A. Abdumalikov Jr., A. M. Zagoskin, Y. A. Pashkin, Y. Nakamura, and J. S. Tsai, *Ultimate on-chip quantum amplifier*, *Phys. Rev. Lett.* **104** (May, 2010) 183603. arXiv:1004.3864 [cond-mat, physics:quant-ph].

[182] A. A. S. Kalaee and A. Wacker, *Positivity of entropy production for the three-level maser*, *Phys. Rev. A* **103** (Jan., 2021) 012202.

[183] S.-W. Li, M. B. Kim, G. S. Agarwal, and M. O. Scully, *Quantum statistics of a single-atom Scovil-Schulz-DuBois heat engine*, *Phys. Rev. A* **96** (Dec., 2017) 063806. arXiv:1710.00902 [quant-ph].

[184] D. Gelbwaser-Klimovsky, R. Alicki, and G. Kurizki, *Work and energy gain of heat-pumped quantized amplifiers*, *EPL* **103** (Oct., 2013) 60005. Publisher: EDP Sciences, IOP Publishing and Società Italiana di Fisica.

[185] D. Gelbwaser-Klimovsky, W. Niedenzu, P. Brumer, and G. Kurizki, *Power enhancement of heat engines via correlated thermalization in a three-level "working fluid"*, *Sci Rep* **5** (Sept., 2015) 14413. Publisher: Nature Publishing Group.

[186] S. Seah, S. Nimmrichter, and V. Scarani, *Work production of quantum rotor engines*, *New J. Phys.* **20** (Apr., 2018) 043045.

[187] E. Geva and R. Kosloff, *Three-level quantum amplifier as a heat engine: A study in finite-time thermodynamics*, *Phys. Rev. E* **49** (May, 1994) 3903–3918.

[188] E. Boukobza and D. J. Tannor, *Thermodynamics of bipartite systems: Application to light-matter interactions*, *Phys. Rev. A* **74** (Dec., 2006) 063823.

[189] W. Greiner and B. Müller, *Quantum Mechanics*. Springer, Berlin, Heidelberg, 1994.

[190] H. Georgi, *Lie Algebras In Particle Physics: from Isospin To Unified Theories*. CRC Press, Boca Raton, June, 2019.

[191] R. Goodman and N. R. Wallach, *Symmetry, Representations, and Invariants*, vol. 255 of *Graduate Texts in Mathematics*. Springer, New York, NY, 2009.

[192] A. W. Harrow, *Applications of coherent classical communication and the Schur transform to quantum information theory*, Dec., 2005. arXiv:quant-ph/0512255.

[193] C. ITZYKSON and M. NAUENBERG, *Unitary Groups: Representations and Decompositions*, *Rev. Mod. Phys.* **38** (Jan., 1966) 95–120. Publisher: American Physical Society.

[194] J. J. de Swart, *The Octet Model and its Clebsch-Gordan Coefficients*, *Rev. Mod. Phys.* **35** (Oct., 1963) 916–939.

[195] L. C. Biedenharn, *The generalisation of the Wigner-Racah angular momentum calculus*, *Physics Letters* **3** (Jan., 1963) 254–256.

[196] L. C. Biedenharn, *INVARIANT OPERATORS OF THE CASIMIR TYPE*, *Phys. Letters* **Vol: 3** (Dec., 1962). Institution: Duke Univ., Durham, N.C.

[197] M. Markiewicz and J. Przewocki, *Duality of averaging of quantum states over arbitrary symmetry groups revealing Schur-Weyl duality*, Nov., 2022. arXiv:2208.07689 [quant-ph].

[198] I. Marvian and R. W. Spekkens, *A Generalization of Schur–Weyl Duality with Applications in Quantum Estimation*, *Commun. Math. Phys.* **331** (Oct., 2014) 431–475.

[199] V. Cavina, A. Soret, T. Aslyamov, K. Ptaszyński, and M. Esposito, *Symmetry shapes thermodynamics of macroscopic quantum systems*, *Phys. Rev. Lett.* **133** (Sept., 2024) 130401. arXiv:2402.04214 [cond-mat, physics:quant-ph].

[200] R. Alicki and M. Fannes, *Entanglement boost for extractable work from ensembles of quantum batteries*, *Phys. Rev. E* **87** (Apr., 2013) 042123. Publisher: American Physical Society.

[201] R. Alicki, S. Rudnicki, and S. Sadowski, *Symmetry properties of product states for the system of N n-level atoms*, *Journal of Mathematical Physics* **29** (May, 1988) 1158–1162.

[202] M. Keyl and R. F. Werner, *Estimating the spectrum of a density operator*, *Phys. Rev. A* **64** (Oct., 2001) 052311.

[203] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, *Julia: A fresh approach to numerical computing*, *SIAM review* **59** (2017), no. 1 65–98.

[204] U. Thiel, M. Horn, and T. Schmidt, "JuLie.jl." `https://github.com/ulthiel/JuLie.jl`, 2023.

[205] A. Lenard, *Thermodynamical proof of the Gibbs formula for elementary quantum systems*, *J Stat Phys* **19** (Dec., 1978) 575–586.

[206] E. Bäumer, M. Perarnau-Llobet, P. Kammerlander, H. Wilming, and R. Renner, *Imperfect Thermalizations Allow for Optimal Thermodynamic Processes*, *Quantum* **3** (June, 2019) 153. Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften.

[207] A. Levy, L. Diósi, and R. Kosloff, *Quantum flywheel*, *Phys. Rev. A* **93** (May, 2016) 052119. Publisher: American Physical Society.

[208] F. Campaioli, S. Gherardini, J. Q. Quach, M. Polini, and G. M. Andolina, *Colloquium: Quantum batteries*, *Rev. Mod. Phys.* **96** (July, 2024) 031001. Publisher: American Physical Society.

[209] W. Niedenzu, V. Mukherjee, A. Ghosh, A. G. Kofman, and G. Kurizki, *Quantum engine efficiency bound beyond the second law of thermodynamics*, *Nat Commun* **9** (Jan., 2018) 165. Publisher: Nature Publishing Group.

[210] A. Mari, A. Farace, and V. Giovannetti, *Quantum optomechanical piston engines powered by heat*, *J. Phys. B: At. Mol. Opt. Phys.* **48** (July, 2015) 175501. Publisher: IOP Publishing.

[211] A. Seegebrecht and T. Schilling, *Work, Heat and Internal Energy in Open Quantum Systems: A Comparison of Four Approaches from the Autonomous System Framework*, Aug., 2023. arXiv:2308.08215 [quant-ph].

[212] J. Monsel, M. Fellous-Asiani, B. Huard, and A. Auffèves, *The energetic cost of work extraction*, *Phys. Rev. Lett.* **124** (Mar., 2020) 130601. arXiv:1907.00812 [quant-ph].

[213] C. Gardiner and P. Zoller, *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics*. Springer Series in Synergetics. Springer, 2004.

[214] C. W. Gardiner and M. J. Collett, *Input and output in damped quantum systems: Quantum stochastic differential equations and the master equation*, *Phys. Rev. A* **31** (June, 1985) 3761–3774. Publisher: American Physical Society.

[215] R. Glauber, *Optical Coherence and Photon Statistics*, in *Quantum Theory of Optical Coherence*, pp. 23–182. John Wiley & Sons, Ltd, 2006. Section: 2 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9783527610075.ch2.

[216] H. Carmichael, *An Open Systems Approach to Quantum Optics: Lectures Presented at the Université Libre de Bruxelles, October 28 to November 4, 1991*. Springer Science & Business Media, May, 1993. Google-Books-ID: El5gxgxWhpgC.

[217] H. J. Carmichael, R. J. Brecha, M. G. Raizen, H. J. Kimble, and P. R. Rice, *Subnatural linewidth averaging for coupled atomic and cavity-mode oscillators*, *Phys. Rev. A* **40** (Nov., 1989) 5516–5519. Publisher: American Physical Society.

[218] *Quantum trajectories*, in *Quantum Measurement and Control* (G. J. Milburn and H. M. Wiseman, eds.), pp. 148–215. Cambridge University Press, Cambridge, 2009.

[219] S. P. Prasad, M. Maffei, P. A. Camati, C. Elouard, and A. Auffèves, *Closing Optical Bloch Equations in waveguide QED: Dynamics, Energetics*, Apr., 2024. arXiv:2404.09648 [quant-ph].

[220] C. Gerry and P. Knight, *Introductory Quantum Optics*. Cambridge University Press, Cambridge, 2004.

[221] D. Walls and G. J. Milburn, eds., *Quantum Optics*. Springer, Berlin, Heidelberg, 2008.

[222] P. Zoller and C. W. Gardiner, *Quantum Noise in Quantum Optics: the Stochastic Schr\"odinger Equation*, Feb., 1997. arXiv:quant-ph/9702030.

[223] T. Mori, *Floquet States in Open Quantum Systems*, *Annu. Rev. Condens. Matter Phys.* **14** (Mar., 2023) 35–56.

List of Publications

The results discussed in this dissertation are based on the following publications:

[A] J. Boeyens, S. Seah, and S. Nimmrichter, *Uninformed Bayesian quantum thermometry*, *Physical Review A* **104** (Nov., 2021) 052214

[B] J. Boeyens, B. Annby-Andersson, P. Bakhshinezhad, G. Haack, M. Perarnau-Llobet, S. Nimmrichter, P. P. Potts, and M. Mehboudi, *Probe thermometry with continuous measurements*, *New Journal of Physics* **25** (Dec., 2023) 123009.

[C] J. Boeyens, B. Yadin S. Nimmrichter, *How exchange symmetry impacts performance of collective quantum heat engines*, *New Journal of Physics* **27** (Jan., 2025) 013021

[D] J. Boeyens, J. Glatthard, E. Gandar, S. Nimmrichter, L. Correa , J. Rubio, *The role of symmetry and geometry in global quantum sensing*, *arXiv:2502.14817* **[quant-ph]** (Feb., 2025)

Acknowledgements

During the course of this research I have been helped significantly along the way:

- Firstly and most importantly, I am indebted to my supervisor Prof. Stefan Nimmrichter. Thank you for your patience, guidance, encouragement and support. Thank you also for brightening things up with your cynical sense of humour in the moments when research tends to get a bit too intense.
- Thank you to all my other co-authors: Stella Seah, Björn Annby-Andersson, Pharnam Bakhshinezhad, Géraldine Haack, Martí Perarnau-Llobet, Patrick. P. Potts, Mohammad Mehboudi, Benjamin Yadin, Jonas Glatthard, Edward Gandar, Jesús Rubio and Luis A. Correa.
- Thank you to Luis and the rest of the DQDD crew for a scientifically productive and at times life threatening research stay in Tenerife.
- Thank you to Benjamin, Chad and Lina for proof-reading parts of this thesis.
- I also would like to thank the group secretary Daniela for always being there to get me through the famous German bureaucracy.
- To all of the colleagues and friends I made in the TQO group, thank you for fostering a great working environment.
- I gratefully acknowledge financial support from the House of Young Talents of Siegen University.

On a personal note, I would like to thank my family for supporting me in moving to another continent to pursue this work. To my sisters, Caitie and Jackie, my mom, dad, René and Siobhan for always making time for a catch up and words of support when I needed them most.

Thank you to Kiara for being my first friend in Germany and sticking with me ever since, thanks for helping me not to take things too seriously but taking things seriously when I needed you to.

Finally, to Chad, any thanks I can think of is absurdly lacking in both scope and depth but thank you. Your unwavering belief in me and your daily support have meant everything. I wouldn't be here if it weren't for you.