Matthias D. Lang

Candidate

Physics and Astronomy

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Carlton M. Caves , Chairperson

Akimasa Miyake

David H. Dunlap

Rolando Somma

Measures of Nonclassical Correlations and Quantum-Enhanced Interferometry

by

Matthias D. Lang

M.S., Physics, University New Mexico, 2011

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Physics

The University of New Mexico

Albuquerque, New Mexico

July 2015

Dedication

To Kendra. Your relentless nagging finally got me to write this.
And to my parents, for supporting me in most of my endeavors.

Acknowledgments

I want to thank all the people that in some way or another helped me in getting this dissertation done.

First and foremost I am indebted to Carl Caves for the guidance, advice and financial support he provided. He was always patient, letting me work at my own pace and gave me all the freedom any grad student could dream of. His integrity in research is an inspiring example to all of his students.

Furthermore I have to thank Shashank Pandey and Zhang Jiang for all the help with various aspects of continuous variable quantum mechanics when I got started on my second project and Joshua Combes for answering all kinds of questions. Thanks to Alex Tacla for letting me listen to lengthy (sometimes even interesting) monologues, Jonas Anderson for showing me how to climb, Chris Ferrie for losing at darts all the time, Chris Cesare for brewing beer with me and Julian Antolin for drinking beer with me, Rob Cook, Ben Baragiola, Leigh Norris, Jonathan Gross, Ninnat Dangniam, Ivan Deutsch, Akimasa Miyake and the rest of the hospitable CQuIC- and PandA-crowd.

Outside of the physics department I need to thank my fiancée Kendra for putting up with me on a daily basis and my parents for making it possible for me to study physics in the first place. Moreover I am grateful for support in various forms from the rest of the Lang and the Lesser family.

More directly related to this dissertation, I have to thank Anil Shaji, my collaborator for the work presented in Chap. 4, as well as M. Piani and S. Vinjanampathy for useful discussions pertaining to the work of Chap. 4, which was supported in part by US National Science Foundation Grant Nos. PHY-1005540 and PHY-0903953. The work of Chap. 5 was supported by NSF Grant Nos. PHY-0653596 and PHY-0903953. For the work presented in Chap. 7 I would like to thank S. Szigeti and S. Haine for useful discussions and acknowledge support from NSF Grant Nos. PHY-1212445, PHY-1004440 and PHY-1314763 as well as Office of Naval Research Grant No. N00014-11-1-0082.

Measures of Nonclassical Correlations and Quantum-Enhanced Interferometry

by

Matthias D. Lang

M.S., Physics, University New Mexico, 2011

Ph.D., Physics, University of New Mexico, 2015

Abstract

In the first part of this dissertation a framework for categorizing entropic measures of nonclassical correlations in bipartite quantum states is presented. The measures are based on the difference between a quantum entropic quantity and the corresponding classical quantity obtained from measurements on the two systems. Three types of entropic quantities are used, and three different measurement strategies are applied to these quantities. Many of the resulting measures of nonclassical correlations have been proposed previously. Properties of the various measures are explored, and results of evaluating the measures for two-qubit quantum states are presented. To demonstrate how these measures differ from entanglement we move to the set of Bell-diagonal states for two qubits, which can be depicted as a tetrahedron in three dimensions. We consider the level surfaces of entanglement and of the correlation measures from our framework for Bell-diagonal states. This provides a complete picture of the structure of entanglement and discord for this simple case and, in particular, of their nonanalytic behavior under decoherence. The pictorial approach also indicates how to show that all of the proposed correlation measures are neither convex nor concave.

In the second part we look at two practical interferometric setups that use nonclassical states of light to enhance their performance. First we consider an

interferometer powered by laser light (a coherent state) into one input port and ask the following question: what is the best state to inject into the second input port, given a constraint on the mean number of photons this state can carry, in order to optimize the interferometer's phase sensitivity? This question is the practical question for high-sensitivity interferometry. We answer the question by considering the quantum Cramér-Rao bound for such a setup. The answer is squeezed vacuum. Then we analyze the ultimate bounds on the phase sensitivity of an interferometer, given the constraint that the state input to the interferometer's initial 50:50 beam splitter B is a product state of the two input modes. Requiring a product state is a natural restriction: if one were allowed to input an arbitrary, entangled two-mode state $|\Xi\rangle$ to the beam splitter, one could generally just as easily input the state $B|\Xi\rangle$ directly into the two modes after the beam splitter, thus rendering the beam splitter unnecessary. We find optimal states for a fixed photon number and for a fixed mean photon number. Our results indicate that entanglement is not a crucial resource for quantum-enhanced interferometry. Initially the analysis for both of these setups is performed for the idealized case of a lossless interferometer. Then the analysis is extended to the more realistic scenario where the interferometer suffers from photon losses.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

With the rapid technological development of the last century, the performance of nearly any device one can think of has been dramatically improved. The generic example here is the exponential increase in transistor densities of integrated circuits, known as *Moore's law*. While this predicted growth has been observed for almost fifty years, the first signs of its slowdown might soon become apparent [ZHB03]. This is not really surprising given the discrete structure of matter: A transistor cannot be smaller than one atom. One can still increase the computing power at this point, however, by simply using more resources and increasing the size of the computer.

Quite generally, limits on how well any device can perform, given a defined set of resources, are imposed by Nature. As Nature is described using quantum mechanics, the most fundamental limitations on device performance need a quantum-mechanical treatment.

While the description needs to be quantum mechanical to deliver accurate predictions, an interesting distinction can be made between devices using resources of a classical nature and devices using a more general class of resources that do not have a classical description. The goal of research in quantum information theory (QIT), for

example, is to classify and pin down the differences between information-processing protocols that use quantum resources and those protocols that have a full description in classical information theory.

Arguably, most of the researchers in quantum information theory are drawn to this field because of the exciting prospect it offers to tackle information-processing tasks that have been deemed infeasible on classical information processors no matter what their size is. The famous example that every student of the field cites when justifying their research to people not familiar with QIT, is *Shor's algorithm*. This algorithm allows factoring of integers on a quantum computer in polynomial 'runtime', a problem on whose classical hardness the security of the widely used RSA cryptosystem is based on. While *Shor's algorithm* presents the most striking selling point for research in QIT, current technology is still decades away from implementing a version of *Shor's algorithm* that could be used by people interested in breaking RSA encryption.

Moving from the ambitious goal of performing a task not feasible classically, such as *Shor's algorithm* or the idea of *quantum simulation*, to the more modest goal of simply performing a task which is feasible classically, but that can be done better using quantum resources, we can see various examples, where research in quantum information theory has led to ideas that either have been implemented or are at least implementable with current or near-future technology. Quantum cryptography systems are commercially available, and the use of quantum resources in metrological setups allowed for the design of 'quantum clocks', accurate to one second in 3.4 billion years and thus able to demonstrate gravitational time dilation within a single room. Nonclassical resources are also used to upgrade LIGO, the world's most sophisticated and powerful laser interferometer, in the hopes that this will push its sensitivity far enough to make the first direct observation of gravitational waves a reality.

On the one hand, research in QIT is devoted to analyzing practical situations and determining how quantum resources have to be deployed in these situations to provide a benefit. On the other hand, scientists working in the field explore the foundations of the theory, trying to get a better grasp on what it is in the very structure of quantum mechanics that is responsible for these *quantum enhancements*.

The research presented in this dissertation touches on both of these branches. In the first part nonclassical correlations beyond entanglement are discussed. These correlations might turn out to be a key piece in the puzzle of pinning down the resource that enables quantum systems to process information more efficiently than their classical counterpart. The second part is devoted to analyzing the ultimate limits on a practical interferometric setup that is allowed the use of nonclassical states of light. Such an interferometer can be used to measure various quantities, such as small distances, more precisely than possible with a standard laser interferometer and therefore could provide new insights in various fields of research.

1.2 Nonclassical correlations

In order to represent quantum information efficiently a quantum-information-processing system has to be composed of parts [BKCD02]. For multi-partite systems, correlations between these parts can arise that are conceptually different from correlations between classical systems.

One kind of nonclassical correlation is entanglement [HHHH09]. Entanglement is the crucial resource for such quantum-information-processing protocols as quantum key distribution, teleportation, and super-dense coding [HHHH09]. Moreover, Josza *et al.* [JL03] and Vidal [Vid03] showed that entanglement is a resource for pure-state quantum computation. A pure quantum state is unentangled if it is a product of pure states for each part. Josza *et al.* and Vidal provided a classical algorithm that efficiently simulates any pure state quantum computation where entanglement is negligible (more precisely, where entanglement is restricted to blocks of processing elements of a fixed size). While the exact amount of entanglement required for a pure-state quantum algorithm that promises an exponential speedup over a classical one performing the same task can be quite small [VdN13], Josza *et al.*'s and Vidal's results nevertheless prove that entanglement is necessary. The constructions by Josza *et al.* [JL03] and Vidal [Vid03] relies on the fact that there exists an efficient description of pure states that are not entangled.

For mixed states, the situation is more complicated. A mixed state is unentangled (separable) if it can be written as an ensemble of pure product states. Operational measures of entanglement are notoriously difficult to calculate for mixed states and even the boundary between separability and entanglement is difficult to characterize. In general, one can say, however, that the set of separable states is a convex set, is invariant under local unitary operations, and has dimension (as a manifold) as large as the space of mixed states [HHHH09]. This dimension, unfortunately, is the reason why for unentangled mixed states, no efficient description is available and the construction from [JL03] cannot be applied. It is unknown whether entanglement is necessary for an exponential speedup in an arbitrary and thus mixed-state quantum computation.

In the realm of mixed states, separable states can have nonclassical correlations even though they are unentangled. A state with only classical correlations, often called a classical state, is one that is diagonal in a product basis, for then the correlations are described by a joint probability distribution for classical variables of the parts. These purely classical states are a set of measure zero; this is suggested by the fact that any classical state can be perturbed infinitesimally to become nonclassical by making two of the eigenvectors infinitesimally entangled and is proved rigorously in [FAC$^+$10].

The fact that the states without any nonclassical correlations are a set of measure zero begs the question of whether there is an efficient description of these kinds of states. If so, a similar construction to the one proposed by Josza and Vidal might be used to show that a more general type of nonclassical correlation, not just entanglement, is a requirement for exponential speedups in quantum algorithms. Eastin [Eas10] followed this agenda and investigated whether a computation only involving classical states can be simulated efficiently on a classical computer. While he was able to show that computations comprised of qubits and gates involving not more than two qubits, provided the state remains a product state after all gates, can indeed be simulated efficiently on a classical computer, his result is apparently not extendable to a more general setting, and no one has taken up the challenge of finding

a more general result in this direction.

Another result hinting at a possible connection between nonclassical correlations beyond entanglement and quantum speedups is related to the model 'deterministic quantum computation with one quantum bit' (DQC1) [KL98]. While this model is not universal, it can perform at least one specific task for which there is no efficient classical algorithm known. It was shown in [DFC05] that in this model only a negligible amount of entanglement is present; Datta *et al.* [DSC08], however, demonstrated the presence of another type of nonclassical correlation, called *quantum discord*, in typical instances of DQC1.

While the connection of nonclassical correlations to quantum speedups remains tenuous [DacVB10], even this tenuous connection sparked a considerable interest. A variety of measures have been proposed to quantify nonclassical correlations for bipartite systems [OZ01, OHHH02, Luo08a, PCMH09, WPM09, BT10, LCS11, MPS$^+$10], in ways that can be nonzero for separable, but nonclassical states.

Our contribution to this field is to present a framework that unifies several of the proposed measures of quantum correlations that go beyond entanglement. While some new measures emerge from this framework, its main purpose is to investigate how the proposed measures relate to one another and to determine whether there is anything special about several existing measures, including quantum discord. Moving to Bell-diagonal states, a specific set of two-qubit states parametrized by three real parameters, allows us then to illustrate some of the properties of these correlation measures; conveniently all the different measures of non classical correlations agree on this subset of two-qubit states, and we can use a pictorial approach to explain some of the properties that are common to all these measures.

1.3 Interferometry with nonclassical states of light

The second topic that will be discussed within this thesis moves away from the foundational topic of quantum correlations to a more practically oriented problem in the field of quantum metrology. Specifically we will investigate how to use the quantum resource of nonclassical light fields most efficiently to optimize the performance of a practical interferometer.

The discovery that squeezed vacuum, injected into the normally unused port of an interferometer, provides phase sensitivity below the shot-noise limit [Cav81] led to thirty years of technology development, beginning with initial proof-of-principle experiments [XWK87, GSYL87] and culminating recently in the use of squeezed light to beat the shot-noise limit in the GEO 600 gravitational-wave detector [C$^+$11] and the Hanford LIGO detector [AAA$^+$13].

In the last decade much work has been devoted to exploring ultimate quantum limits on estimating the differential phase shift between two optical paths and to finding the states that achieve these limits. Given exactly N photons, the optimal state, in the absence of photon loss, is a *N00N* state, $(|N,0\rangle + |0,N\rangle)/\sqrt{2}$ [Ger00, BKA$^+$00, GBC02, LKD02], i.e., a superposition of all photons proceeding down one path with all photons proceeding down the other path. The *N00N* state is the optical analogue of the cat state that is optimal for atomic (Ramsey) interferometry [BIWH96]. Ideally, the sensitivity of a phase estimation setup using *N00N* states scales linearly with the energy used to produce the state. This provides a quadratic improvement over the best scaling possible when only employing classical resources, where the particle nature of photons gives rise to the shot-noise limit. Since the *N00N* state is extremely sensitive to photon loss, considerable effort has gone into determining optimal N-photon input states and corresponding sensitivities in the presence of photon loss [DDDS$^+$09, KSD11, EdMFD11a, EdMFD11b].

While these states indeed deliver optimal or near-optimal performance, given a

fixed input energy, we argue that they are not of practical relevance because they are very hard to produce with current technology and are therefore only available with quite low photon numbers. Consequently, the phase resolution obtained from using these optimal states cannot compete, even remotely, with the resolution obtained from a classical interferometer operating at or near the shot-noise limit with a strong, commercially available laser. This does not mean, however, that nonclassical states are useless for metrology. The use of squeezed states to enhance the sensitivity of the GEO 600 and LIGO interferometers is testimony to the efficacy of squeezed light in a situation where the lasers powering the interferometer have been made as powerful as design constraints allow.

In the work we present here, we turn the focus away from interferometry with states which can only be created with very small numbers of photons and instead start from the assumption of a more practical interferometric setting. We will still search for nonclassical input states that perform best, but we will do so under additional constraints that ensure that the states we find are not just of academic interest. Inspired by Caves' original proposal, we will look at an interferometer whose main source of power is a laser and ask the question: What is the optimal nonclassical state to inject into the second input port to maximize the interferometers phase sensitivity? In this way the main power production is separated from the production of nonclassical light. One can still make use of strong commercially available lasers, while the exotic state of light does not need to be powerful to obtain a significant improvement of the devices performance. We will prove that squeezed vacuum is the optimal choice in this setting, thus showing that in this sense the current LIGO setup is optimal.

The main problem in the production of $N00N$ states is the challenge to entangle two optical modes. To avoid this problem, another sensible question of practical importance is the following: what is the optimal unentangled input state to an interferometer? If we pursue this approach, any entanglement necessary for enhanced resolution gets produced by the first beam splitter of the interferometer. We analyze this situation for the two cases of a fixed photon number and mean photon number

constraint. Making this simple assumption, that the state in the two input modes of the interferometer is of product form, we find for the case of fixed photon number, that the optimal sensitivity is given by *identical* or *fraternal twin-Fock states*, depending on whether the photon number is even or odd. For the case of a mean-photon-number constraint, we will see that squeezed vacuum, now in both input modes, again yields the best sensitivity. While this product state of two squeezed vacua has essentially the same sensitivity as a *N00N* state, it happens to be an eigenstate of the first beam splitter of the interferometer. As a consequence, no modal entanglement is generated at all before the state is subjected to the phase shift. Hence, contrary to what one might expect given the discussion of *N00N* states, our findings indicate that modal entanglement is not a crucial resource in quantum-enhanced interferometry.

The analysis of the two scenarios described above is first performed under the assumption of an idealized interferometer, that is an interferometer which is not subjected to photon losses. We make use quantum Fisher information [Hel76, Hol11, BC94] as the figure of merit to quantify the performance of a particular input state to determine the phase shift occurring in the interferometer.

The main source of 'noise' that any linear optical setup is subjected to is loss of photons. Consequently any real world interferometer has to deal with photon losses. In recent years it has been shown by several groups [FI08, DDDS$^+$09, KSD11, EdMFD11a] that losses can be very detrimental to the performance of any state in an interferometric setup. Most important was the observation that asymptotically, as the mean photon number employed to produce a state goes to infinity, no state is able to achieve the sought after 'Heisenberg scaling', the quadratic improvement over the best classical scaling displayed by *N00N* states in the lossless regime.

In the light of this, an analysis that includes photon losses is desirable whenever ultimate limits of sensitivities are discussed. Our case, being chiefly motivated by practical concerns, makes a close account of the effects of losses mandatory. In the last part of this dissertation, we revisit the analysis of the optimal state for a laser-powered interferometer and for an interferometer powered by an arbitrary product state, given a mean photon number constraint, and extend the analysis to the case where photon

losses occur in the setup.

Losses transform pure states to mixed states and unlike for the pure-state case, there is no general, explicit expression for the quantum Fisher information available. This makes an analysis including losses much harder, as it is not possible to incorporate losses into calculations analogous to those performed for the lossless case. Using some additional tools, however, we are able to show that for the cases of practical interest, the optimal states for the lossless case remain optimal (or very nearly so) even when photon losses degrade the performance of the interferometer.

1.4 Organization of this dissertation

As stated above, this dissertation is comprised of contributions to two topics in quantum information theory, so it is divided into two distinct parts. The first part, Chaps. 2–5, deals with the foundational topic of nonclassical correlations that go beyond the notion of quantum entanglement, while the second part, Chaps. 6–8, treats a more practical question in the field of interferometry.

Within the first part, Chap. 2 presents some concepts of classical information theory and their counterparts in quantum information theory. It provides the background and introduces notation used in the first part of the dissertation. Chapter 3 discusses some results on measures of nonclassical correlations, some of which are investigated in more detail in Chap. 4, the central chapter for this first part. In this central chapter, a framework is introduced that demonstrates the relations between various measures of quantum correlations that have been proposed independently and that highlights the special properties of particular measures including quantum discord. The first part concludes with Chap. 5, where we investigate specific properties of these nonclassical correlation measures on a convenient subset of two-qubit states.

The second part of the dissertation starts in Chap. 6 with a basic discussion of parameter estimation and, more specifically, why Fisher information can be used as a figure of merit when one is interested in the sensitivity of a particular estimation

scheme. Chapters 7 and 8 give the central results of the analysis. Chapter 7 investigates optimal input states to an interferometer under certain constraints motivated by practical considerations in the absence of losses. This relatively simple, but important analysis is extended in Chap. 8 to the situation where the interferometer is subject to photon losses. The analysis that includes losses requires considerably more technical apparatus than the lossless analysis, but confirms its main conclusions about which input states are optimal.

1.5 List of publications

The following provides a list of publications encompassing most of the work presented in this dissertation.

1. M. D. Lang and C. M. Caves, *Quantum discord and the geometry of Bell-diagonal states*, Physical Review Letters **105**, 150501 (2010).

2. M. D. Lang, C. M. Caves and A. Shaji, *Entropic measures of non-classical correlations*, International Journal of Quantum Information **9**, 1553-1586 (2011).

3. M. D. Lang and C. M. Caves, *Optimal quantum-enhanced interferometry using a laser power source*, Physical Review Letters **111**, 173601 (2013).

4. Z. Jiang, M. D. Lang and C. M. Caves, *Mixing nonclassical pure states in a linear-optical network almost always generates modal entanglement*, Physical Review A **88**, 044301 (2013).

5. M. D. Lang and C. M. Caves, *Optimal quantum-enhanced interferometry*, Physical Review A **90**, 025802 (2014).

6. S. A. Haine, S. S. Szigeti, M. D. Lang, and C. M. Caves, *Heisenberg-limited metrology with information recycling*, Physical Review A, **91**, 041802(R) (2015).

The first two items in the list report the results of the first part of the dissertation, i.e., the results on measures of nonclassical correlations. Items 3–6 are related to the

second part of the dissertation, i.e., the results on practical interferometry. Items 3 and 5 are concerned with lossless interferometry in the two scenarios discussed in Chap. 7. Item 4 presents related results on how entanglement is generated in linear optical networks, and item 6 discusses applications of quantum Fisher information to an interferometric technique called information recycling. The results in items 4 and 6 are not discussed in the dissertation. The work on lossy interferometers, reported in Chap. 8, is in preparation for publication.

Chapter 2

Information Theory in the classical and the quantum world

2.1 Introduction

In this chapter I will present some basic concepts and quantities important in classical information theory and their generalization to the quantum world. My goal here is to describe these concepts such that someone outside of the field of quantum information can follow the main points of this dissertation. I do not intend to be rigorous when introducing some mathematical concepts, but will instead try to provide simple examples to encourage an intuitive understanding of these concepts. A detailed discussion of most of the following can be found in [NC00].

2.2 Bits

Arguably the most fundamental concept in all of information theory is the concept of a binary digit or as it is commonly called a bit. A bit is some system that can be in either of two states. Bits are used to store information. Physically, say in a computer, this is typically done on an integrated circuit, where voltage in a particular part of the

circuit is used to represent a bit. The two states a bit can take are represented by a 'high' and a 'low' voltage. In the following we will not be concerned about the physical realization of a bit, but rather talk about the abstract concept. This abstraction away from particular physical realizations is the hallmark of an information theory. As is standard practice, I will label the two states a bit can take by '0' and '1'.

2.2.1 Entropy

The most basic quantity in classical information theory, used to quantify information, e.g., the information carried by a bit, is the Shannon entropy H. Given a random variable A with an associated probability distribution (or probability mass function) p_a, the Shannon entropy $H(A)$ of this random variable is given by [NC00]

$$H(A) \equiv -\sum_a p_a \log p_a. \tag{2.1}$$

We will use the convention here and in the following that logarithms are to be taken to base 2. This will fix the unit of information to be bits. The above quantity tells us the amount of information we gain, on average, from learning the value of the random variable A.

An example of this would be a coin toss. Having two possible outcomes, 'heads' and 'tails', a coin toss can be thought of as a random bit A. Even though I will be talking about a coin toss, I will use the standard convention for bits '0' and '1' to label the outcomes. For a fair coin toss, we have

$$p_0 = p_1 = \frac{1}{2} \qquad \rightarrow \qquad H(A) = 1. \tag{2.2}$$

This means we gain 1 bit of information by looking at the outcome of the fair toss. On the other hand, if we look at the outcome of a maximally biased coin, with

$$p_0 = 0, \quad p_1 = 1 \qquad \rightarrow \qquad H(A) = 0, \tag{2.3}$$

we will gain no information, as there was no uncertainty to start with: the outcome of the toss was determined before we tossed the coin.

A strong operational motivation of why it is expression (2.1) that is commonly used to quantify the information is provided by *Shannon's source coding theorem* [CT12], which links $H(A)$ to the minimal average length of bit strings that faithfully represent the information in the random source A.

If we consider one additional system, say another bit B, we will use the *joint entropy* $H(A, B)$ to quantify the overall information in this joint system. By analogy with the above, the joint entropy is defined as

$$H(A, B) \equiv -\sum_{a,b} p_{a,b} \log p_{a,b}, \tag{2.4}$$

where $p_{a,b}$ is the joint probability density (mass) function, for system A being in the state a and system B being in the state b. Again take a coin toss as an example. We can imagine two independent and fair coins tossed together. We have

$$p_{0,0} = \frac{1}{4} = p_{0,1} = p_{1,0} = p_{1,1} \quad \rightarrow \quad H(A, B) = 2. \tag{2.5}$$

It makes sense intuitively that the information gained from 2 independent fair coins is twice the amount from a single fair coin. In the next section we will see what happens when the two coins are not independent.

2.2.2 Correlation and Mutual information

The two random variables A and B can be correlated, in which case they share some information. As an example for perfect correlation, one could image two fair coins glued together by their edges, both facing 'heads' up. The outcome of coin A is still completely random, but once we look at coin A, we know the outcome of coin B. Looking again at the probabilities and entropies, we see

$$p_{0,0} = \frac{1}{2} = p_{1,1}, \quad p_{0,1} = 0 = p_{1,0} \quad \rightarrow \quad H(A) = 1 = H(B) = H(A, B). \tag{2.6}$$

Here we see that the total information we can gain is present in either coin; looking at the other reveals nothing new. We can write this as

$$H(A|B) = 0 = H(B|A), \tag{2.7}$$

where $H(A|B)$ is to be read as the entropy of A given that that we know the state of B,

$$H(A|B) \equiv H(A,B) - H(B). \tag{2.8}$$

In this last example, one coin shares all its information with the other. A quantifier for this shared information is known as the mutual information,

$$H(A:B) \equiv H(A) + H(B) - H(A,B). \tag{2.9}$$

For two perfectly correlated coins we have $H(A:B) = H(A,B)$, which restates the fact that all of the information in this bipartite system is shared information. As should be obvious from the motivation, each of the quantities introduced above is a positive quantity. This enables us to display how these quantities relate in the Venn diagram of Fig. 2.1.

2.3 Qubits

A quantum bit, or a qubit for short, is the quantum version of a classical bit and the smallest conceivable quantum system. While a classical bit can be in one of two states, a qubit's state can be represented by any (normalized) vector of a two-dimensional Hilbert space $\mathcal{H}_2$. Using the standard 'bra-ket' notation we use the symbol $|a\rangle$ to denote a vector $\vec{a}$ and $\langle a|$ to denote its dual vector. To write down the state of a pure qubit, we typically pick an orthonormal basis. The most common one is called the computational basis and uses the basis vectors $|0\rangle$ and $|1\rangle$. We will use this basis in the following, unless we explicitly specify another one. Now we can write an arbitrary (pure) qubit state $|\psi\rangle$ as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{2.10}$$

where α and β are complex numbers such that $|\alpha|^2 + |\beta|^2 = 1$. It turns out that the global phase of a quantum state $|\psi\rangle$ is not observable and can be neglected,

$$|\psi\rangle \equiv e^{i\varphi}|\psi\rangle, \tag{2.11}$$

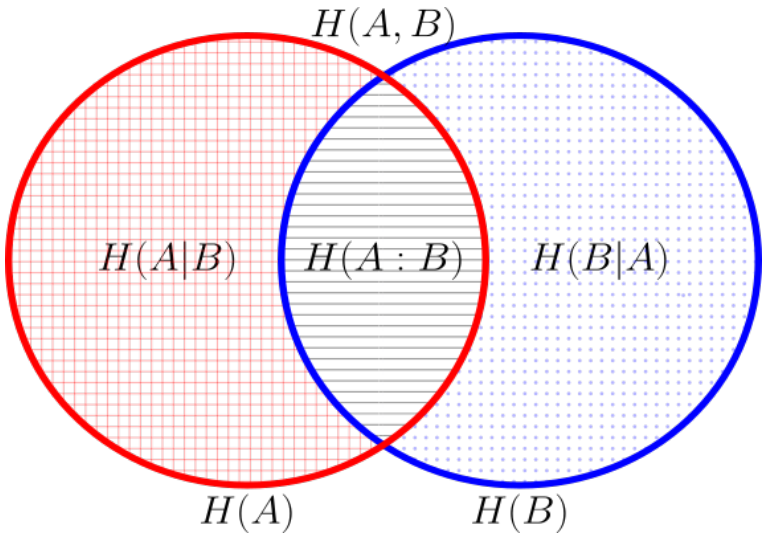

Figure 2.1: (Color online) The (red) circle on the left denotes the entropy associated with system A; the (blue) circle on the right denotes the entropy associated with system B. The area on the right filled in with (blue) dots is the information missing about B given complete information about A; this area denotes the conditional entropy $H(B|A)$. Similarly, the area on the left filled in with the (red) grid denotes $H(A|B)$. The overlap between the two circles, filled with horizontal lines, denotes the mutual information $H(A:B)$, which is the information contained in A about B and vice versa. The combined envelope of the two circles is the joint entropy $H(A,B)$. From the diagram, we have $H(B|A) = H(A,B) - H(A) = H(B) - H(A:B)$ and $H(A|B) = H(A,B) - H(B) = H(A) - H(A:B)$. For a classical joint probability distribution, the entropic measures are all Shannon entropies or relative Shannon entropies—thus they are guaranteed to be nonnegative—and they are related as the diagram depicts. For a bipartite quantum state, the joint quantum von Neumann entropy, $S(A,B)$, and the marginal von Neumann entropies, $S(A)$ and $S(B)$, replace $H(A,B)$, $H(A)$, and $H(B)$. The measures are related as depicted in the diagram, because the quantum conditional entropies, $S(B|A)$ and $S(A|B)$, and the quantum mutual information, $S(A:B)$, are *defined* by these relations. The difference is that $S(B|A)$ and $S(A|B)$, as so defined, can be negative, and thus the quantum mutual information $S(A:B)$ can be bigger than the marginal entropies, $S(A)$ and $S(B)$, and bigger than the joint entropy $S(A,B)$.

which enables us to parametrize the qubit state in terms of two angles θ and ϕ,

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle. \tag{2.12}$$

Given this parametrization we can geometrically think of the state of a pure qubit as a point on a sphere, the so-called Bloch sphere. Orthogonal states in this picture

lie at antipodal points on this sphere. A convention we will adopt is to identify the state $|0\rangle$ with the north pole of the sphere and $|1\rangle$ with its south pole.

Any spin-$\frac{1}{2}$ object, say the spin of an electron, is an example of this simplest of quantum objects. More commonly, however, in the lab two isolated levels of a more complex system such as an atom are used as a qubit.

The qubit we have been talking about up to now was always in a *pure state*. In analogy to a random classical bit, we will allow qubits to be in a statistical mixture of pure states, say an equal mixture of $|0\rangle$ and $|1\rangle$. These states are simply known as mixed states. Mixed states are convex combinations of more than one state and we will use the density matrix formalism to describe them. The density matrix ρ is a positive semi-definite matrix that describes a quantum state. The density matrix ρ of a pure state $|\psi\rangle$ is simply the outer product of is state vector with itself:

$$\rho = |\,\psi\,\rangle\langle\,\psi\,|. \tag{2.13}$$

To get the density matrix ρ_{mix} for the example above, we form the convex combination of states $|\,0\,\rangle\langle\,0\,|$ and $|\,1\,\rangle\langle\,1\,|$:

$$\rho_{mix} = p|\,0\,\rangle\langle\,0\,| + (1-p)|\,1\,\rangle\langle\,1\,|. \tag{2.14}$$

Written in the computational basis this becomes:

$$\rho_{mix} = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}. \tag{2.15}$$

For an equal mixture we have $p = 1/2$. Notice that this is different from the superposition $(|0\rangle + |1\rangle)/\sqrt{2}$, whose density matrix is

$$\rho_{sup} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \tag{2.16}$$

In the picture of the Bloch sphere, mixed states sit in the inner part of the sphere. The equal mixture of $|\,0\,\rangle$ and $|\,1\,\rangle$ for example is the point associated with the center of the sphere.

A common way of writing the density matrix of a qubit makes use of the Pauli matrices $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ as a hermitian operator basis:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.17}$$

With this, a qubit state ρ is specified by a vector $\vec{v}$ in the Bloch-sphere picture,

$$\rho = \frac{1}{2}(I_2 + \vec{v} \cdot \vec{\sigma}), \tag{2.18}$$

where $\vec{v}$ is called the *Bloch vector*. The Bloch vector of the mixed qubit from Eq. (2.14) is $\vec{v}_{mix} = (0, 0, 2p - 1)$, while the superposition in Eq. (2.16) has a Bloch vector of $\vec{v}_{sup} = (1, 0, 0)$. As these examples illustrate, the Bloch vector of a pure state is normalized, while the Bloch vector of a mixed states are sub-normalized.

If we have a composite system, e.g. two qubits, the vector space $\mathcal{H}_{comp}$ that the new state 'lives' in will be the tensor product of the two Hilbert spaces of its constituents, a 4-dimensional Hilbert space for two qubits $\mathcal{H}_4 = \mathcal{H}_2 \otimes \mathcal{H}_2$. To denote state vectors of a joint system, we will often omit the tensor product sign and write $|a\, b\rangle$ for $|\, a \,\rangle \otimes |\, b \,\rangle$, if it is obvious from the context what this notation will mean. Using this convention, the standard computational basis for the combined space will be $\{|\, 00 \,\rangle, |\, 01 \,\rangle, |\, 10 \,\rangle, |\, 11 \,\rangle\}$. Using the two states ρ_{mix} and ρ_{sup} as an example, their joint density matrix ρ_{joint} is:

$$\rho_{joint} = \rho_{mix} \otimes \rho_{sup} = \begin{pmatrix} p \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} & 0 \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \\ 0 \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} & (1-p) \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \end{pmatrix}. \tag{2.19}$$

Conversely we will use the *partial trace* to obtain the density matrix for one of the *marginal states* from the density matrix of the joint system. An index on the trace specifies which system is 'traced out'. Using the example above we have

$$\mathrm{tr}_B(\rho_{joint}) = \begin{pmatrix} p(\frac{1}{2} + \frac{1}{2}) & 0(\frac{1}{2} + \frac{1}{2}) \\ 0(\frac{1}{2} + \frac{1}{2}) & (1-p)(\frac{1}{2} + \frac{1}{2}) \end{pmatrix} = \rho_{mix}\,, \tag{2.20}$$

and

$$\mathrm{tr}_A(\rho_{joint}) = \begin{pmatrix} \frac{1}{2}(p+1-p) & \frac{1}{2}(p+1-p) \\ \frac{1}{2}(p+1-p) & \frac{1}{2}(p+1-p) \end{pmatrix} = \rho_{sup}\,. \tag{2.21}$$

Similar to Eq. (2.18) we can write arbitrary two qubit states as

$$\rho = \frac{1}{4}\Big(I_4 + \vec{a}\cdot\vec{\sigma}\otimes I_2 + I_2\otimes\vec{b}\cdot\vec{\sigma} + \sum_{i,j} c_{i,j}\,\sigma_i\otimes\sigma_j\Big). \tag{2.22}$$

Here $\vec{a}$ and $\vec{b}$ are the Bloch vectors of the respective marginal states; $c_{i,j} = \langle\sigma_i^A\otimes\sigma_j^B\rangle$ is a matrix defined in terms of expectation values of products of Pauli operators and contains information about the correlation of the two subsystems.

Again for the sake of a more compact notation, we will sometimes omit the tensor product sign. Especially when an operator O acts only on one system, say the Hilbert space of system $\mathcal{H}^{\mathcal{A}}$ of a composite system with Hilbert space $\mathcal{H}^{\mathcal{A}}\otimes\mathcal{H}^{\mathcal{B}}$, we will write O^A meaning $O\otimes I$.

2.3.1 Quantum measurements

Unlike a classical system, whose state we can measure without disturbing it, a quantum mechanical measurement is intrinsically invasive and will generally change the quantum state of the system being measured. We will use the 'positive-operator-valued-measure' (POVM) formalism to mathematically describe a measurement. While the POVM formalism describes more general measurements than the projective measurements usually discussed in textbooks, it only tells us about the statistics of the measurement and not about the state of the quantum system after the measurement has been performed. A full description of the measurement process requires using what are called quantum operations, which is beyond the scope of this brief introduction. This, however, should not overly concern us as these post-measurement states will not be relevant for most of the following discussion, although in one section quantum operations pop up briefly. A complete discussion of the general measurement formalism can be found in [NC00].

A POVM is defined by a set of positive operators, $\{E_m\}$. The members of the set are called *POVM elements*. The POVM elements satisfy the completeness relation

$$\sum_m E_m = I. \tag{2.23}$$

The probability p_m for measurement outcome m, when measuring on a quantum state ρ, is given by

$$p_m = \mathrm{tr}(E_m \rho). \tag{2.24}$$

If the set of POVM elements consists of orthogonal projectors, the measurement is called a projective measurement.

The POVM elements for a measurement in the computational basis are an example of a projective measurement with rank-one projectors:

$$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad E_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.25}$$

If we measure the equal mixture of the state $(|\,0\rangle\langle\,0\,| + |\,1\rangle\langle\,1\,|)/2$ in this basis, we get equal probabilities $p_0 = 1/2 = p_1$. At this point we can highlight the crucial difference between mixed states and superpositions of states. Looking at the measurement statistics for the state ρ_{sup} in Eq. (2.16), we see that it equals the statistics of the mixed state above. Changing the measurement basis to the eigenbasis of ρ_{sup}, however,

$$E_+ = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \qquad E_- = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}, \tag{2.26}$$

reveals that these two states differ drastically: $p_+ = \frac{1}{2}$, $p_- = \frac{1}{2}$ for the mixed state, but $p_+ = 1$, $p_- = 0$ for ρ_{sup}. Pure states can always be measured in their eigenbasis with certainty, but for mixed states, any rank-one POVM has at least two outcomes with nonzero probability.

2.3.2 Von Neumann entropy

Mixed states are the quantum analogue of a random bit, whereas pure states are the analogue of a bit in a defined state. Given this analogy, we expect a quantum entropy

measure to be zero for a pure state, but greater than zero for a mixed state. The von Neumann entropy $S(A)$ of a quantum system A with density matrix ρ,

$$S(A) = S(\rho) \equiv -\operatorname{tr}(\rho \log \rho), \tag{2.27}$$

fulfills these expectations and is the quantum analogue of the Shannon entropy. As is standard convention, functions of operators are applied to the eigenvalues in the eigenbasis of the operator. It turns out that the von Neumann entropy shares many of the properties the Shannon entropy [NC00]. Similar to its classical counterpart, *Schumacher's quantum noiseless channel coding theorem* gives the von Neumann entropy an operational meaning by connecting it the the minimum average number of qubits needed to reliably represent the quantum information contained in ρ.

Similar to the classical case we can add a second system and ask about the quantum information shared by these two systems. The generalization of joint entropy is straightforward. Mutual information can be written analogously to the classical case,

$$S(A : B) \equiv S(A) + S(B) - S(A, B). \tag{2.28}$$

The most obvious choice for the quantum conditional information $S(A|B)$ is

$$S(A|B) = S(A, B) - S(B). \tag{2.29}$$

One has to be careful, however, when looking at Fig. 2.1 to visualize these quantities. Most notable is that the quantum analogue of the conditional information is no longer necessarily a positive quantity. We will see an example of this in the next section on *entanglement*. Moreover, interpreting the conditional information as the information about system X given that we learned what state system Y is in, is not as straightforward as in the classical case. 'Learning the state of a system' involves a quantum measurement, which generally alters the state itself, and it seems we would have to use the quantum measurement formalism to define an appropriate conditional information. In Chap. 3, we will meet an alternative definition for the the conditional entropy based on this thought.

2.3.3 Quantum entanglement

Entanglement is a form of correlation between two quantum systems, which has no classical counterpart. The defining characteristic of a pure entangled state is that both of the marginal states are mixed. In fact the entropy of the marginal state is used to quantify the entanglement of the overall state. The canonical example for an entangled state is a Bell state

$$|\beta_{00}\rangle = (|00\rangle + |11\rangle)/\sqrt{2}. \tag{2.30}$$

We have $\mathrm{tr}_A(\rho_{\beta_{00}}) = I/2 = \mathrm{tr}_B(\rho_{\beta_{00}})$ and $S(\mathrm{tr}_A(\rho_{\beta_{00}})) = 1$. We say this state carries one bit of entanglement. The entropy of the marginal states is an entanglement measure called the *entropy of entanglement*. Unfortunately this only works for pure states. If we go to the realm of mixed states, entanglement is more difficult to quantify. What we can say is that a state ρ_{sep} is not entangled when its density matrix can be decomposed into product states:

$$\rho_{sep} = \sum_i p_i \rho_i^A \otimes \rho_i^B. \tag{2.31}$$

We say such a state is separable. While this boundary between entangled and unentangled states is clearcut, there are different ways of quantifying the entanglement. Several operationally motivated quantities such as entanglement of formation, squashed entanglement, entanglement cost, and a variety of others have been proposed [HHHH09]. These operational measures limit to the entropy of entanglement for pure states, but are notoriously difficult to calculate for mixed states; even the boundary between separabiltiy and entanglement is difficult to characterize, despite its seemingly simple definition. From Eq. (2.31), however, it is obvious that the set of separable states is a convex set. Moreover, it is known that the set is invariant under local unitary operations and has dimension as large as the space of mixed states [HHHH09].

Entanglement is the crucial resource for several quantum-information-processing protocols, such as quantum key distribution, teleportation and super-dense coding [HHHH09].

As we alluded to in the previous section, entangled states are an example of where the analogy to the classical entropic quantities breaks down. The conditional entropy of the Bell entangled state, for example, is negative, i.e., $S(A,B)-S(A)=0-1=-1$. So in some sense the overall state is defined while its constituents are completely random. This seems counter-intuitive, which is not surprising given that there is no analogue in the classical world to this phenomenon.

2.3.4 Decoherence

Decoherence is a process that reduces the purity of a state due to a system's interaction with other degrees of freedom, which one does not consider part of the system one is interested in. Commonly one calls those degrees of freedom collectively the environment. Decoherence turns pure states into mixed states and during this process may destroy quantum correlations. It is responsible for the fact that we do not encounter quantum effects in macroscopic systems, such as a superposition of a living and a dead cat. Decoherence poses a fundamental obstacle when trying to control a quantum system for a specific purpose. Quantum systems that are intrinsically well insulated from the environment—and therefore robust against decoherence—are generally hard to control because of their inaccessibility. On the other hand, quantum systems that are easy to control usually also couple strongly to the environment and therefore lose their quantumness easily.

Mathematically, decoherence can be described by the decay of off-diagonal elements of the density matrix in some basis preferred by the environment [Zur93]. As this is a nonunitary process, one typically invokes Lindblad-type master equations to describe quantum systems subjected to decoherence. As we will just briefly talk about decoherence, I will not describe the full formalism. Instead, I will limit myself to describing the effect of particular decoherence processes on particular quantum system when that situation arises within the dissertation.

One of the correlations that gets destroyed by decoherence is entanglement. Given the importance of entanglement for quantum information-processing tasks,

the preservation of entanglement is paramount to anyone trying to exploit quantum systems to perform tasks not feasible classically. Unfortunately entanglement displays a very unsettling behaviour when subjected to decoherence. While one might expect correlations like entanglement to decay in an exponential way it turns out that entanglement does show, generally, a decay following a nonanalytical curve, which makes it vanish completely in a finite amount of time. This phenomenon, dubbed 'entanglement sudden death' (ESD) [YE09, Col10], might pose specific challenges for the development of robust quantum information protocols, but it is not really surprising in the view of the geometry of separable states, since separable states have nonzero measure in the space of all states [ŻHSL98]. In a decoherence process that involves decay to a separable equilibrium state that does not lie on the boundary between separability and entanglement, the decohering state will cross that boundary before reaching the equilibrium state.

2.4 Maxwell's demon and Landauer's principle

Maxwell proposed the following Gedankenexperiment: Consider an isolated box filled with some gas in thermal equilibrium. The box is divided into two sections connected by a trapdoor. An intelligent being, a demon, sits in the box and can open and close the trapdoor. In particular he can see the single gas molecules. Maxwell supposed that such a demon could in principle lower the system's entropy by looking at the molecules that fly towards the trapdoor and opening or closing the trapdoor so that slower molecules collect on one side of the box and faster molecules on the other side. Maxwell intended to show up the limitations of the Second Law of Thermodynamics and to emphasize its probabilistic nature.

Maxwell's idea was debated for over 100 years in an effort to 'save' the Second Law, thus leading to interesting connections between thermodynamics and information theory [LR02]. The resolution generally accepted today was put forward by Bennett [Ben82] and is based on Landauer's principle. Landauer showed that the erasure of information is tied to an increase in entropy [Lan61]. This is usually

referred to as *Landauer's principle* today. The minimal cost for the erasure of one bit of information is $kT \ln 2$. Bennett argued that Maxwell's demon acquires information when measuring the particle's velocities in order to operate the trapdoor. As the system we are considering is isolated and the demon only has a finite memory, he needs to erase his memory periodically. The entropy increase from erasure balances the reduction of entropy from the separation of fast and slow particles.

In a famous paper of 1929 [Szi29], Szilárd envisioned a different incarnation of Maxwell's demon: Consider a cylinder containing only one particle in contact with a heat bath (the walls of the cylinder). A divider could be inserted, separating the cylinder into two equal parts. Upon acquiring the information in which part the particle is, a demon could insert a piston into the empty part, remove the divider, and let the particle isothermally push the piston outwards, doing some work W on a load coupled to the piston. The heat bath transfers energy Q=W to the particle in the process. After the particle pushes the piston out, it has the same volume accessible to it as in the beginning of the process. The heat bath having transferred energy to the particle, which does work, this cyclic process becomes a conversion of heat to work, in violation of the Second Law. The resolution to this paradox again involves accounting for the information acquired by the demon, which has to be erased at some point, given a finite memory, in order for the demon to be ready for a new cycle of the process. Interestingly Szilárd introduced the modern concept of information and used what was later called a bit by Shannon, as the demon has to acquire one bit of information in order to extract work.

The demons we will be taking about in Chap. 4 will be less ambitious than Maxwell's and Szilárd's. They won't try to challenge the Second Law of Thermodynamics, but work in accordance with it to transform a low entropy system to a high entropy system while extracting work and then will move on to another copy of the system. We will be interested in the net work these demons can extract on average keeping close account of the cost they have to pay for the erasure of information in the process.

This concludes the brief introduction to (quantum) information theory. The

focus of this chapter was to provide the background necessary to follow Chaps. 4 and 5 of this dissertation for a more detailed discussion of the ideas presented above see [NC00].

Chapter 3

Discord and other nonclassical correlations beyond entanglement

3.1 Introduction

At the heart of quantum information theory lies the objective to pinpoint the features that make it distinct from a classical theory. A complete understanding of these features would help us exploit this nonclassicality to perform information-processing-tasks that are not possible with a classical system. One obvious feature that makes a quantum system distinct from a classical system is the kind of correlations, having no classical counterpart, that can be shared by several quantum systems. Entanglement is one type of nonclassical correlation which has proven to be a crucial resource in several quantum information-processing-tasks. Separable states, however, can have nonclassical correlations even though they are unentangled.

Bennett *et al.* [BDF+99] found an orthogonal set of product states, that cannot be reliably distinguished by separate observers. As a mixture of these states is not entangled, this discovery hinted at the fact that some other type of correlation, different from entanglement, has to account for a distinctly nonclassical behavior of some systems.

A variety of measures have been proposed in an attempt to quantify nonclassical correlations for bipartite systems [OZ01, OHHH02, Luo08a, PCMH09, WPM09, BT10, LCS11, MPS+10] that can be nonzero for separable, but nonclassical states. In this chapter I will mainly focus on *Quantum Discord*, which is the quantity that attracted the most attention. In fact, sometimes all quantum correlations other than entanglement get lumped together under the name of quantum discord. I will try to be precise and use the name Quantum Discord, or simply just discord, only when referring to the quantity originally proposed by Ollivier and Zurek [Zur00, OZ01] and independently by Henderson and Vedral [HV01]. This section will not be comprehensive; I will only point out some results that are either of general interest or needed in the later chapters. A structured approach to the definition of these types of correlation measures is the focus of the next chapter, while a more complete account on the properties and various other results related to these correlation measures can be found in Modi *et al.'s* review article [MBC+12].

3.2 Discord

Ollivier and Zurek conceived quantum discord as a measure of disagreement between two forms of mutual information, $S(X:Y)$ and $J(Y \to X)$, that are equivalent classically. $S(X:Y)$ is the mutual information from Eqs. (2.9) and (2.28), while $J(Y \to X)$ is based on the identity for classical entropies:

$$H(X:Y) = H(X) - H(X|Y). \tag{3.1}$$

They noted that in order to generalize the expression $H(X|Y)$ to the quantum case one needs to specify a measurement basis in which information about system Y was obtained. Given a set of one-dimensional projectors $\{\Pi_y\}$ on system Y, the state of the combined system $\rho_{X|y}$ after obtaining measurement outcome y is

$$\rho_{X|y} = \frac{\Pi_y \rho_{X,Y} \Pi_y}{p_y}, \tag{3.2}$$

with $p_y = \mathrm{tr}(\Pi_y \rho_{X,Y})$.

Now Ollivier and Zurek [Zur00, OZ01] argued that an appropriate form for quantum version of the conditional entropy is

$$\sum_y p_y S(\rho_{X|y}) \equiv H_{\{\Pi_y\}}(X|Y), \tag{3.3}$$

and used this to define $S_{\{\Pi_y\}}(X:Y)$ in analogy to Eq. (3.1):

$$S_{\{\Pi_y\}}(X:Y) = S(X) - H_{\{\Pi_y\}}(X|Y). \tag{3.4}$$

This quantity is identical to the quantity that Henderson and Vedral [HV01] identified with the amount of classical correlations that can be extracted from a quantum system through the measurement $\{\Pi_y\}$. As a measure that characterizes all the classical correlations accessible in a quantum state, Henderson and Vedral proposed to maximize this quantity over all measurements. Intuitively a measure of quantum correlations is the difference between the total correlations as measured by $S(X:Y)$ and the classically accessible correlations measured by $\max_{\{\Pi_y\}} S_{\{\Pi_y\}}(X:Y) \equiv J(Y \to X)$. This is the quantity that Ollivier and Zurek dubbed quantum discord $\mathcal{D}(Y \to X)$:

$$\mathcal{D}(Y \to X) = S(X:Y) - J(Y \to X). \tag{3.5}$$

The arrow emphasises that this quantity is inherently asymmetric, as only one of the systems here is being measured. As is desirable for a correlation measure that is supposed to measure the truly quantum correlations, this quantity turns out to be non-negative[1]. Since it is based on entropic quantities, it is invariant under local unitary transformations, which intuitively should not change correlations between two systems. For the case of pure states, it is identical to the entanglement measures 2.3.3. The intriguing part about discord however is that in the realm of mixed states, it tries to capture more correlations than just entanglement. An example of a two qubit

[1]The Appendix contains a proof of the non-negativity of discord, but I urge the reader to wait until the next chapter before looking at the proof. We will slightly adapt the notation and introduce a second measurement when defining discord. While the resulting quantity is the same, it is seen from a different perspective.

state ρ_{disc} that is not entangled, but has nonzero discord can be found in [DacVB10]:

$$\rho_{disc} = \frac{1}{4}\left(|\,0\,\rangle\langle\,0\,|\otimes|\,+\,\rangle\langle\,+\,|+|\,1\,\rangle\langle\,1\,|\otimes|\,-\,\rangle\langle\,-\,|\right.$$
$$\left.+\,|\,+\,\rangle\langle\,+\,|\otimes|\,1\,\rangle\langle\,1\,|+|\,-\,\rangle\langle\,-\,|\otimes|\,0\,\rangle\langle\,0\,|\right). \tag{3.6}$$

Over the last 10 years discord has received a great deal of attention. Unfortunately, a big part of the research devoted to discord is not well motivated, because this quantity has no intrinsic operational meaning attached to it. While some progress has been made in providing operational interpretations [MBC+12], many of these attempts are arguably contrived.

In the following, I will briefly describe the operational interpretation for discord, that I, personally, find the most convincing. Here discord is linked a quantum communication protocol [MD11, MD13]. Suppose Alice and Bob share a bipartite quantum state. Alice's goal is to transfer her part of the state to Bob. To accomplish this task she can use local operations and classical communication (LOCC) as a classical resource as well as shared entanglement as a quantum resource. We will assume LOCC 'comes for free' and focus on the quantum resource needed to achieve their goal. One way to merge these states is for Alice to use the quantum teleportation protocol [BBC+93]. It turns out, however, that she can do better and only spend $S(A|B)$ bits of entanglement to accomplish her goal [HOW05]. Moreover, this protocol, called *quantum state merging*, also provides a suggestive interpretation of what it means that $S(A|B)$ can be negative. In that case, Alice does not have to spend entanglement bits to merge the states, but rather recovers ebits in the process.

If Bob measures his part of the system before he receives Alice's part, the cost for the state merging process will increase: Bob will destroy some correlations during his measurements, which Alice otherwise could have take advantage of. Madhok and Datta [MD11] showed that the amount the state merging cost increases when Bob measures his system first is quantified by the discord $D(B \to A)$ of the original state. This measurement could also be thought of as a decoherence process, where the environment 'measures' Bob's system and destroys correlations [MD13].

3.2.1 Other forms of nonclassical correlations

In 2003 Zurek [Zur03] showed that the advantage a nonlocal quantum demon has over his local classical counterparts in extracting work from a quantum system can be quantified through a form of discord. Unfortunately with this he started a trend in attaching the same name, discord, to different quantities that share some common traits. To emphasise that the quantity involved in the discussion of demons is different form the original discord, we will refer to it as 'demon discord'.

Another quantity that is similar to, but different from discord was proposed by Luo in 2008 [Luo08a]. The idea behind this quantity, which Luo called measurement-induced disturbance (MID), is to quantify how much correlation in a system gets disturbed when the system is measured. Unlike discord this quantity is symmetric under exchange of the two systems. An advantage over discord is that it is easier to compute. The measurements involved here are fixed to be performed in the eigenbasis of the marginal density operators, so unless the marginal states are degenerate, no optimisation procedure needs to be performed.

These are just two examples of a plethora of similar quantities [MPS+10] that try to capture more quantum correlations than entanglement does. We will meet some more of these measures in the next chapter and show how they are related. For now we will focus on a more general point, i.e., the set of states for which all these correlation measures vanish.

3.3 The set of classical states

A state is considered to be classical when its density matrix ρ_{cc} is diagonalized in a basis that is the product of basis elements for each subsystem [Luo08a]:

$$\rho_{cc} = \sum_{a,b} p_{a,b} |\, e_a \,\rangle\langle\, e_a \,| \otimes |\, f_b \,\rangle\langle\, f_b \,|. \tag{3.7}$$

A state written in this form is characterized by purely classical correlations. Moreover we can measure the state without disturbing it if we perform the measurements in

basis $\{|\, e_a \,\rangle\}$ for the first system and basis $\{|\, f_b \,\rangle\}$ for the second system. One kind of correlation measure, necessarily symmetric, will be, zero if and only if the joint state can be written in form of Eq. (3.7). The set of states for which discord vanishes is a bit bigger than that. In fact, if we consider system A being measured, in order for the joint state to have zero discord, one needs to be able to diagonalize the state ρ_{cq} in a conditional product basis [Dat10]:

$$\rho_{cq} = \sum_{a,b} p_{a,b} |\, e_a \,\rangle\langle\, e_a \,| \otimes |\, f_{b|a} \,\rangle\langle\, f_{b|a} \,|, \tag{3.8}$$

This means that the state is block-diagonal in a product basis:

$$\rho_{cq} = \sum_{a} p_a |\, e_a \,\rangle\langle\, e_a \,| \otimes \rho_{B|a}. \tag{3.9}$$

In this case only system A is undisturbed by an appropriate local measurement. Hence, these latter states ρ_{cq} are sometimes called 'classical-quantum', whilst the former ρ_{cc} are dubbed 'classical-classical'. Clearly the classical-classical states are a subset of the 'classical-quantum' state, which in turn are a subset of the separable states ρ_{sep} for Eq. (2.31).

Ferraro *et al.* [FAC+10] pointed out a big conceptual difference between the set of separable states and the set of classical-quantum states: They proved that the set of states with zero discord has measure zero in the space of all states and is nowhere dense. As a consequence a randomly selected bipartite quantum state will have non-zero discord. Moreover, a decohering state will not cross this classicality boundary before reaching an equilibrium state. As a consequence, there will be no phenomenon like 'entanglement sudden death' for discord. In Chap. 5, we will see a visualisation of this behaviour. While the state space we consider there will be restricted, it is well suited to demonstrate the points above. Moreover, we will see that unlike the set of separable states, the set of classical states is not convex.

While this chapter reported some properties on nonclassical correlations and presented discord in its original form, we will take on a slightly different perspective in the next chapter. There we take a more logical approach to nonclassical correlations and define a framework from which several measures on noncalssical correlations, one of them being quantum discord, emerge.

Chapter 4

A framework for entropic measures of nonclassical correlations

In this chapter we unify some of the measures discussed in Chap. 3 in a single framework, with the goal of clarifying the relations between these several quantities. Some ordering relations will follow naturally from the formulation of our framework; others we will have to prove explicitly.

4.1 Introduction

Maxwell demons observe a physical system and use the information obtained to extract work from the system [Max91]. For multipartite systems, we can distinguish quantum Maxwell demons, which have knowledge of the entire density operator and can manipulate and make measurements on the joint system, from classical demons, which can only perform operations and make measurements on the subsystems of the multipartite system. Because a single classical demon cannot be everywhere at the same time, it must recruit local demons to gather, process, and use information about the local systems; thus it is better to think of a classical demon as a collection of local demons. These local demons might or might not be allowed to communicate with each other using classical channels. The amount of work that the two kinds

of demons, quantum and classical, can extract from a given multipartite quantum state by employing protocols within each demon's means is a way of comparing quantum-information-processing protocols with classical ones.

This demonology [Zur00, OHHH02, Zur03, BT10] is but one of several attempts [Zur00, OZ01, HV01, RR02, Zur03, PHH08, Luo08a, WPM09, BT10, MPS+10] to track down and quantify the correlations that exist in multipartite quantum states. The nonclassical part of these correlations is not just quantum entanglement, even though entanglement is a part of it. The open question of pinning down why mixed-state quantum algorithms can solve certain problems exponentially faster than the best known classical ones [JL03], even in the absence of any significant entanglement, is one of the main motivations behind studying the nonclassical correlations in quantum states other than entanglement [DV07, DSC08, Dat08, DG09, Eas10].

We consider only bipartite states in this work. For our numerical work, the discussion is specialized yet further to states of two qubits. Correlations between systems can be quantified in terms of correlation coefficients and covariance matrices or in terms of entropic measures like mutual information. We choose the latter approach as the preferred one in information theory. The aim of this work is to formulate a framework in terms of which the several entropy-based measures of nonclassical correlations that have been proposed can be classified and understood. Constructing the framework leads to two new measures we have not seen previously in the literature. The focus here is not so much on unifying various measures, as in Ref. [MPS+10], but rather on clarifying the relationships among them.

The setting for our framework is two systems, A and B, with a joint quantum state ρ_{AB}. We consider three types of nonclassical-correlation measures, $\mathcal{M}(\rho_{AB})$, between A and B:

1. Mutual-information-based measures.

2. Conditional-entropy-based measures.

3. Demon-based (joint-entropy-based) measures.

The type-2 correlation measures can be asymmetric between A and B because conditional entropy is typically asymmetric.

As Landauer pointed out, when talking about demons, erasure of the demon's memory—and the associated thermodynamic cost—is an essential feature for assessing what a demon can do [Lan61]. As we mentioned above, a classical demon that works on a bipartite quantum system is best thought of as two local demons working in concert. Whether the two demons can communicate impacts their ability to coöperate. So the demon-based measures are thus further divided into two classes:

i. Erasure without communication between the demons.

ii. Erasure with communication between the demons.

All the measures of nonclassical correlations we consider here are constructed as the difference between a quantum entropic measure, $\mathcal{Q}(\rho_{AB})$, and its classical counterpart, $\mathcal{C}(\rho_{AB})$, which is derived from the probabilities for results of local measurements on one or both of the subsystems. The thinking behind this construction is that $\mathcal{Q}$ quantifies some notion of all the correlations in the system, whereas the corresponding classical $\mathcal{C}$ captures only the corresponding classical correlations. The difference, $\mathcal{M} = \mathcal{Q} - \mathcal{C}$, is therefore a way of quantifying the nonclassical correlations in the quantum state.

The results of local measurements are all that local classical observers (demons) can access, and these measurement results are used to probe the correlations (if any) between A and B. We do not want, however, our measure of nonclassical correlations to depend on the specifics of the measurement performed. Hence, in its construction, the classical measure, $\mathcal{C}(\rho_{AB})$, is maximized over all possible measurements within specific measurement strategies that are defined beforehand. In some instances, when maximization is necessary, we are able to show that the maximum is attained on rank-one POVMs; in other cases, we restrict the maximization to rank-one POVMs. We give a full discussion of these different situations and the issues surrounding rank-one POVMs after we have developed our framework.

We thus imagine that there are classical observers A and B—demons or otherwise—who have access to the two parts of the bipartite system. We allow these observers to employ one of three measurement strategies:

a. Local, rank-one-projector measurements in the eigenbases of the marginal density operators.

b. Unconditioned local measurements.

c. Conditioned local measurements.

For strategy (a), the local measurements are unique modulo degeneracies in the marginal density operators. The other two strategies require maximization of the classical measure $\mathcal{C}$ over the measurements allowed by the strategy. The first two measurement strategies do not require the observers to communicate with each other, but the last one does. Consequently, the first two strategies are symmetric between A and B. For the third strategy, A performs a measurement and communicates the result to B, who can then condition his measurement on the result communicated by A. This makes the nonclassical correlation measures that are based on the third measurement strategy asymmetric between A and B.

We now have three types of correlation measures and three measurement strategies, and we can label the resulting correlation measures with the type and the strategy. For example, $\mathcal{M}_{1b}$ refers to the nonclassical correlation measure constructed as the difference between quantum and classical mutual informations, where unconditioned local measurements are used to construct the classical mutual information.

There is a natural hierarchy in the three types of measurements strategies. Allowing arbitrary, unconditioned local measurements, as in strategy (b), is a restriction of the conditioned local measurements of strategy (c), since to get (b) from (c), observer B simply chooses to ignore any communication A might have sent regarding her measurement results. Likewise, measuring in the local eigenbases of the marginal density operators, as in strategy (a), is a restriction of the arbitrary, unconditioned local measurements of strategy (b). Thus, when we maximize over the measurements

in a particular strategy, the classical measure $\mathcal{C}$ cannot decrease—and generally it increases—as we move from (a) to (b) to (c). This is saying that the more general the measurements the local observers are allowed to do, the more they can expect to discover about any classical correlations that exist between the subsystems. Since our nonclassical-correlation measure $\mathcal{M}$ is the difference between $\mathcal{Q}$ and $\mathcal{C}$, $\mathcal{M}$ cannot increase—and generally it decreases—as we move from (a) to (b) to (c), i.e., $\mathcal{M}_{ja} \geq \mathcal{M}_{jb} \geq \mathcal{M}_{jc}$ for $j = 1, 2, 3$.

In Sec. 4.2 we formulate our framework: Sec. 4.2.1 reviews the bipartite entropic information measures that we use in constructing our framework; Sec. 4.2.2 spells out the description of local measurements for strategies (a)–(c); Sec. 4.2.3 defines the nonclassical-correlation measures and discusses relations among them; and Sec. 4.2.4 considers the issues raised by assuming the local measurements are described by rank-one POVMs and also whether one can specialize further to measurements described by rank-one projectors. In Sec. 4.3 we present numerical results comparing the various measures for two-qubit states, assuming that the local measurements can be described by orthogonal rank-one projection operators. App. A provides additional information.

4.2 Framework for entropic measures of nonclassical correlations

In this section we develop our framework for measures of nonclassical correlations and explore properties of the various measures the framework leads to.

4.2.1 Entropic measures of information and correlation

Entropic measures of information quantify how much information can be extracted from a system or, more poetically, how much information is "missing" about the fine-grained state of the system.

Figure 2.1 is a useful pictorial representation of the relationships among the entropies and entropic measures of correlation that apply to bipartite systems. The figure provides an accurate representation for classical entropies. In the quantum case, some of the quantities cannot be represented or are misrepresented by this diagram, but even so, the diagram is a useful tool because it captures correctly the relationships among the various entropies.

For a bipartite state ρ_{AB} of systems A and B, the quantum entropic quantities that will be used in the ensuing discussion are the following:

1. $S(A,B) = S(\rho_{AB}) = -\mathrm{tr}(\rho_{AB}\log\rho_{AB})$, the joint von Neumann entropy of the whole system.

2. $S(A) = S(\rho_A) = -\mathrm{tr}_A(\rho_A\log\rho_A)$ and $S(B) = S(\rho_B) = -\mathrm{tr}_B(\rho_B\log\rho_B)$, the von Neumann entropies of the marginal density operators.

3. $S(B|A) = S(A,B) - S(A)$ and $S(A|B) = S(A,B) - S(B)$, the quantum conditional entropies.

4. $S(A:B) = S(A) + S(B) - S(A,B)$, the quantum mutual information, which is related to the quantum conditional entropies by $S(A:B) = S(B) - S(B|A) = S(A) - S(A|B)$. The quantum mutual information can also be written as a quantum relative entropy,

$$S(A:B) = S(\rho_{AB}||\rho_A \otimes \rho_B)\ , \tag{4.1}$$

where the relative entropy is defined by

$$S(\rho||\sigma) = -S(\rho) - \mathrm{tr}(\rho\log\sigma)\ . \tag{4.2}$$

Local measurements on the bipartite quantum system are described by a joint probability distribution p_{ab} for outcomes labeled by a and b. Bayes's theorem relates the joint, conditional, and marginal distributions: $p_{b|a}p_a = p_{ab} = p_{a|b}p_b$. These distributions are used to define the classical information measures:

1. $H(A,B) = H(p_{ab}) = -\sum_{a,b} p_{ab} \log p_{ab}$, the Shannon entropy of the joint distribution p_{ab}.

2. $H(A) = H(p_a) = -\sum_a p_a \log p_a$ and $H(B) = H(p_b) = -\sum_a p_b \log p_b$, the Shannon entropies of the marginal distributions, p_a and p_b.

3. $H(B|A) = H(A,B) - H(A) = \sum_a p_a H(B|a)$ and $H(A|B) = H(A,B) - H(B) = \sum_a p_b H(A|b)$, the classical conditional entropies. $H(B|a) = -\sum_b p_{b|a} \log p_{b|a}$ and $H(A|b) = -\sum_a p_{a|b} \log p_{a|b}$ are the Shannon entropies of the conditional distributions $p_{b|a}$ and $p_{a|b}$; the conditional entropies are averages of $H(B|a)$ over p_a and $H(A|b)$ over p_b.

4. $H(A:B) = H(A) + H(B) - H(A,B) = \sum_{a,b} p_{ab} \log(p_{ab}/p_a p_b)$, the classical mutual information. $H(A:B)$ is the relative information of the joint distribution p_{ab} with respect to the product of the marginals, $p_a p_b$,

$$H(A:B) = H(p_{ab}||p_a p_b) \; ; \tag{4.3}$$

the classical relative information, which is always nonnegative, is defined by

$$H(p_j||q_j) = \sum_j p_j \log(p_j/q_j) = -H(p_j) - \sum_j p_j \log q_j \; . \tag{4.4}$$

We also have $H(A:B) = H(B) - H(B|A) = H(A) - H(A|B)$.

Figure 2.1 in Chap. 2 summarizes the relations among the classical entropies; it works because the classical conditional entropies and the classical mutual information are all nonnegative. This leads to several inequalities that can be read off Fig. 2.1. For example, we can see that

$$\max\big(H(A), H(B)\big) \leq H(A,B) \leq H(A) + H(B) \; . \tag{4.5}$$

The lower bound on $H(A,B)$ is saturated when knowing one subsystem completely determines the other (the two circles in Fig. 2.1 are either identical or become nested), i.e., $H(A:B) = \min\big(H(A), H(B)\big)$. The upper bound is saturated when there are no correlations between A and B, i.e., $H(A:B) = 0$, so determining one subsystem gives

no information about the other (the two circles in Fig. 2.1 are disjoint). For quantum entropies the lower bound in Eq. (4.5) does not hold, which is equivalent to saying the quantum conditional entropies can be negative. The simplest counter-example is a two-qubit Bell state: the joint state is pure and, hence, has zero entropy, but the marginal states are completely mixed, so their entropies are maximal and both equal to one.

4.2.2 Local measurements

We now spell out the general description of the local measurements that applies to measurement strategies (a)–(c). Although we only need measurement statistics—and, hence, only need POVMs—to evaluate the classical entropic measures, we start our description with quantum operations, partly to be general and partly so we can deal with post-measurement states in a subsequent discussion of Maxwell demons.

The measurement on A is described by quantum operations [NC00] that are labeled by the possible outcomes a of the measurement on A:

$$\mathcal{A}_a = \sum_\alpha A_{a\alpha} \odot A_{a\alpha}^\dagger \ . \tag{4.6}$$

The quantum operation is applied to a density operator by inserting the density operator in place of the $\odot$. The operators $A_{a\alpha}$, the Kraus operators of $\mathcal{A}_a$, combine to give the POVM element for outcome a,

$$E_a = \sum_\alpha A_{a\alpha}^\dagger A_{a\alpha} \ , \tag{4.7}$$

and the POVM elements satisfy a completeness relation, $I_A = \sum_a E_a$.

The absence of communication in strategies (a) and (b) makes them quite straightforward. The measurement on B is described by a set of quantum operations,

$$\mathcal{B}_b = \sum_\beta B_{b\beta} \odot B_{b\beta}^\dagger \ . \tag{4.8}$$

These give POVM elements

$$F_b = \sum_\beta B_{b\beta}^\dagger B_{b\beta} \ , \tag{4.9}$$

which satisfy a completeness relation $I_B = \sum_b F_b$. The state of the joint system after measurements with outcomes a and b is $\rho_{AB|ab} = \mathcal{A}_a \otimes \mathcal{B}_b(\rho_{AB})/p_{ab}$, where

$$p_{ab} = \mathrm{tr}\big(\mathcal{A}_a \otimes \mathcal{B}_b(\rho_{AB})\big) = \mathrm{tr}(E_a \otimes F_b \rho_{AB}) \tag{4.10}$$

is the joint probability for outcomes a and b. The post-measurement joint state and the joint probability marginalize to the subsystems in the standard way.

We need to be more careful with strategy (c) because of the communication from A to B. We handle strategy (c) in a general way that allows us to interpolate between (b) and the extreme case of (c) in which every outcome a leads to a different measurement on B. We do this by introducing a set C whose elements c label the possible measurements to be made on B. We let A stand for the set of outcomes a, and we define a function $c(a)$ that maps an outcome a to the corresponding value in C. We let $A_c = \{a \mid c(a) = c\}$ be the subset of A that leads to the B measurement labeled by c. The subsets A_c partition A into disjoint subsets. We can regard C as another variable in our analysis; it is a coarse graining of the measurement on A. Formally, we have that C is perfectly correlated with A, i.e., $p_{c|a} = \delta_{c,c(a)}$, implying that $H(C|A) = 0$ and $H(A:C) = H(C)$. Should there be only one possible measurement on B, i.e., only one value of c, then there is no communication, and the situation reduces to strategy (b). The extreme case of (c) corresponds to having a different value of c for each outcome a, in which case there is no difference between the outcome set A and the set C.

The state of the joint system after the measurement on A yields outcome a is $\rho_{AB|a} = \mathcal{A}_a(\rho_{AB})/p_a$, where

$$p_a = \mathrm{tr}\big(\mathcal{A}_a(\rho_{AB})\big) = \mathrm{tr}_A(E_a \rho_A) \tag{4.11}$$

is the probability for outcome a. The state of system B, conditioned on outcome a, is

$$\rho_{B|a} = \mathrm{tr}_A(\rho_{AB|a}) = \frac{\mathrm{tr}_A\big(E_a \rho_{AB}\big)}{p_a} \ ; \tag{4.12}$$

notice that this is determined by the POVM element E_a. The probability for making

measurement c on B follows formally from

$$p_c = \sum_a p_{c|a} p_a = \sum_{a \in A_c} p_a = \mathrm{tr}_A(E_c \rho_A) \ . \tag{4.13}$$

Here we introduce coarse-grained POVM elements for the measurement on A, labeled by the measurement to be made on B:

$$E_c = \sum_{a \in A_c} E_a \ . \tag{4.14}$$

Notice that if there is only one possible measurement on B, i.e., only one value of c, then $E_c = I_A$; when there is a different measurement for each outcome a, the POVM elements E_c are the same as the POVM elements E_a. We also have the state of B conditioned on the coarse-grained outcome c:

$$\rho_{B|c} = \frac{\mathrm{tr}_A\big(E_c \rho_{AB}\big)}{p_c} \ . \tag{4.15}$$

Notice that Eqs. (4.12) and (4.15) imply that

$$\rho_B = \sum_a p_a \rho_{B|a} = \sum_c p_c \rho_{B|c} \ . \tag{4.16}$$

We turn our attention now to the measurements on B. We let B stand for the set of all outcomes on B for all the possible measurements on B. We define a function $c(b)$ that maps an outcome b to the measurement c in which it occurs, and we define $B_c = \{b \mid c(b) = c\}$ to be the subset of B outcomes for the measurement labeled by c. The subsets B_c partition the set of all possible outcomes on B into disjoint subsets. We again have perfect correlation, i.e., $p_{c|b} = \delta_{c,c(b)}$, implying that $H(C|B) = 0$ and $H(B:C) = H(C)$.

The measurement on B that is labeled by c is described by quantum operations

$$\mathcal{B}_{b|c} = \sum_\beta B_{b\beta|c} \odot B^\dagger_{b\beta|c} \ , \tag{4.17}$$

The Kraus operators give the POVM elements for this measurement,

$$F_{b|c} = \sum_\beta B^\dagger_{b\beta|c} B_{b\beta|c} \ , \tag{4.18}$$

and these satisfy a completeness relation $I_B = \sum_{b\in B_c} F_{b|c}$. In sums over b, we can let the sum run over the outcomes of all the possible measurements on B by the artifice of defining $B_{b\beta|c} = 0$ for $b \notin B_c$ and, hence, $F_{b|c} = 0$ for $b \notin B_c$.

The state of the joint system, conditioned on outcomes a and b, is

$$\rho_{AB|ab} = \frac{\mathcal{A}_a \otimes \mathcal{B}_{b|c(a)}(\rho_{AB})}{p_{ab}} = \frac{\mathcal{B}_{b|c(a)}(\rho_{AB|a})}{p_{b|a}} \,, \tag{4.19}$$

where

$$p_{ab} = \mathrm{tr}\big(\mathcal{A}_a \otimes \mathcal{B}_{b|c(a)}(\rho_{AB})\big) = \mathrm{tr}(E_a \otimes F_{b|c(a)}\rho_{AB}) = p_a \mathrm{tr}_B(F_{b|c(a)}\rho_{B|a}) \tag{4.20}$$

is the joint probability for a and b and

$$p_{b|a} = \mathrm{tr}\big(\mathcal{B}_{b|c(a)}(\rho_{AB|a})\big) = \mathrm{tr}(F_{b|c(a)}\rho_{B|a}) \tag{4.21}$$

is the conditional probability for b given a. Notice that p_{ab} and $p_{b|a}$ are nonzero only if $b \in B_{c(a)}$ or, equivalently, only if $a \in A_{c(b)}$.

For our purposes, it is easier to work with the coarse-grained outcomes c, which specify the measurements on B. Indeed, the joint probability for b and c is

$$p_{bc} = \sum_a p_{c|ab}p_{ab} = \sum_{a\in A_c} p_{ab} = \mathrm{tr}(E_c \otimes F_{b|c}\rho_{AB}) = p_c \mathrm{tr}_B(F_{b|c}\rho_{B|c}) \,. \tag{4.22}$$

Notice that p_{bc} is nonzero only if $b \in B_c$. Thus the conditional probability of b given c takes the form

$$p_{b|c} = \frac{p_{bc}}{p_c} = \mathrm{tr}_B(F_{b|c}\rho_{B|c}) \,, \tag{4.23}$$

and the unconditioned probability for b is

$$p_b = \sum_c p_{bc} = \mathrm{tr}(E_{c(b)} \otimes F_{b|c(b)}\rho_{AB}) = p_{c(b)} \mathrm{tr}_B(F_{b|c(b)}\rho_{B|c(b)}) \,. \tag{4.24}$$

4.2.3 Measures of nonclassical correlations

In this subsection we formulate our framework for entropic measures of nonclassical correlations, considering in turn the three types of measures introduced in Sec. 4.1

and for each type, the three local measurement strategies, (a), (b), and (c). For strategy (a), the local measurements are in the eigenbases of the marginal density operators. *For strategies (b) and (c), we assume that the measurements are described by rank-one POVMs*, which means that E_a and $F_{b|c}$ are multiples of rank-one projection operators. We discuss this assumption in Sec. 4.2.4.

To compare and relate the various measures, we rely on two inequalities that relate the quantum and the classical entropies: the *POVM inequality* (see App. A.1 for a proof) and the *ensemble inequality* [NC00].

The POVM inequality relates the quantum entropy for a state ρ to the classical entropy for probabilities $p_j = \mathrm{tr}(E_j\rho)$ obtained from (nonzero) POVM elements E_j:

$$H(p_j) + \sum_j p_j \log(\mathrm{tr}E_j) = -\sum_j p_j \log\left(\frac{p_j}{\mathrm{tr}E_j}\right) \geq S(\rho) \; . \tag{4.25}$$

A rank-one POVM is one such that all the POVM elements are rank-one, i.e., $E_j = \mu_j P_j$, where P_j is a rank-one projection operator and $0 \leq \mu_j = \mathrm{tr}E_j \leq 1$. The trace of the completeness relation implies that $\sum_j \mu_j =$ (dimension of the quantum system). For a rank-one POVM, we have

$$H(p_j) \geq S(\rho) - \sum_j p_j \log \mu_j \geq S(\rho) \; . \tag{4.26}$$

The ensemble inequality [NC00] says that the Shannon information of a set of ensemble probabilities q_j exceeds the Holevo quantity of the ensemble:

$$H(q_j) \geq S\left(\sum_j q_j\rho_j\right) - \sum_j q_j S(\rho_j) \; . \tag{4.27}$$

For strategy (a), where the local measurements are in the eigenbases of the marginal density operators, we have immediately that $H(A) = S(A)$ and $H(B) = S(B)$. For both (b) and (c), we can apply the POVM inequality in its rank-one form to $p_a = \mathrm{tr}(E_a\rho_A)$ to conclude that $H(A) \geq S(A)$. Similarly, for strategy (b), the POVM inequality applied to $p_b = \mathrm{tr}(F_b\rho_B)$ gives $H(B) \geq S(B)$. For strategy (c), we need a

chain of inequalities to conclude that $H(B) \geq S(B)$:

$$\begin{aligned} H(B) = H(C,B) &= H(C) + H(B|C) \\ &= H(p_c) + \sum_c p_c H(B|c) \\ &\geq H(p_c) + \sum_c p_c S(\rho_{B|c}) \qquad (4.28) \\ &\geq S\left(\sum_c p_c \rho_{B|c}\right) \qquad (4.29) \\ &= S(\rho_B) = S(B)\ . \qquad (4.30) \end{aligned}$$

The first inequality (4.28) is a consequence of applying the POVM inequality to Eq. (4.23), the second inequality (4.29) is an example of the ensemble inequality, and the final equality uses Eq. (4.16).

Type 1: Mutual-information-based measures

For type-1 measures, we choose $\mathcal{Q}_1 = S(A:B)$ and $\mathcal{C}_1 = H(A:B)$, giving the difference measure

$$\mathcal{M}_1 = S(A:B) - H(A:B)\ . \qquad (4.31)$$

We now apply the three measurement strategies introduced in Sec. 4.1 to obtain the classical mutual information $H(A:B)$; this leads to three different type-1 measures.

For strategy (a), the local measurements are made in the eigenbases of the marginal density operators, and this gives a nonclassical-correlation measure that we denote by $\mathcal{M}_{1a}$. If the marginal density operators have nondegenerate eigenvalues, the marginal eigenbases are unique; in the case of degeneracy, one needs to maximize $H(A:B)$ over the rank-one, projection-valued measurements in the degenerate subspaces to get a unique measure $\mathcal{M}_{1a}$. The measure $\mathcal{M}_{1a}$ was introduced by Luo in [Luo08a] and called there the *measurement-induced disturbance* (MID). The same measure, in a different guise, had been proposed by Rajagopal and Randall in [RR02]; they defined what they called the *quantum deficit* as $H(A,B) - S(A,B)$, where $H(A,B)$ is obtained from measurements in the marginal eigenbases. The quantum deficit and

MID are the same because they differ by the terms $H(A) - S(A)$ and $H(B) - S(B)$, which are zero for measurements in the marginal eigenbases.

When strategy (b) is used, we obtain the measure

$$\mathcal{M}_{1b} = S(A:B) - \max_{\text{(b)}} H(A:B) \ , \tag{4.32}$$

where the classical mutual information has to be maximized over the unconditioned local measurements of strategy (b). The maximum classical mutual information was introduced in [PHH08] as a measure of classical correlations, and the same paper suggested $\mathcal{M}_{1b}$ as a measure of nonclassical correlations. This measure was investigated in detail by Wu, Poulsen, and Mølmer (WPM) in [WPM09], and we refer to it as the *WPM measure*, while denoting it as $\mathcal{M}_{1b}$. The optimal unconditioned local measurements are not necessarily orthogonal-projection-valued. An example of a case in which the maximization requires POVMs and not just projective measurements was given in [WPM09]; we review and extend this example in App. A.3. In addition, the optimal local measurements do not generally occur in the marginal eigenbases, which implies that $\mathcal{M}_{1a}^{(\text{MID})} \geq \mathcal{M}_{1b}^{(\text{WPM})}$.

For strategy (c), the classical mutual information $H(A:B)$ can be made arbitrarily large, thus allowing $\mathcal{M}_{1c}$ to be arbitrarily negative. This is easy to see by considering the extreme case of (c) in which every outcome a leads to a different measurement on system B; then, as noted in Sec. 4.2.2, $H(A:B) = H(A)$, which can be as big as desired by giving the measurement on A an arbitrarily large number of outcomes. We conclude that $\mathcal{M}_{1c}$ has nothing to do with quantifying nonclassical correlations, so we drop $\mathcal{M}_{1c}$ from our array of possible measures.

Type 2: Conditional-entropy-based measures

For type-2 measures, we choose $\mathcal{Q}_2 = -S(B|A)$ and $\mathcal{C}_2 = -H(B|A)$. The result is the difference measure

$$\mathcal{M}_2 = H(B|A) - S(B|A) \ . \tag{4.33}$$

We notice immediately that

$$\mathcal{M}_2 = \mathcal{M}_1 + \big[H(B) - S(B)\big] \geq \mathcal{M}_1 \ . \tag{4.34}$$

This shows that a type-1 measure is always less than or equal to the type-2 measure that uses the same measurement strategy, with equality only when B is measured in the marginal eigenbasis.

Measurements in the eigenbases of the marginal density operators have $H(B) = S(B)$, so for strategy (a), we have $\mathcal{M}_{2a} = \mathcal{M}_{1a}$, and our measure is again MID.

Strategy (b) gives the measure

$$\mathcal{M}_{2b} = \min_{\text{(b)}} H(B|A) - S(B|A) \ , \tag{4.35}$$

where we have to minimize $H(B|A)$ over all unconditioned local measurements. We can conclude from general considerations that $\mathcal{M}_{1a}^{(\text{MID})} = \mathcal{M}_{2a}^{(\text{MID})} \geq \mathcal{M}_{2b} \geq \mathcal{M}_{1b}^{(\text{WPM})}$. Notice also that the unconditioned local measurements that minimize $H(B|A)$ need not be the same as those that minimize $H(A|B)$. This means that $\mathcal{M}_{2b}$ is intrinsically asymmetric between subsystems A and B even though the measurement strategy is symmetric.

Strategy (c) gives the measure

$$\mathcal{M}_{2c} = \min_{\text{(c)}} H(B|A) - S(B|A) \ . \tag{4.36}$$

The POVM inequality immediately gives a bound on $H(B|A)$,

$$H(B|A) = \sum_a p_a H(B|a) \geq \sum_a p_a S(B|a) \equiv H_{\{E_a\}}(B|A) \ . \tag{4.37}$$

When we are allowed to make conditional measurements on B, the bound can be achieved by measuring B, for outcome a, in the eigenbasis of $\rho_{B|a}$. Hence, with the conditional measurements on B specified, the minimization of the classical conditional entropy, $H(B|A)$, is reduced to choosing a measurement on A that minimizes the conditional entropy $H_{\{E_a\}}(B|A)$:

$$\min_{\text{(c)}} H(B|A) = \min_{\{E_a\}} H_{\{E_a\}}(B|A) \equiv \widetilde{H}(B|A) \ . \tag{4.38}$$

The quantity $\widetilde{H}(B|A)$ is a special sort of classical conditional entropy. The resulting measure is equivalent to the *quantum discord* [Zur00, OZ01] introduced in Chap. 3:

$$\mathcal{M}_{2c} = \widetilde{H}(B|A) - S(B|A) \equiv \mathcal{D}(A \to B) \ . \tag{4.39}$$

In App. A.3, we exhibit joint states that show that to find the minimum $\widetilde{H}(B|A)$—and, hence, to find the quantum discord—sometimes requires rank-one POVMs, not just orthogonal-projection-valued measurements.

Notice that in Chap. 3 we introduced discord as a difference between mutual information-like quantities in accordance with the original definition by Ollivier and Zurek [OZ01]: $\mathcal{D}(A \to B) = S(A : B) - J(A \to B)$. This is equivalent to Eq. (4.39), as one can readily subtract $S(B)$ from $S(A : B)$ and $J(A \to B)$: $(S(A : B) - S(B)) - (J(A \to B) - S(B)) = -S(B|A) + \widetilde{H}(B|A)$. Moreover the definition of discord in Chap. 3 does not make use of conditioned measurements measurements on B, but rather assumed that the quantity to be minimized over measurements on A is the conditional entropy $H_{\{E_a\}}(B|A)$.

We can conclude from general considerations that $\mathcal{M}_{1a}^{(\mathrm{MID})} = \mathcal{M}_{2a}^{(\mathrm{MID})} \geq \mathcal{M}_{2b} \geq \mathcal{M}_{2c}^{(\mathrm{discord})}$. Our present considerations do not, however, provide an ordering of the WPM measure and quantum discord. We return to the ordering of WPM and discord in Sec. 4.2.3 and show in App. A.2 that $\mathcal{M}_{1b}^{(\mathrm{WPM})} \geq \mathcal{M}_{2c}^{(\mathrm{discord})}$.

Type 3: Demon-based measures

Type-3 measures quantify the difference in the work that can be extracted from a quantum system by quantum and classical demons. The demons extract work by transforming the initial joint state ρ_{AB} to the fully mixed joint state using any means at their disposal, including measurements. We assume here that all states of the system have the same energy so that all the work that the demons extract arises from the entropy difference between the initial and final states of the system; it is natural to choose $k_B T \ln 2$ as the unit of work. Throughout this chapter, whenever we talk

about extractable work and erasure cost, we actually mean average work and average erasure cost.

The maximum work that can be extracted by a quantum demon by any means is given by the entropy difference between the initial and final states,

$$W_q = \log(d_A d_B) - S(A,B) \;, \tag{4.40}$$

where d_A and d_B are the dimensions of the two subsystems. The demon could extract this amount of work by devising an optimal process that directly transforms the joint state ρ_{AB} to the maximally mixed state. It could, instead, make a measurement in the joint eigenbasis of ρ_{AB}, extract work $\log(d_A d_B)$ as the post-measurement pure eigenstate is transformed to the maximally mixed state, and then pay a price $S(A,B)$ to erase its memory of the $S(A,B)$ bits acquired in the measurement. The demon would then be ready to pick up another copy of the system and repeat the process.

In contrast to a quantum demon, a local, classical demon can only manipulate the subsystem in its possession. In Sec. 4.1 we introduced two cases for the local demons that are dealing with our bipartite system. In case (i) the two demons are not allowed to communicate with each other. In this case, the maximum amount of work demon A can extract from subsystem A is $\log d_A - S(A)$. This can be achieved by an optimal process that directly transforms the marginal state ρ_A to the maximally mixed state or by measuring in the marginal eigenbasis, extracting work $\log d_A$ as the post-measurement pure state is transformed to the maximally mixed state, and then erasing the $S(A)$ bits of measurement record at cost $S(A)$. Since demon B is in the same situation, the maximum work the two local demons can extract is

$$W_c = \log(d_A d_B) - S(A) - S(B) \;. \tag{4.41}$$

The difference in the amount of work that can be extracted by the quantum and classical demons, called the *work deficit* [OHHH02, Zur03, BT10], is the quantum mutual information:

$$W_q - W_c = S(A) + S(B) - S(A,B) = S(A:B) \equiv \mathcal{M}_{3(\mathrm{i})} \;. \tag{4.42}$$

Brodutch and Terno [BT10] have noted that the work deficit in the case of erasure without communication between the local demons provides an operational interpretation of the quantum mutual information.

In case (ii) the local demons can communicate their measurement results and thus reduce their cost of erasure. In particular, the demons make local measurements, which in accord with the assumptions of this section are described by rank-one POVMs and thus leave the two subsystems in pure states. They can then extract work

$$W^+ = \log d_A + \log d_B \tag{4.43}$$

as their respective systems are transformed to the maximally mixed state. They must then erase their memories of the measurement record so they are ready to handle another copy of the joint state ρ_{AB}.

In the absence of communication, the total erasure cost is $W^- = H(A) + H(B) \geq S(A)+S(B)$, with the minimum attained for measurements in the marginal eigenbases; the net work the demons can extract is that of case (i), i.e., $W_c = W^+ - W^- = \log d_A d_b - S(A) - S(B)$. If the demons can communicate, however, as in case (ii), then they can take advantage of correlations between their measurement results to reduce their erasure cost to the joint classical information in their measurement records, $W^- = H(A, B)$, which gives net work

$$W_c = W^+ - W^- = \log(d_A d_B) - H(A, B) \ . \tag{4.44}$$

Thus in case (ii), the work deficit becomes

$$W_q - W_c = H(A, B) - S(A, B) = \mathcal{M}_3 \ , \tag{4.45}$$

giving us joint-entropy-based measures of nonclassical correlations, with $\mathcal{Q}_3 = -S(A, B)$, $\mathcal{C}_3 = -H(A, B)$, and $\mathcal{M}_3 = \mathcal{Q} - \mathcal{C}$.

We now have to consider the three measurement strategies for the local demons, but before embarking on that, we note that

$$\mathcal{M}_3 = \mathcal{M}_2 + \big[H(A) - S(A)\big] = \mathcal{M}_1 + \big[H(B) - S(B)\big] + \big[H(A) - S(A)\big] \ , \tag{4.46}$$

so for each measurement strategy, we have $\mathcal{M}_3 \geq \mathcal{M}_2 \geq \mathcal{M}_1$, as we have noted earlier.

For strategy (a), measurement in the marginal eigenbases, we have $H(A) = S(A)$ and $H(B) = S(B)$, so we again get the MID measure, i.e., $\mathcal{M}_{3a} = \mathcal{M}_{2a} = \mathcal{M}_{1a}$; this is the form in which Rajagopal and Randall [RR02] defined what they called the quantum deficit. For strategy (b), we have to minimize $H(A, B)$ over all unconditioned local measurements,

$$\mathcal{M}_{3b} = \min_{\text{(b)}} H(A, B) - S(A, B) \; ; \tag{4.47}$$

in general, the result is not the same as $\mathcal{M}_{2b}$ or $\mathcal{M}_{1b}$.

For strategy (c), we have to minimize $H(A, B)$ over all conditioned local measurements. The minimization over the conditioned measurements on B is simple, since as in Eq. (4.37), we have

$$H(A, B) = H(A) + H(B|A) \geq H(A) + \sum_a p_a S(B|a) \equiv H_{\{E_a\}}(A, B) \; , \tag{4.48}$$

with equality if and only if the measurement on B, given outcome a, is in the marginal eigenbasis of $\rho_{B|a}$. Hence, with the conditional measurements on B specified, the minimization of the classical joint entropy, $H(A, B)$, is reduced to choosing a measurement on A that minimizes the joint entropy $H_{\{E_a\}}(A, B)$:

$$\min_{\text{(c)}} H(A, B) = \min_{\{E_a\}} H_{\{E_a\}}(A, B) \equiv \widetilde{H}(A, B) \; . \tag{4.49}$$

The quantity $\widetilde{H}(A, B)$ is a special sort of classical joint entropy. The resulting measure of nonclassical correlations is

$$\mathcal{M}_{3c} = \widetilde{H}(A, B) - S(A, B) \; . \tag{4.50}$$

This measure was hinted at in Zurek's original paper on discord [Zur00]. Ollivier and Zurek [OZ01] defined quantum discord as the quantity $\mathcal{M}_{2c}$, but Zurek [Zur03] resurrected $\mathcal{M}_{3c}$ as a modified form of discord in his paper on discord and Maxwell demons. Brodutch and Terno [BT10] have also pointed out that $\mathcal{M}_{3c}$ is the measure that applies to demons that can communicate and use strategy (c) for their measurements. Hence, we can call $\mathcal{M}_{3c}$ the *demon discord* (dd).

As noted in Sec. 4.1, we have $\mathcal{M}_{3a}^{\text{MID}} \geq \mathcal{M}_{3b} \geq \mathcal{M}_{3c}^{(\text{dd})}$.

Properties of nonclassical-correlation measures

The following array neatly summarizes the measures of nonclassical correlations that we have found and the relations we have found among them:

$$\begin{array}{ccccccccc} & & & & \mathcal{M}_{1a}^{(\text{MID})} & \geq & \mathcal{M}_{1b}^{(\text{WPM})} & & \\ & & & & \| & & \wedge| & & \\ & & & & \mathcal{M}_{2a}^{(\text{MID})} & \geq & \mathcal{M}_{2b} & \geq & \mathcal{M}_{2c}^{(\text{discord})} \\ & & & & \| & & \wedge| & & \wedge| \\ S(A,B) & = & \mathcal{M}_{3(\text{i})} & \geq & \mathcal{M}_{3a}^{(\text{MID})} & \geq & \mathcal{M}_{3b} & \geq & \mathcal{M}_{3c}^{(\text{dd})} \end{array} \tag{4.51}$$

The vertically oriented inequalities are best read by leaning your head to the left; in the absence of leaning, the wedges point toward the smaller quantity, as is standard.

Of the potential measures we started with, the demon-based measure that assumes erasure without communication is special and gives the quantum mutual information. Of the remaining nine potential measures, we discarded one, $\mathcal{M}_{1c}$, as meaningless; we found that the three measures in the left column of the array are all identical to the MID measure; we determined that three of the other measures are the WPM measure, quantum discord, and demon discord; and we are thus left with two new measures, $\mathcal{M}_{2b}$ and $\mathcal{M}_{3b}$, although $\mathcal{M}_{3b}$ is very closely related to—and perhaps identical to—a discord-like measure introduced by Modi *et al.* [MPS$^+$10].

Modi *et al.* [MPS$^+$10] introduced a set of measures of quantum and classical correlations based on the relative-entropy distance (4.2) between a multi-partite state ρ and the nearest state σ_ρ that is diagonal in a product basis, or between ρ and the nearest product state. The only one of these measures relevant to our discussion is their "discord," which when specialized to bipartite states, is the distance $\mathcal{D}_{\text{Modi}} = \min_{\sigma_{AB}} S(\rho_{AB}||\sigma_{AB})$, where σ_{AB} is diagonal in a product basis. Modi *et al.* show that the minimum is attained on a state obtained by projecting ρ_{AB} into a product basis, i.e., $\sigma_{AB} = \sum_{a,b} |e_a, f_b\rangle\langle e_a, f_b|\rho_{AB}|e_a, f_b\rangle\langle e_a, f_b|$, in which case,

$S(\rho_{AB}||\sigma_{AB}) = S(\sigma_{AB}) - S(\rho_{AB})$. Thus we have

$$\mathcal{D}_{\text{Modi}} = \min_{\{|e_a, f_b\rangle\}} S(\sigma_{AB}) - S(\rho_{AB}) \; . \tag{4.52}$$

Since $S(\sigma_{AB})$ is the classical joint entropy of a measurement made on ρ_{AB} in the product basis $|e_a, f_b\rangle$, this would be the same as our $\mathcal{M}_{3b}$ if we knew that the optimal local measurements for $\mathcal{M}_{3b}$ were described by *orthogonal* rank-one projectors.

Brodutch and Terno [BT10] define three kinds of "discord": their D_1 is the standard discord $\mathcal{M}_{2c}^{(\text{discord})}$; their D_2 is the demon discord $\mathcal{M}_{3c}^{(\text{dd})}$; and their D_3 is a discord-like quantity that uses a different conditional measurement strategy. This strategy allows conditioned local measurements, but with the measurement on A constrained to be in the marginal eigenbasis of ρ_A. The Brodutch-Terno measurement strategy is a restriction of strategy (c), and (a) is a restriction of the Brodutch-Terno strategy. Measures based on it could thus be placed in the array (4.51) as an alternative intermediate column whose ordering with strategy (b) is indeterminate.

All the measures in the array (4.51), except the quantum mutual information, are bounded above by MID, and MID is bounded above by the quantum mutual information. Similarly, MID, $\mathcal{M}_{2b}$, and $\mathcal{M}_{3b}$ are bounded below by both the WPM measure and the quantum discord, and the demon discord $\mathcal{M}_{3c}$ is bounded below by discord. The WPM measure and quantum discord have a special status in that they are the most parsimonious of the measures in quantifying nonclassical correlations.

WPM showed that their measure is nonnegative, and Datta [Dat08] showed that discord is nonnegative, allowing us to conclude that all the other measures are also nonnegative. Both proofs rely on the strong subadditivity of quantum entropy [NC00]; we review the proofs in App. A.2. Careful consideration of the conditions for saturating the strong-additivity inequality [HJPW04], not presented here, give the conditions for WPM and discord to be zero: the WPM measure is zero if and only if ρ_{AB} is diagonal in a *product basis*, i.e., an orthonormal basis of the form $|e_a\rangle \otimes |f_b\rangle$, and discord is zero if and only if ρ_{AB} is diagonal in a *conditional product basis* (pointing from A to B), i.e., an orthonormal basis of the form $|e_a\rangle \otimes |f_{b|a}\rangle$.

Since MID, like WPM, is zero if and only if ρ_{AB} is diagonal in a product basis,

the relations in the array (4.51) imply that $\mathcal{M}_{2b}$ and $\mathcal{M}_{3b}$ are zero if and only if ρ_{AB} is diagonal in a product basis. Similarly, the inequality $\mathcal{M}_{3c}^{(\mathrm{dd})} \geq \mathcal{M}_{2c}^{(\mathrm{discord})}$ shows that having ρ_{AB} diagonal in a conditional product basis is necessary to make $\mathcal{M}_{3c}$ zero, and a moment's contemplation of Eqs. (4.48)–(4.50) shows that this is also a sufficient condition.

For pure states, we have $S(A,B) = 0$, $S(A) = S(B) = -S(B|A) = -S(A|B)$, and $S(A:B) = 2S(A) = 2S(B)$. It is easy to show that the optimal measurement for all the measures in the array is measurement in the Schmidt basis of the pure state (marginal eigenbasis for each subsystem), which gives $H(A) = H(B) = H(A,B) = H(A:B) = S(A) = S(B)$ and $H(B|A) = H(A|B) = 0$. Thus all the measures in the array, except the quantum mutual information, are equal to the marginal quantum entropy, $S(A) = S(B)$, which is the entropic measure of entanglement for bipartite pure states.

The remaining gap in our understanding left by the relations in the array is whether there is an inequality between the WPM measure and discord. The WPM measure is strictly bigger than zero for states that are diagonal in a conditional product basis that is not a product basis and so is bigger than the quantum discord for such states. If there is an inequality, it must be that the WPM measure is bounded below by quantum discord. Indeed, it is not hard to come up with a proof, using the method of Piani *et al.* [PHH08]. The proof, given in App. A.2, is part of the two-step demonstration that WPM and discord are nonnegative. We conclude that

$$\mathcal{M}_{1b}^{(\mathrm{WPM})} \geq \mathcal{M}_{2c}^{(\mathrm{discord})} = \mathcal{D}(A \to B) \,. \tag{4.53}$$

The proof allows us to identify the equality condition: the WPM measure is equal to discord if and only if ρ_{AB} is diagonal in a conditional product basis that points from B to A.

We emphasize that product bases and conditional product bases do not exhaust the set of orthonormal bases that are made up of product states. There are orthonormal bases made up entirely of product states that are neither product bases nor conditional product bases; these have been studied, for example, in the context of nonlocality

without entanglement [BDF+99]. Not surprisingly, we refer to such a basis as a basis of product states, to be distinguished from a product basis or a conditional product basis.

4.2.4 Rank-one POVMs and projective measurements

In Sec. 4.2.3 we assumed that all the measurements were described by rank-one POVMs. This assumption does not affect the demon-based work deficit (4.42) for the case of erasure without communication, for that case, which leads to the quantum mutual information, does not rely on any assumptions about how the subsystems are measured. Nor does this assumption affect MID, which is derived from measurement strategy (a), a strategy that from the outset prescribes orthogonal-projection-valued measurements in the eigenbases of the marginal density operators. The assumption must be carefully examined, however, for measurement strategies (b) and (c). On the face of it, there is a problem for the second and third rows of our array. For type-2 measures, the task is to minimize a classical conditional entropy, and for type-3 measures, the task is to minimize a classical joint entropy. In both cases, the minimum is achieved by making no measurements at all.

For the demon-based measures in the right two columns of the third row, it is clear what the problem is. The contribution of $H(A, B)$ to the classical work comes from the erasure cost; the local demons can minimize their erasure cost by not having a measurement record. Of course, if the local demons make no measurements, they also cannot extract the work attendant on knowing more about their system's state. The upshot is that formula (4.44) for the net classical work needs to be modified if one does not assume measurements described by rank-one POVMs. App. A.4 shows, not surprisingly, that, once modified, the net classical work is always optimized on rank-one POVMs, so one can restrict the demons in this way without affecting their performance.

For the measure $\mathcal{M}_{2b}$, we know of no reason to restrict to rank-one POVMs more compelling than declaring that the measure would be nonsense without this restriction.

For the quantum discord, we can do better: the original definition of discord [OZ01] did not discuss conditioned measurements on B, but rather formulated the discord directly in terms of minimizing the classical conditional entropy as in Eqs. (4.37)–(4.39); this is equivalent to our assuming rank-one POVMs for the measurement on B. We are still left with a question—why should the measurements on A be restricted to rank-one POVMs?—and this same question applies to both local measurements for the WPM measure. We now address this question by showing in both situations that the optimum can always be attained on rank-one POVMs. It is important to show this, because the proofs regarding nonnegativity and ordering of the WPM measure and discord, given in App. A.2, assume rank-one POVMs.

We deal with the WPM measure first. The key point is obvious: making coarse-grained POVM measurements on A and B should not uncover as much mutual information as making fine-grained, rank-one POVM measurements. We start with POVMs $\{E_a\}$ and $\{F_b\}$ for systems A and B, and we imagine that these are a coarse graining of POVMs $\{E_{aj}\}$ and $\{F_{bk}\}$, i.e.,

$$E_a = \sum_j E_{aj} \ , \qquad F_b = \sum_k F_{bk} \ . \tag{4.54}$$

A POVM element can always be fine-grained to the rank-one level by writing it in terms of its eigendecomposition. The joint probability for the fine-grained outcomes aj and bk is $p_{ajbk} = p_{jk|ab}p_{ab}$, with similar relations for the marginals for the two subsystems. It is now trivial to show that fine graining never decreases the classical mutual information:

$$H(A, J : B, K) = H(A : B) + \sum_{a,b} p_{ab} H(J : K|a, b) \ . \tag{4.55}$$

This means that in maximizing the classical mutual information, we need only consider rank-one POVMs.

For the discord, the reduction to rank-one POVMs has been demonstrated by Datta [Dat08]; it is sufficiently brief that we repeat it here. Since the conditional measurements on B are already specified, we need only worry about fine graining the measurement on A. We need the conditional state of B given the coarse-grained

outcome a in terms of the conditional states given the fine-grained outcome aj:

$$\rho_{B|a} = \frac{\mathrm{tr}_A(E_a \rho_{AB})}{p_a} = \sum_j \frac{\mathrm{tr}_A(E_{aj} \rho_{AB})}{p_a} = \sum_j p_{j|a} \rho_{B|aj} \ , \tag{4.56}$$

where

$$p_{j|a} = \frac{p_{aj}}{p_a} = \frac{\mathrm{tr}_A(E_{aj} \rho_A)}{\mathrm{tr}_A(E_a \rho_A)} \ . \tag{4.57}$$

The quantity to be minimized over measurements on A is the conditional entropy (4.37). For it, we can write

$$\begin{aligned} H_{\{E_a\}}(B|A) &= \sum_a p_a S(\rho_{B|a}) \\ &= \sum_a p_a S\Big(\sum_j p_{j|a} \rho_{B|aj} \Big) \qquad (4.58) \\ &\geq \sum_{a,j} p_{aj} S(\rho_{B|aj}) = H_{\{E_{aj}\}}(B|A) \ , \qquad (4.59) \end{aligned}$$

where the inequality follows from the concavity of the von Neumann entropy. Thus fine graining never increases this conditional entropy, so we are assured that the minimum is attained on rank-one POVMs.

We have now settled the question of restricting to rank-one POVMs for all the measures except the measure in the middle, $\mathcal{M}_{2b}$, and for it, we simply assert that it makes sense only if we restrict to rank-one POVMs. A remaining question is whether we can further restrict to orthogonal-projection-valued measurements. Searching over the entire set of rank-one POVMs is a daunting task, considerably more onerous than searching just over projection-valued measurements. On this score, we can report that WPM drew attention to an example where the WPM measurements are optimized on a rank-one POVM that is not projection-valued; we extend this example to quantum discord and generalize it in App. A.3.

These examples, however, require that at least one system have dimension bigger than two; For evaluating quantum discord for a two-qubit system, Chen *et al.* [CZY+11] found, that there are some states for which three-element POVMs on system A do better than two-outcome, orthogonal-projection-valued measurements.

Exploring the situation numerically, Galve, Giorgi, and Zambrini [GGZ11] confirmed this finding, but suggested that the corrections to the two-qubit discord obtained by using POVMs, instead of orthogonal projectors, are negligible. In the next section, we do a wholesale evaluation of the various measures for two-qubit states; in the need for manageable numerics, we restrict the search over measurements to rank-one, orthogonal projection operators; according to [GGZ11], this should have no significant effect on the result.

Several groups of investigators have considered Gaussian versions of nonclassical-correlation measures for Gaussian states of two harmonic-oscillator modes; the local measurements are restricted to Gaussian measurements, i.e., measurements whose POVM elements are the phase-space displacements of a particular single-mode Gaussian state. Giorda and Paris [GP10] and Addesso and Datta [AD10] focused on a Gaussian version of discord and showed that the optimal Gaussian measurements are rank-one POVMs, but that for some Gaussian states, the optimal measurement is not orthogonal-projection-valued and thus requires the use of POVMs. Mišta *et al.* [MTG+11] investigated Gaussian versions of MID and WPM (which they called AmeriolatedMID); their investigation showed that for some Gaussian states, the optimal Gaussian measurement for the WPM measure is not the globally optimal measurement when one allows nonGaussian POVMs.

4.3 Numerical results for two-qubit states

One purpose of our framework is to clarify relations among the various measures of nonclassical correlations beyond entanglement. The ordering of the measures is of particular interest. The framework provides by construction some ordering relations between the measures; in addition, we have proved, using the method of Piani *et al.*, the important relation that the WPM measure is bounded below by the discord. Nonetheless, questions remain, in particular, of whether there is an ordering between $\mathcal{M}_{2b}$ and $\mathcal{M}_{3c}^{(\mathrm{dd})}$, as well as between $\mathcal{M}_{1b}^{(\mathrm{WPM})}$ and $\mathcal{M}_{3c}^{(\mathrm{dd})}$.

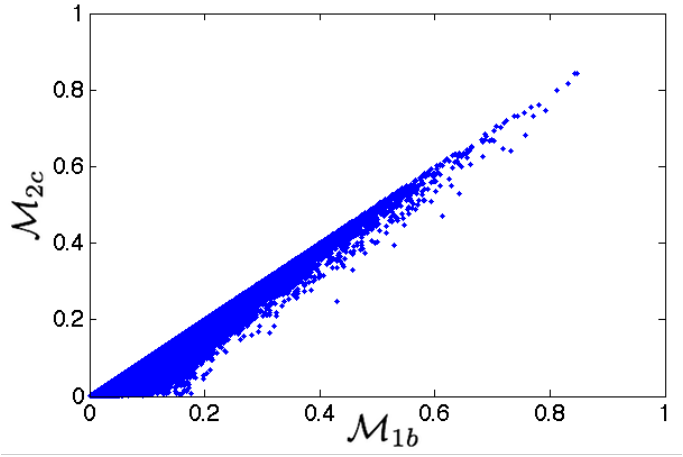

Figure 4.1: $\mathcal{M}_{2c}^{(\text{discord})} = \mathcal{D}(A \to B)$ plotted against $\mathcal{M}_{1b}^{(\text{WPM})}$ for one million randomly generated joint density matrices, using orthogonal projectors for the measurements. As expected, the WPM measure is never smaller than the discord; also evident is that discord is zero for a larger class of states than the WPM measure, those being the states that are diagonal in a conditional product basis pointing from A to B.

In this section we illustrate and investigate the various orderings by presenting numerical evaluations of the several measures for randomly selected two-qubit states. It should be noted, however, that in order to do the optimizations over measurements

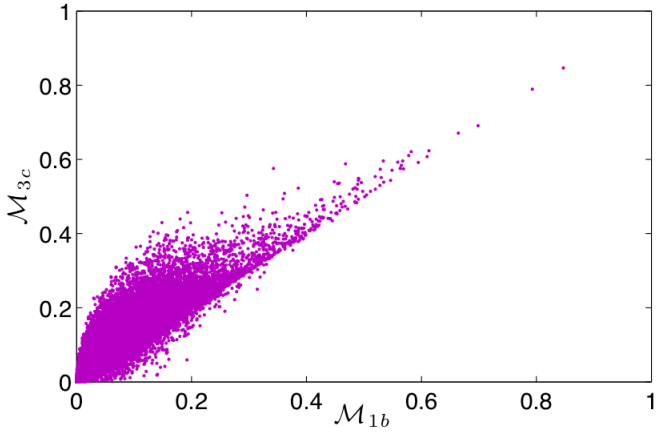

Figure 4.2: $\mathcal{M}_{3c}^{(\text{dd})}$ plotted against $\mathcal{M}_{1b}^{(\text{WPM})}$ for 100,000 randomly generated joint density matrices. Since $\mathcal{M}_{3c}^{(\text{dd})} \geq \mathcal{M}_{2c}^{(\text{discord})}$, the points from Fig. 4.1 move upwards. Many points pass the diagonal, and the ordering of Fig. 4.1 disappears.

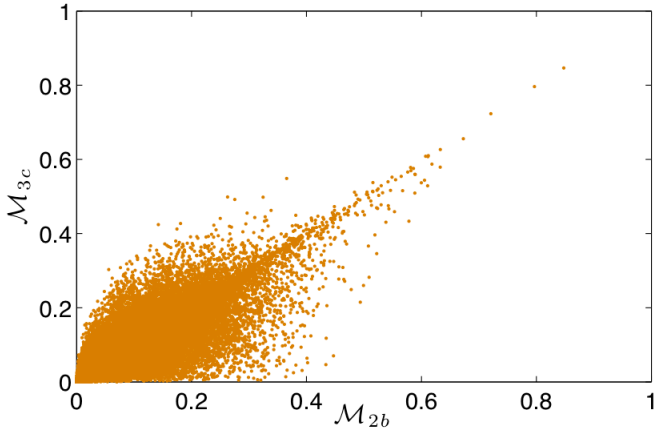

Figure 4.3: $\mathcal{M}_{3c}^{(\mathrm{dd})}$ plotted against $\mathcal{M}_{2b}$ for 100,000 randomly generated joint density matrices. Relative to Fig. 4.2, the points move right, due to the relation $\mathcal{M}_{2b} \geq \mathcal{M}_{1b}$. Since not all of them pass the diagonal, there is no ordering relation between $\mathcal{M}_{3c}$ and $\mathcal{M}_{2b}$.

numerically, we have had to restrict ourselves to orthogonal projectors instead of the more general POVMs, so in some situations, we might not be finding the optimal measurements.

To calculate the various correlation measures, we follow the approach of Al-Qasimi and James [AJ11]. The measurement operators E_a and F_b are orthogonal projectors,

$$E_a = |e_a^A\rangle\langle e_a^A| \,, \qquad F_b = |e_b^B\rangle\langle e_b^B| \,, \qquad a, b \in \{0, 1\} \,, \tag{4.60}$$

$$|e_0^X\rangle = \cos\theta^X|0\rangle + e^{i\phi^X} \sin\theta^X|1\rangle \,, \quad |e_1^X\rangle = -\sin\theta^X|0\rangle + e^{i\phi^X} \cos\theta^X|1\rangle \,. \tag{4.61}$$

The required optimization is done by a numerical search over the angles $\{\theta^X, \phi^X\}$ for $X \in \{A, B\}$. For measurement strategy (b), we must search over the four angles for both qubits, but for strategy (c), we need only search over the two angles for subsystem A.

Figure 4.1 compares the WPM measure and discord, confirming the expectation that the WPM measure is never smaller than discord. Figures 4.2 and 4.3 display the aforementioned pairs of correlation measures where our framework does not imply an ordering relation; the numerical data show that there is no ordering for these pairs.

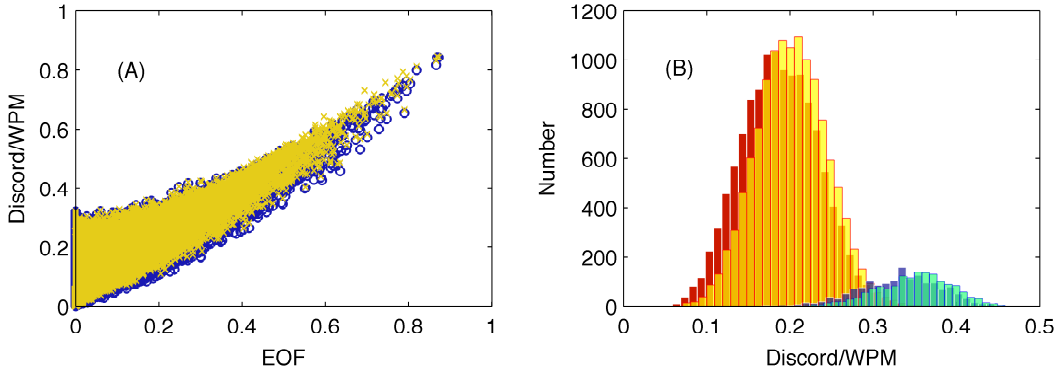

Figure 4.4: (A) Discord (blue circles) and the WPM measure (yellow crosses) for one million randomly chosen two-qubit states, plotted against entanglement of formation, E_f. As the correlations increase, the spread between entanglement and WPM or discord decreases. (B) Two superimposed histograms showing the distribution of discord and the WPM measure for ranges of values of E_f: left histogram shows discord (red) and WPM (yellow) for the states of (A) corresponding to $0.1 \leq E_f \leq 0.2$; right histogram shows discord (blue) and WPM (green) corresponding to $0.3 \leq E_f \leq 0.4$.

Another relation we have explored numerically is the one between the correlation measures and entanglement. Figure 4.4 shows discord and the WPM measure plotted against entanglement of formation, reproducing the plot in [AJ11] for discord, but providing new data for the WPM measure. The entanglement of formation is calculated using Wootters's analytical expression [Woo98], $E_f(\rho) = h\big((1+\sqrt{1-\mathcal{C}^2(\rho)})/2\big)$. Here $h(x)$ is the binary entropy, $h(x) = -x\log x - (1-x)\log(1-x)$, and $C(\rho)$ is the concurrence, given by $\mathcal{C}(\rho) = \max(0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4)$, where the λ_js are the eigenvalues in decreasing order of the operator $\sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}}$, with $\tilde{\rho} = (\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y)$.

To avoid the slow numerical optimization procedures, analytical expressions for the correlation measures would be desirable. Yet only for very restricted classes of joint states are such expressions available. Girolami *et al.* [GPA10] [1] suggested that there is an analytical expression for the WPM measure for general two-qubit states. To understand their claim, we write the joint two-qubit state in terms of Pauli

[1] In the published version of their paper Girolami *et al.* restriced the validity of their analytical expression to the set of two-qubit X-states [GPA11].

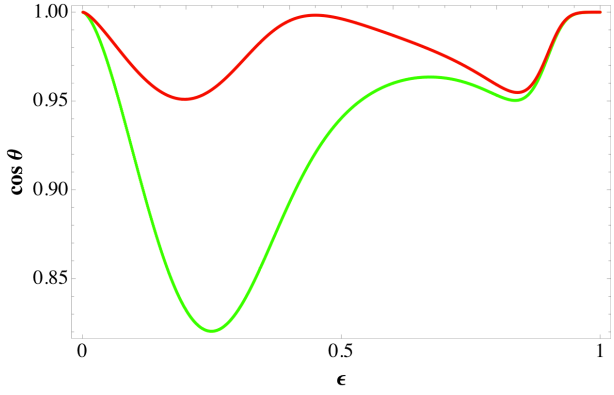

Figure 4.5: Deviation of the numerically obtained, optimal measurement vectors from the maximal singular vectors of the correlation matrix for the WPM measure. The joint states are a mixture of a pure product state with marginal spin (Bloch) vectors $\boldsymbol{a} = (1,\, 0,\, 0)$ and $\boldsymbol{b} = (1/\sqrt{2},\, -1/2,\, 1/2)$ and a mixed Bell-diagonal (zero marginal spin vectors) state, with correlation matrix $c = \mathrm{diag}(-0.9,\, -0.8,\, -0.7)$. The mixing parameter is ϵ, with $\epsilon = 0$ corresponding to the product state and $\epsilon = 1$ to the Bell-diagonal state. The green curve shows the cosine of the angle between the maximal right singular vector and the measurement vector on system B. The red curve is the cosine of the angle between maximal left singular vector and the measurement vector on system A.

operators:

$$\rho_{AB} = \frac{1}{4}\Big(I_{AB} + \boldsymbol{a}\cdot\boldsymbol{\sigma}^A \otimes I_B + I_A \otimes \boldsymbol{b}\cdot\boldsymbol{\sigma}^B + \sum_{j,k=1}^{3} c_{jk}\sigma_j^A \otimes \sigma_k^B\Big)\,. \tag{4.62}$$

The correlation matrix, $c_{jk} = \mathrm{tr}(\sigma_j^A \otimes \sigma_k^B \rho_{AB})$, is not symmetric, but $c^T c$ can be diagonalized as $c^T c \boldsymbol{m_j} = \lambda_j^2 \boldsymbol{m}_j$. The eigenvalues, λ_j^2, are the squares of the singular values of c, and the eigenvectors are the right singular vectors of c. The correlation matrix maps the right singular vectors to the left singular vectors, $c\boldsymbol{m}_j = \lambda_j \boldsymbol{n}_j$, and the left singular vectors, $\boldsymbol{n}_j$, are the eigenvectors of cc^T.

The claim of Girolami *et al.* was that the maximal left and right singular vectors, i.e., those corresponding to the largest singular value of c, specify the optimal measurements for the WPM measure. We can confirm that the measurement vectors for generic (randomly generated) two-qubit states are close to the maximal singular vectors of c for all three correlation measures that are based on measurement strategy (b), but it can be shown analytically that in general the singular vectors

are not the optimal measurement vectors. Moreover, there are examples where the deviation becomes obvious in the numerics. Figure 4.5 shows an example where the angle between the measurement vectors and the maximal singular vectors is noticeable in the calculation of the WPM measure. Similar plots can be obtained for the measures $\mathcal{M}_{2b}$ and $\mathcal{M}_{3b}$.

In this chapter we considered several entropic measures of nonclassical correlations. We presented a framework from which these different measures emerge logically. This enabled us to investigate ordering relations between these measures, by putting them onto equal footing. While some of these ordering relations are a direct consequence of the different measurement strategies employed in the definition of the framework, others were either proven analytically or numerical evidence was presented demonstrating the nonexistence of an ordering between some measures. In the next chapter we investigate some properties of these measures further by using a pictorial approach for a restricted set of two qubit states.

Chapter 5

Quantum Discord and the Geometry of Bell-Diagonal States

Maintenance of quantum coherence is clearly important for quantum-information-processing protocols. Noise and decoherence, by turning pure states into mixed states, generally destroy quantum coherence. Efficient representation of quantum information requires that a quantum-information-processing system be composed of parts [BKCD02]. For multi-partite systems, quantum coherence is related to nonclassical correlations between the parts.

One can use decoherence mechanisms to explore the nooks and crannies of nonclassical-correlation measures. There is no sudden death of discord [FAC$^+$10], as is suggested by the absence of open sets of classical states, but the nonanalyticity of nonclassical measures points to the possibility of sudden changes in derivatives. Investigation of the behavior of nonclassical measures under decoherence has begun [Col10, MCSV09, MWF$^+$10, MPM10], with a focus on the action of decoherence within the class of two-qubit states that are diagonal in the Bell basis. This focus is motivated by the fact that entanglement measures and nonclassical-correlation measures can be calculated explicitly for the Bell-diagonal states, thus allowing one to determine how these measures change under decoherence.

The Bell-diagonal states are a three-parameter set, whose geometry, including the subsets of separable and classical subsets, can be depicted in three dimensions [HHHH09, HH96]. Level surfaces of entanglement and nonclassical measures can be plotted directly on this three-dimensional geometry. The result is a complete picture, for this simple case, of the structure of entanglement and nonclassicality. We suggest that it is more illuminating to use this picture to explain how measures of entanglement and nonclassicality change along the one-dimensional trajectories traced out by decohering states, rather than the other way around. Hence we review and expand the pictorial approach here.

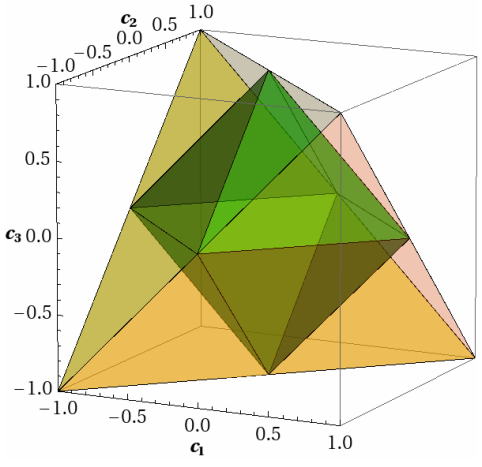

Figure 5.1: Geometry of Bell-diagonal states. The tetrahedron $\mathcal{T}$ is the set of valid Bell-diagonal states. The Bell states $|\beta_{ab}\rangle$ sit at the four vertices, the extreme points of $\mathcal{T}$. The green octahedron $\mathcal{O}$, specified by $|c_1|+|c_2|+|c_3| \leq 1$ ($\lambda_{ab} \leq 1/2$), is the set of separable Bell-diagonal states. There are four entangled regions outside $\mathcal{O}$, one for each vertex of $\mathcal{T}$, in each of which the biggest eigenvalue λ_{ab} is the one associated with the Bell state at the vertex. Classical states, i.e., those diagonal in a product basis, lie on the Cartesian axes.

In Sec. 2.3 we have seen that an arbitrary two-qubit state can be written as:

$$\rho = \frac{1}{4}\Big(I_4 + \vec{a}\cdot\vec{\sigma}\otimes I_2 + I_2\otimes\vec{b}\cdot\vec{\sigma} + \sum_{i,j} c_{i,j}\,\sigma_i\otimes\sigma_j\Big). \tag{5.1}$$

Any two-qubit state, satisfying $\vec{a} = 0 = \vec{b}$, i.e., having maximally mixed marginal density operators $\rho_A = I/2 = \rho_B$, can be brought to Bell-diagonal form by using

local unitary operations on the two qubits to diagonalize the correlation matrix $c_{i,j} = \langle \sigma_j^A \otimes \sigma_k^B \rangle$ [Luo08b].

Bell-diagonal states of two qubits, A and B, consequently have a density operators of the form

$$\rho_{AB} = \frac{1}{4}\left(I + \sum_{j=1}^{3} c_j\, \sigma_j^A \otimes \sigma_j^B\right) = \sum_{a,b} \lambda_{ab} |\beta_{ab}\rangle\langle\beta_{ab}| \,. \tag{5.2}$$

The eigenstates are the four Bell states $|\beta_{ab}\rangle \equiv (|0,b\rangle + (-1)^a |1, 1 \oplus b\rangle)/\sqrt{2}$, with eigenvalues

$$\lambda_{ab} = \frac{1}{4}\Big(1 + (-1)^a c_1 - (-1)^{a+b} c_2 + (-1)^b c_3\Big) \,. \tag{5.3}$$

A Bell-diagonal state is specified by a 3-tuple (c_1, c_2, c_3). The density operator ρ_{AB} must be a positive operator, i.e., $\lambda_{ab} \geq 0$; the resulting region of Bell-diagonal states is the state tetrahedron $\mathcal{T}$ in Fig. 5.1. Separable Bell-diagonal states are those with positive partial transpose [HHHH09]. Partial transposition changes the sign of c_2, so operators with positive partial transpose occupy the reflection of $\mathcal{T}$ through the plane $c_2 = 0$; the region of separable Bell-diagonal states is the intersection of the two tetrahedra, which is the octahedron $\mathcal{O}$ of Fig. 5.1 [HH96].

The entanglement of formation $\mathcal{E}$ [HHHH09, Woo98] is a monotonically increasing function of Wootters's concurrence C [Woo98], which for Bell-diagonal states, is given by $C = \max(0, 2\lambda_{\max} - 1)$, where $\lambda_{\max} = \max \lambda_{ab}$. The concurrence and the entanglement of formation are convex functions on $\mathcal{T}$. They are zero for the separable states in the octahedron $\mathcal{O}$. In each of the four entangled regions outside $\mathcal{O}$, C and $\mathcal{E}$ are constant on planes parallel to the bounding face of $\mathcal{O}$ and increase as one moves outward through these planes toward the Bell-state vertex.

We will talk about discord when investigating correlation measures in this setting because it has been a focus of recent work on decoherence and nonclassical correlations [Col10, MCSV09, MWF$^+$10, MPM10]. As we will show, however, everything we will talk about directly applies to all of the correlation measures discussed in the previous chapter, since for the set of Bell-diagonal states, all of these measures are equal.

Discord is defined as the difference between the mutual information $\mathcal{I}$ and the accessible classical correlation $\mathcal{C}$ [HV01],

$$\mathcal{D} = \mathcal{I} - \mathcal{C} = \tilde{S}(B|A) - S(B|A) \ . \tag{5.4}$$

The quantum mutual information $\mathcal{I}$ is regarded as quantifying the *total* correlations in the joint state ρ_{AB}.

The quantum mutual information of Bell-diagonal states,

$$\mathcal{I} = 2 - S(\rho_{AB}) = \sum_{a,b} \lambda_{ab} \log_2(4\lambda_{ab}) \ , \tag{5.5}$$

is a convex function on $\mathcal{T}$. It has smooth level surfaces that bulge outward toward the vertices of $\mathcal{T}$ (see Fig. 5.2).

The next step is to quantify *purely classical* correlations $\mathcal{C}$ in terms of information from measurements. As we restricted ourselves to Bell-diagonal states, we can make use of an analytical expression for $\mathcal{C}$, derived by Luo [Luo08b], and circumvent the cumbersome optimization procedure:

$$\mathcal{C} = 1 - H_2\left(\frac{1+c}{2}\right) = \frac{1+c}{2}\log_2(1+c) + \frac{1-c}{2}\log_2(1-c) \ , \tag{5.6}$$

where $H_2(p) = -p\log_2 p - (1-p)\log_2(1-p)$ is the binary entropy and $c = \max|c_j|$.

Luo calculated $\mathcal{C}$ for Bell-diagonal states under the assumption that the measurement is described by one-dimensional orthogonal projection operators. We will relax this assumption and extend Luo's result to rank-one POVMs here. If the measurement on system A is to be described by arbitrary rank-one POVM elements

$$E_k = q_k(I + \boldsymbol{n}_k \cdot \boldsymbol{\sigma}), \tag{5.7}$$

we have $p_k = q_k$ and the state $\rho_{B|k}$ of system B after the measurement outcome k is

$$\rho_{B|k} = (I + \boldsymbol{d_k} \cdot \boldsymbol{\sigma})/2, \tag{5.8}$$

with $d_{kj} = c_j n_{kj}$. Now, we have

$$S(\rho_{B|k}) = H_2[(1 + |\boldsymbol{d_k}|)/2 \geq H_2[(1+c)/2], \tag{5.9}$$

since $|\boldsymbol{d}_k| \leq c$. This shows that $\tilde{S}(B|A) \geq H_2[(1+c)/2]$, with equality for measurement of orthogonal projectors along the direction of maximum c_j.

From Eqs. (5.8) and (5.9) we see that the optimal measurement on system B is always aligned with the measurement on system A and does not depend on the outcome of the first measurement. As both of the optimal measurements are specified by orthogonal projectors, and the marginal density operators for either subsystem of Bell-diagonal state are degenerate, the measurements are performed in the marginal eigenbasis of both parts. This causes the hierarchy of correlation measures in Eq. (4.51) to collapse and demonstrates, for the set of Bell diagonal states, the equality of all the correlation measures considered in the previous chapter.

The accessible classical information $\mathcal{C}$, a convex function on $\mathcal{T}$, is constant on the surfaces of cubes (or the portion of such a cube in $\mathcal{T}$) centered at the origin in Fig. 5.2 —this introduces nonanalyticity—and $\mathcal{C}$ increases monotonically with the size of the cube.

As we have shown that discord equals the symmetric correlation measures, due to the symmetry of the Bell-diagonal states, it vanishes if and only if ρ_{AB} is diagonal in a product basis $|e_j^A\rangle \otimes |f_k^B\rangle$. In our three dimensional picture of the Bell-diagonal state space, these purely classical states fall on the Cartesian axes [DacVB10] (see Fig. 5.1).

Figure 5.3 plots level surfaces of discord for Bell-diagonal states. From these plots, it is clear that discord is quite a different beast from entanglement of formation, quantum mutual information, and the measure of classical correlations. Whereas $\mathcal{E}$, $\mathcal{I}$, and $\mathcal{C}$ generally increase outward from the origin, $\mathcal{D}$ increases away from the Cartesian axes, capturing an entropic notion of distance from classical states [MPS$^+$10, DacVB10]. In particular, as one moves outward along one of the constant-discord tubes of Fig. 5.3, the classical correlations and the total correlations of the quantum mutual information increase, but their difference, the nonclassical correlations as measured by discord, remains constant. At the vertices of $\mathcal{O}$, $\mathcal{I} = \mathcal{C} = 1$ and $\mathcal{D} = \mathcal{E} = 0$. At the Bell-state vertices of $\mathcal{T}$, $\mathcal{I} = 2$ and $\mathcal{C} = \mathcal{D} = \mathcal{E} = 1$, this being the

maximum value of discord for two qubits. In addition, $\mathcal{E}$, $\mathcal{I}$, and $\mathcal{C}$ are all convex, whereas discord is neither concave nor convex, as is evident from the plots in Fig. 5.3: one can mix two positive-discord states to get a zero-discord classical state, and one can mix two zero-discord classical states on different axes to get a positive-discord state. This argument and its conclusion are not special to Bell-diagonal states: mixing two discordant states can lead to a state that is diagonal in a product basis, and mixing two states that are diagonal in incompatible product bases generally leads to a discordant state.

Mazzola, Piilo, and Maniscalco [MPM10] recently investigated the dynamics of classical and nonclassical correlations, as measured by discord, for two qubits under decoherence processes that preserve Bell-diagonal states. In particular, they considered independent phase-flip channels for the two qubits. The phase flips are implemented mathematically by random applications of σ_z operators to the qubits. This decoherence process leaves c_3 unchanged, but flips the signs of c_1 and c_2 randomly, leading to exponential decay of c_1 and c_2 at the same rate. Mazzola and collaborators found that for the initial conditions they considered, the entanglement of formation decays to zero in a finite time—sudden death of entanglement [YE09]—but that the discord remains constant for a finite time and then decays, reaching zero at infinite time. This situation is depicted in terms of the surfaces of constant discord in Figure 5.4. The decohering-state trajectory is a straight line that runs along a tube of constant discord, until it encounters an intersecting tube, after which the discord decreases to zero when the state becomes fully classical.

This behavior is generic for flip channels and initial conditions on edges of the state tetrahedron. We focus here on the phase-flip channel with initial conditions in the $(+,-,+)$-octant, but analogous considerations apply to the other flip channels (bit and bit-phase) and to initial conditions on the other edges of $\mathcal{T}$. Consider then initial conditions anywhere along the edge of $\mathcal{T}$ in this octant: $c_1(0) = 1$ and $0 \leq -c_2(0) = c_3(0) \leq 1$. The trajectory under phase flips is a straight line $c_3 = c_3(0) = -c_2/c_1$. Along this straight line, the eigenvalues λ_{ab} factor into products of probabilities, $(1 \pm c_1)/2$ and $(1 \pm c_3)/2$, thus making $S(\rho_{AB})$ the entropy of two

independent binary random variables with these probabilities. This yields a quantum mutual information $\mathcal{I} = 2 - H_2[(1+c_3)/2] - H_2[(1+c_1)/2]$. Furthermore, along the trajectory $c = \max(c_1, c_3)$. The result is that the trajectory initially runs along a tube of constant discord

$$\mathcal{D} = 1 - H_2\left(\frac{1+c_3}{2}\right), \tag{5.10}$$

for $c_1 \geq c_3$. When $c_1 = c_3$, the trajectory encounters another tube, after which, for $c_1 \leq c_3$, the discord decreases monotonically as $\mathcal{D} = 1 - H_2[(1+c_1)/2]$ as c_1 decreases. Meanwhile, the entanglement of formation decreases monotonically from its initial value to a sudden death at $c_1 = (1-c_3)/(1+c_3)$.

The situation investigated in [MPM10] is surely interesting: under decoherence, nonclassical correlations remain constant for a finite time interval. This situation is, however, a special one, as can be seen from the surfaces of constant discord; the trajectories considered here are the only straight lines in parameter space that stay on a surface of constant discord. Indeed, the pictorial approach can provide a complete understanding of how entanglement and nonclassicality change under decoherence within the set of Bell-diagonal states.

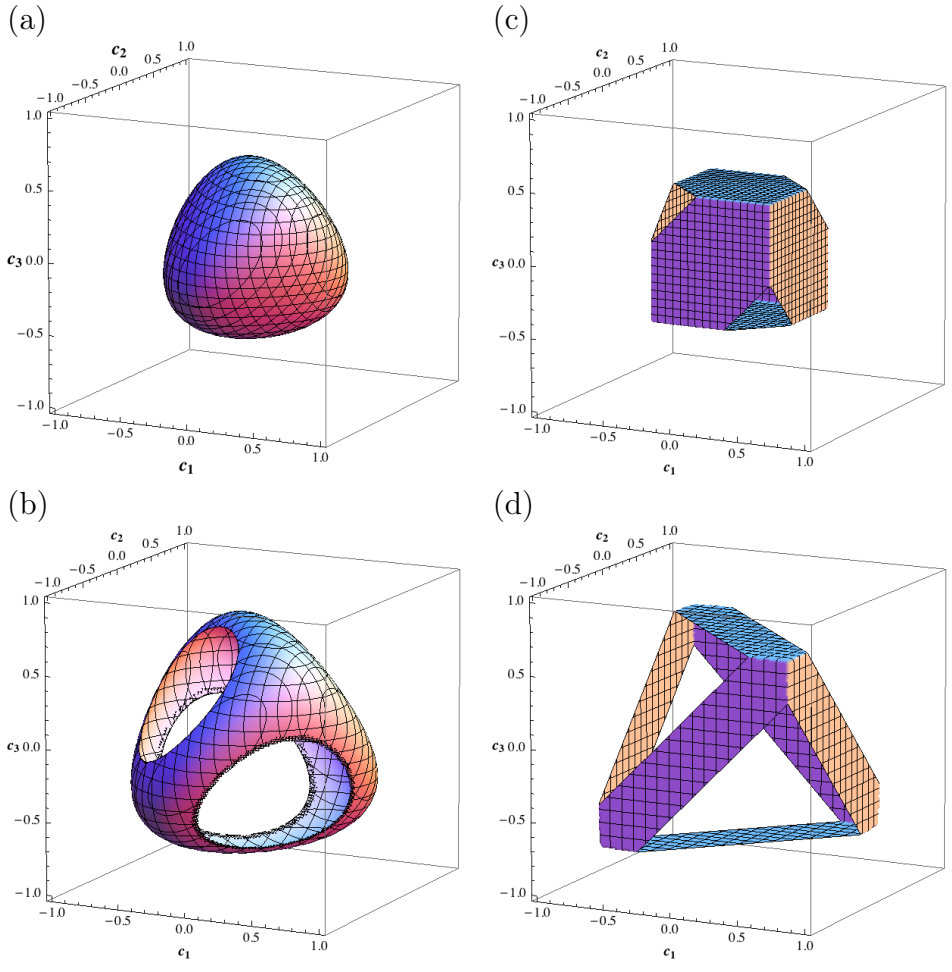

Figure 5.2: Level surfaces of the quantum mutual information $\mathcal{I}$ (left column) and the accessible classical information $\mathcal{C}$ (right column): (a) $\mathcal{I} = 0.1$; (b) $\mathcal{I} = 0.55$; (c) $\mathcal{C} = 0.15$; (d) $\mathcal{C} = 0.4$. The smooth surface of $calI = 0.1$ bulges towards the vertices of the tetrahedron $\mathcal{T}$ from Fig. 5.1; as the mutual information grows, this surface becomes inflated and eventually intersects $\mathcal{T}$, giving rise to the windows seen in (b) for $\mathcal{I} = 0.55$. The level surfaces for the accessible classical information $\mathcal{C}$ are cubes centered at the origin. As $\mathcal{C}$ increases corners of the cube get cut of as they poke through the surface of $\mathcal{T}$.

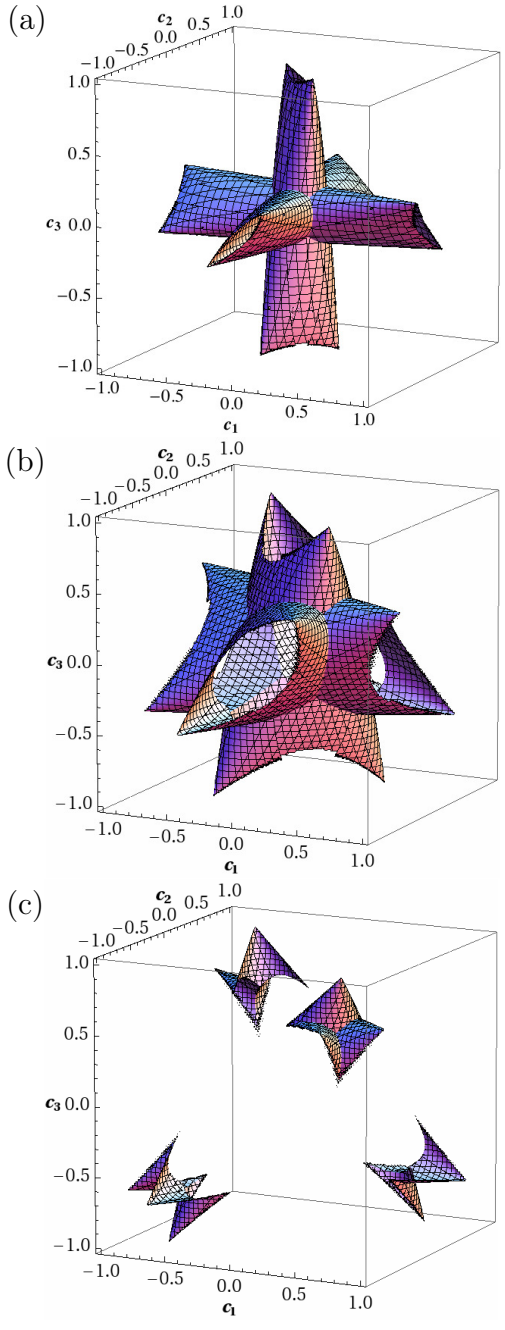

Figure 5.3: Surfaces of constant discord: (a) $\mathcal{D} = 0.03$; (b) $\mathcal{D} = 0.15$; (c) $\mathcal{D} = 0.35$. The level surfaces consist of three intersecting "tubes" running along the three Cartesian axes. The tubes are cut off by the state tetrahedron $\mathcal{T}$ at their ends, and they are squeezed and twisted so that at their ends, they align with an edge of $\mathcal{T}$. As discord decreases, the tubes collapse to the Cartesian axes [DacVB10]. As discord increases, the tube structure is obscured, as in (c): the main body of each tube is cut off by $\mathcal{T}$; all that remains are the tips, which reach out toward the Bell-state vertices.

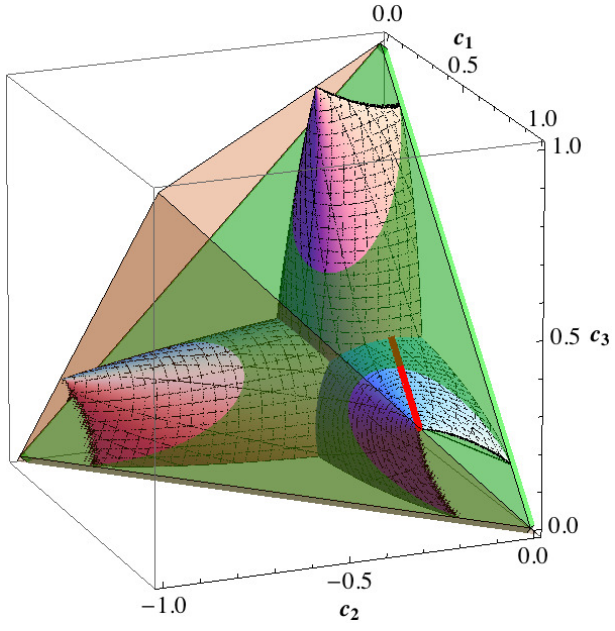

Figure 5.4: Trajectory (red) of a Bell-diagonal state under random phase flips of the two qubits; initial conditions are $c_1(0) = 1$, $-c_2(0) = c_3(0) = 0.3$. The trajectory is the straight line $c_3 = c_3(0) = 0.3 = -c_2/c_1$. For clarity, only the $(+,-,+)$-octant is shown. A constant-discord surface is plotted for the discord value of the initial state. Faces of the yellow state tetrahedron $\mathcal{T}$ and the green separable octahedron $\mathcal{O}$ are also shown. The straight-line trajectory proceeds along a tube of constant discord till it encounters the vertical tube at $c_1 = 0.3$, after which discord decreases monotonically to zero when the trajectory reaches the c_3 axis. Entanglement of formation decreases monotonically to zero when the trajectory enters $\mathcal{O}$ at $c_1 = 0.7/1.3 = 0.54$.

Chapter 6

Parameter estimation

The goal of metrology is to determine some quantity as accurately as possible. A prototypical example for this is an interferometer, in which a change in the relative phase between the two arms is the quantity to be determined. A laser beam is split up and later recombined at a beam splitter, interfering with itself. Any change in the path lengths relative to each other can then be determined precisely in the interference pattern after the beams are recombined. Interferometers are quite universal measuring devices with such application as laser gyroscopes, phase-contrast microscopes, astronomical optical interferometry, etc. One way to improve the sensitivity of these interferometers is to use stronger lasers, but design constraints or availability make this not always a viable option.

Generally, for any estimation scheme, a relevant question is the following: Given a fixed amount of resources (such as a laser of a given power), how can we use this resource efficiently to determine the quantity we are interested in as accurate as possible? The first step is to isolate the apparatus from various sources of noise to improve the signal-to-noise ratio. Eventually, however, quantum mechanics imposes fundamental limits on how well one can do with a classical resource such as laser light even if all classical sources of noise are eliminated.

Here the field of quantum metrology comes in. The idea is to use quantum resources, such as exotic states of light, to mitigate the effects of quantum noise and

to improve the sensitivity of one's precision instruments.

One example we will focus on is arguably the most advanced interferometer of its kind: the Laser Interferometer Gravitational-Wave Observatory or LIGO. The idea is that a passing gravitational wave will cause a minuscule change in the 4 km long arms of the interferometer. To detect this change, a tremendous effort has been put into practically eliminating all kinds of sources of classical noise: The laser beams travel in ultra high vacuum, mirrors have to be suspended to subdue seismic noise, scattered light has to be 'filtered' out, etc. Despite all this effort, no gravitational waves have been detected so far.

In the regime LIGO is currently working in, the relevant part of the noise introduced through the quantum nature of light, is the *photon-counting errors*, which give rise to the *shot-noise limit*. Caves [Cav81] showed in 1981 that, in principle, an interferometer can surpass this limit, making use of an entirely non-classical form of light, a squeezed state of light. After successful test runs [VKL+10, KVL+12, C+11], this technology became part of the current LIGO upgrade [AAA+13], which is expected to push the sensitivity of LIGO into the realm where a successful detection of gravitational waves becomes very likely.

In this chapter I will describe some tools that I will use later on in the analysis of an interferometric scheme to estimate a phase. Central to this introduction will be the use of Fisher information for quantifying the performance of a parameter estimation scheme. A brief, but maybe slightly more rigorous introduction to the concept of Fisher information can be found in [CT12, VTB13]. In-depth discussions of the problem of parameter estimation for the quantum case are located in [Hol11, Hel76]. A more comprehensive review, discussing limits on optical interferometry in general, but also the use of Fisher information in this context, was recently written by Demkowicz-Dobrzanski *et al.* [DDJK14].

6.1 Classical parameter estimation and the Cramér-Rao bound

The problem central to parameter estimation is to infer the state of some system, specified through a set of parameters. One has to rely on a set of data gained from observations on the system to estimate the parameters. Usually these observations are not perfect, i.e., are compromised by random errors, so we have to use statistical tools to accomplish our goal. This setting can also be used to characterize a process, by letting a known state evolve under the unknown process and then observing the change in the system. To keep things simple, I will restrict myself here to the estimation of a single parameter, which I will call θ.

Mathematically, we describe these probabilistic measurement outcomes of imperfect observations through a probability distribution: $p(\vec{x}|\theta)$ is the probability for obtaining the set $\vec{x}$ of measured data, given that the parameter of the system is θ. Thus we get information about θ through sampling from the distribution $p(\vec{x}|\theta)$. We will assume that there are n_{rep} systems to sample from and that they are independent, in which case the distribution factorizes:

$$p(\vec{x}|\theta) = \prod_{i=1}^{n_{rep}} p(x_i|\theta) \tag{6.1}$$

After obtaining a set of data $\vec{x}$, we will need to map this data set to an estimate θ_{est} of the parameter we are interested in. The function $\Theta(\vec{x})$ that does this mapping is called an estimator. The error of the estimator $\Delta\Theta$ is its deviation from the true value of the parameter $\Delta\Theta = \Theta(\vec{x}) - \theta$.

As an example consider the measurement of the position of an object. Say we only have access to a 'coarse' ruler, and we have to guess the exact position. We could model the noise by a Gaussian (normal) distribution with standard deviation $\sigma = 1$ (in some arbitrary units); e.g., the position we measure x is the true position θ plus some Gaussian distributed random variable a: $x = \theta + a$, with $p(a) = N(0, 1)$, where $N(0, 1)$ is the zero-mean Gaussian distribution with unit variance. In this case

we have $p(x|\theta) = N(\theta, 1)$. Assume we measure 3 times and get $\vec{x} = \{8.7, 11.2, 10.5\}$. Now we need to map these outcomes to an estimate θ_{est}. The obvious thing to do here is to take the average $x_{avg} = 10.1 = \theta_{est}$. It turns out that in this case, taking the average is the best we can to as it minimizes the mean-squared error $\Delta\Theta_{MSE}$:

$$\Delta\Theta_{MSE} = \int d\vec{x}\, p(\vec{x}|\theta)\, (\Theta(\vec{x}) - \theta)^2 \tag{6.2}$$

In principle we could have chosen another way of obtaining the estimate, i.e. the geometric mean of our samples or simply the last measurement outcome. For more complicated problems it is by no means trivial or obvious to find the optimal estimator.

We say the estimator is unbiased if its error averaged over the measurement outcomes is zero:

$$\int d\vec{x}\, p(\vec{x}|\theta)\, \Delta\Theta = 0. \tag{6.3}$$

To simplify the following discussion, I will assume the estimators we are talking about are unbiased. Differentiating Eq. (6.3) with respect to the parameter we are trying to estimate we get

$$\frac{d}{d\theta} \int d\vec{x}\, p(x|\theta)\Delta\Theta = -\int d\vec{x}\, p(\vec{x}|\theta) + \int d\vec{x}\, \frac{\partial p(\vec{x}|\theta)}{\partial\theta}\Delta\Theta = 0. \tag{6.4}$$

Rewriting this equation, we get

$$\int dx_1 \cdots dx_n\, p(x_1|\theta) \cdots p(x_n|\theta) \left(\sum_{i=1}^{n} \frac{\partial \ln p(x_i|\theta)}{\partial\theta} \right) \Delta\Theta \tag{6.5}$$

$$= n \int dx\, p(x|\theta) \left(\frac{\partial \ln p(x|\theta)}{\partial\theta} \right) \Delta\Theta = 1. \tag{6.6}$$

Now we can group the terms

$$n \int dx\, \left(\sqrt{p(x|\theta)} \left(\frac{\partial \ln p(x|\theta)}{\partial\theta} \right) \right) \left(\sqrt{p(x|\theta)}\Delta\Theta \right) \tag{6.7}$$

and apply the Schwarz inequality to obtain the (classical) Cramér-Rao lower bound (CRLB) on the variance of the estimator,

$$\left\langle \Delta\Theta^2 \right\rangle \geq \frac{1}{n_{rep}\mathcal{F}(\theta)}, \tag{6.8}$$

where we introduced the Fisher information

$$\mathcal{F}(\theta) \equiv \int dx\, p(x|\theta) \left(\frac{\partial \ln p(x|\theta)}{\partial \theta} \right)^2 . \tag{6.9}$$

While $p(x|\theta)$ describes the probability of getting the outcome x from a measurement, given some fixed value of θ, we can look at it from a different perspective: After having obtained as set of measurement outcomes $\vec{x}_0$ we can think of $p(\vec{x}_0|\theta')$ as a function of θ'. In this case this function describes the likelihood that θ' has given rise to the observed measurement outcomes x_0; hence $p(x|\theta')$ is known as the likelihood function in this context.

The maximum likelihood estimator $\Theta_{MLE}(\vec{x})$ simply picks as an estimate the θ' for which the likelihood function $p(x|\theta')$ is maximized; that is, for a uniform prior on the parameter, the θ' that most likely caused the observed measurement outcomes. In our example above the likelihood function is simply the product of the three Gaussian distributions, which is again a Gaussian with mean $\mu = (8.7 + 11.2 + 10.5)/3 = 10.1$. We see from this example that for the case of Gaussian distributions, taking the mean of the data as an estimate corresponds to MLE.

It can be proven [Fis25, Cra99] that MLE asymptotically converges in probability to the true value:

$$\underset{n_{rep}\to\infty}{\text{p-lim}}\ \Theta_{MLE}(\vec{x}) = \theta. \tag{6.10}$$

Moreover, in this limit the estimate is asymptotically Gaussian $N(\theta, \sigma)$, with a standard deviation $\sigma = (n_{rep}\,\mathcal{F})^{-1/2}$. Hence, as $n_{rep} \to \infty$, MLE achieves the lower bound of Eq. (6.8).

In this work we are interested in quantifying the optimal performance of a particular setup for trying to estimate a parameter. This seems like a daunting task, as one might think we would have to find the optimal estimator, a problem for which no general answer is known. Depending on the specific estimation scheme, there may be estimators that do better than MLE; moreover, MLE might be computationally infeasible. Still, it is the asymptotic optimality of MLE that justifies the use of Fisher

information as a quantifier for the optimal performance in a particular estimation scheme, without even having to specify an estimator. There are certain pitfalls associated with the achievability of the CRLB, which one needs to be aware of when using it as the figure of merit for estimation performance. Generally, it is always a good idea to bolster a Fisher analysis by providing an explicit estimator whose error achieves the CRLB in a finite number of trials.

6.2 Parameter estimation in a quantum setting

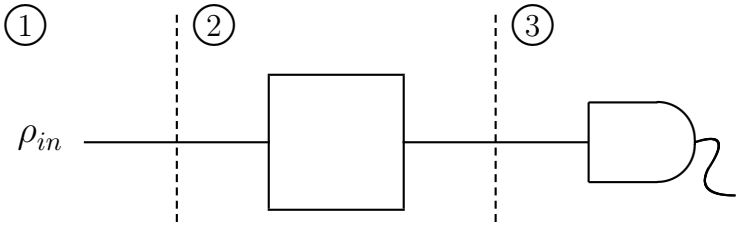

Figure 6.1: It is convenient to think of parameter estimation in a quantum setting as a three-step process: 1) A *probe* is prepared, i.e., the relevant quantum system is prepared into a known state. 2) The known probe state evolves according to a parameter-dependent process. 3) A measurement is performed and from the outcomes of this measurement the parameter of interest is estimated.

The estimation of a parameter in a quantum setting is usually depicted as the three-step process shown in Fig. 6.1: First, a quantum system is prepared in some known state ρ_{in}. Then this state is evolved by some parameter-dependent unitary process $U(\theta)$, e.g., a relative phase shift in two modes of light:

$$\rho(\theta) = U(\theta)\rho_{in}U(\theta)^{\dagger}\,. \tag{6.11}$$

Lastly, the system is subjected to a quantum measurement, modeled by POVM elements $\{E_x\}$. From the outcomes x of this measurement, the unknown parameter is inferred.

As mentioned earlier, it is desirable in estimation problems to use a given resource as efficiently as possible. In interferometry this resource is typically a fixed amount of energy in some given amount of time; one is interested in how to use a limited

number of photons optimally in order to maximize the interferometer's sensitivity to a relative phase shift. Only the first step in Fig. 6.1 consumes photons, but to assess the performance one needs to find the measurement that is optimal for a given input state and then find the optimal estimator to get an estimate of the relative phase give a set of measurement outcomes.

While this seems intimidating, one way to cut down on these optimization tasks is to use Fisher information as a performance quantifier, which eliminates the need to find an optimal estimator: The measurement $\{E_x\}$ generates the probability distribution $p(x|\theta) = \mathrm{tr}(\rho(\theta)E_x)$, from which we can calculate the classical Fisher information $\mathcal{F}(p(x|\theta)) \equiv \mathcal{F}_{E_x}(\rho(\theta)|E_x)$. What we are left with is to find the optimal measurement for a given state and find the state that delivers the best performance.

This remains a formidable task. It turns out, however, that again we can evade part of the work by using a generalization of the Fisher information, simply known as quantum Fisher information $\mathcal{F}_Q$. Quantum Fisher information is derived in [Hel76, Hol11] as a tool for quantum parameter estimation, but it is the derivation of Braunstein and Caves [BC94] that illustrates that quantum Fisher information is exactly the quantity we want in the scenario above. Quantum Fisher information of the state $\mathcal{F}_Q$, after being subjected to the parameter shift, corresponds to the classical Fisher information optimized over all possible quantum measurements:

$$\mathcal{F}_Q(\rho(\theta)) = \max_{\{E_x\}} \mathcal{F}_{E_x}(\rho(\theta)|E_x). \tag{6.12}$$

The quantum Fisher information for the estimation of parameter θ of state $\rho(\theta)$ is defined as the expectation value of the square of a Hermitian operator L_θ,

$$\mathcal{F}_Q(\rho(\theta)) = \mathrm{tr}(L_\theta^2 \rho(\theta)). \tag{6.13}$$

The operator L_θ is called the symmetric-logarithmic-derivative (SLD) operator in analogy to the derivative of the logarithm in the definition of the classical Fisher information. The SLD operator is defined by the implicit equation

$$\partial_\theta \rho(\theta) \equiv \frac{\partial}{\partial\theta}\rho(\theta) = \frac{L_\theta \rho(\theta) + \rho(\theta) L_\theta}{2}. \tag{6.14}$$

With this $\mathcal{F}_Q$ we have a powerful tool to analyze the limits of any quantum parameter estimation. The only variable we have to specify now is the input state ρ_{in}. In particular, we can use $\mathcal{F}_Q$ to answer the specific question: which state ρ_{in} is the optimal choice for estimating parameter θ, given a limited resource, such as a fixed photon number.

The advantage of using $\mathcal{F}_Q$ is not having to worry about a specific measurement setup, nor an estimator to go with it. On the downside there is no systematic way in general to obtain the operator L_θ from Eq. (6.14), which casts doubt on the usefulness of this method. Luckily under certain circumstances the quantum Fisher information can be calculated analytically, making it a viable option for the analysis of estimation schemes.

One of those easy cases, which will be of particular use to us, is the case when $\rho(\theta)$ is a pure state, i.e., $\rho(\theta) = |\,\psi\,\rangle\langle\,\psi\,|$, where I now omit the explicit θ dependence to make the following equation more readable. For a pure state, we have $\rho^2 = \rho$ and therefore

$$\partial_\theta\rho = (\partial_\theta\rho)\rho + \rho(\partial_\theta\rho). \tag{6.15}$$

Comparing this with Eq. (6.14) we see that since $\partial_\theta\rho$ is Hermitian,

$$L_\theta = 2\partial_\theta\rho = 2\big(|\,\partial_\theta\psi\,\rangle\langle\,\psi\,| + |\,\psi\,\rangle\langle\,\partial_\theta\psi\,|\big). \tag{6.16}$$

This allows us to write down an explicit expression for $\mathcal{F}_Q$:

$$\begin{aligned}\mathcal{F}_Q = 4\mathrm{tr}\Big(\rho\big(&|\,\partial_\theta\psi\,\rangle\,\langle\psi|\,\partial_\theta\psi\rangle\,\langle\,\psi\,| + |\,\partial_\theta\psi\,\rangle\,\langle\psi|\,\psi\rangle\,\langle\,\partial_\theta\psi\,| + |\,\psi\,\rangle\,\langle\partial_\theta\psi|\,\partial_\theta\psi\rangle\,\langle\,\psi\,| \\ &+ |\,\psi\,\rangle\,\langle\partial_\theta\psi|\,\psi\rangle\,\langle\,\partial_\theta\psi\,|\big)\Big) \qquad (6.17)\\ = 4\Big(&\langle\psi|\,\partial_\theta\psi\rangle\,\langle\psi|\,\partial_\theta\psi\rangle + \langle\psi|\,\partial_\theta\psi\rangle\,\langle\partial_\theta\psi|\,\psi\rangle + \langle\partial_\theta\psi|\,\partial_\theta\psi\rangle \\ &+ \langle\partial_\theta\psi|\,\psi\rangle\,\langle\partial_\theta\psi|\,\psi\rangle\Big). \qquad (6.18)\end{aligned}$$

From the trace condition, $\mathrm{tr}\rho = 1$, it follows that $\mathrm{tr}(\partial_\theta\rho) = 0 = \langle\psi|\,\partial_\theta\psi\rangle + \langle\partial_\theta\psi|\,\psi\rangle$; this allows us to write Eq. (6.18) in a more compact way:

$$\mathcal{F}_Q = 4\left(\langle\partial_\theta\psi|\,\partial_\theta\psi\rangle - |\langle\partial_\theta\psi|\,\psi\rangle|^2\right). \tag{6.19}$$

As this quantity is readily computable, we are now in the position to apply quantum Fisher information as the quantifier for the performance of estimation schemes. We will see a simple example of this application at the end of the discussion on coherent states in Sec. 6.3.

We should note that in general when more than one parameter, say k parameters, are unknown, the quantum Fisher information becomes a $k \times k$ matrix and the corresponding CRLB is a matrix bound. To keep this introduction simple, I kept it at the one-parameter case. When we need the Fisher information matrix for multiple parmaeters in Chap. 7, we will introduce it there.

6.3 Coherent and squeezed states

In this section I will describe some quantum states that are useful in the quantum mechanical description of light. I will assume some familiarity of the reader with the quantum mechanical description of a harmonic oscillator [CTDL91]. The Hamiltonian for a single mode of the radiation field can be written in terms of the creation and annihilation operators $a^\dagger$ and a:

$$H = a^\dagger a + \frac{1}{2}. \tag{6.20}$$

Here and in the following we will use natural units, with $\hbar = 1$ and $\omega = 1$, to keep the equations clear. The eigenstates $|\, n \,\rangle$ of this Hamiltonian are called number states. We have

$$a^\dagger a |\, n \,\rangle = n |\, n \,\rangle, \tag{6.21}$$

where $\hat{n} = a^\dagger a$ is known as the (photon) number operator. The expectation value of $\hat{n}$ is proportional to the energy needed to create a particular state and hence often a constraint in an estimation scheme. The creation operator $a^\dagger$ creates photons in the mode, while the annihilation opeators a destroys photons:

$$a |\, n \,\rangle = \sqrt{n} |\, n-1 \,\rangle, \qquad a^\dagger |\, n \,\rangle = \sqrt{n+1} |\, n+1 \,\rangle. \tag{6.22}$$

The Hermitian quadrature operators, corresponding to the position and momentum of a harmonic oscillator, can be written as

$$x = \frac{1}{\sqrt{2}}(a + a^\dagger) \qquad \text{and} \qquad p = -\frac{i}{\sqrt{2}}(a - a^\dagger), \tag{6.23}$$

and the respective commutators in our units are

$$[a, a^\dagger] = 1 \qquad \text{and} \qquad [x, p] = i. \tag{6.24}$$

The absence of a well-defined phase of the number states indicates that they are very different from the classical description of radiation. The number states form a complete basis, however, and we can use superpositions of number states to describe a quantum analogue to classical light, capturing its oscillatory behaviour.

Laser light is described quantum mechanically by a coherent state

$$|\,\alpha\,\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} |\,n\,\rangle. \tag{6.25}$$

Here α is a complex number, the analogue of the complex amplitude of a laser beam. A coherent state has the same quadrature fluctuation as the vacuum state $|\,0\,\rangle$. It is a minimum uncertainty state, achieving the Heisenberg inequality,

$$\langle\,\alpha\,|(\Delta x)^2|\,\alpha\,\rangle \langle\,\alpha\,|(\Delta p)^2|\,\alpha\,\rangle = \frac{1}{4}, \tag{6.26}$$

where the fluctuations in both quadratures are equal $\langle\,\alpha\,|(\Delta x)^2|\,\alpha\,\rangle = \langle\,\alpha\,|(\Delta p)^2|\,\alpha\,\rangle$. Coherent states are the same as the vacuum state, but displaced in phase space by the displacement operator, $|\,\alpha\,\rangle = D(a, \alpha)|\,0\,\rangle$, where the displacement operator is given by

$$D(a, \alpha) \equiv e^{\alpha a^\dagger - \alpha^* a}. \tag{6.27}$$

As an example of an estimation problem in a quantum setting, consider the following. The vacuum's x-quadrature has been displaced by an unknown amount α_0, and we want to know how accurately we can determine this unknown parameter and what the optimal measurement is to do so. A displacement of the x-quadrature corresponds to a real displacement amplitude, i.e., $\alpha = \alpha_0$. To calculate the quantum

Fisher information Eq. (6.19), we need to take derivatives with respect to α_0 of the state $|\,\alpha_0\,\rangle$. Using Eq. (6.27), we have

$$|\,\partial_{\alpha_0}\alpha_0\,\rangle = (a^\dagger - a)|\,\alpha_0\,\rangle = -i\sqrt{2}p|\,\alpha_0\,\rangle. \tag{6.28}$$

With this the quantum Fisher information becomes

$$F_Q/4 = \left(\langle\partial_{\alpha_0}\alpha_0|\,\partial_{\alpha_0}\alpha_0\rangle - |\,\langle\partial_{\alpha_0}\alpha_0|\,\alpha_0\rangle\,|^2\right) = 1. \tag{6.29}$$

Finding the optimal measurement is generally a hard problem. Our strategy here will be to make a reasonable guess about the right measurement and then prove that this measurement is indeed optimal. As the shift is in the x-quadrature, a good guess to measure in the position basis. The probability of obtaining measurement outcome x is [Cav]:

$$p(x|\alpha_0) = |\,\langle x|\,\alpha\rangle\,|^2 = \frac{1}{\sqrt{\pi}}e^{-(x-\sqrt{2}\alpha_0)^2}. \tag{6.30}$$

Plugging this into Eq. (6.9) for the classical Fisher information, we see after a short calculation that the classical Fisher information for a position measurement in this case matches the quantum Fisher information. This proves that the position measurement is optimal here. Since the statistics of the position measurement in this case are Gaussian, the Fisher bound can be achieved for any number of trials.

The Fisher information here is a constant. Typically we are interested in how the Fisher information scales with whatever resource we are considering. In the example here, however, our input state was the vacuum state and therefore we did not spend any resource in creating it. If we were to do the same calculation with a coherent state $|\,\beta\,\rangle$ as the input state, instead of vacuum, we would get the same result. In the next section we will use another state, the squeezed vacuum, that is more suited to the estimation of a quadrature shift; we turn to the properties of squeezed vacuum now.

A squeezed vacuum state is created from vacuum using the squeeze operator

$$S(r,\phi) = \exp\left(\frac{1}{2}r\left(a^2e^{-2i\phi} - (a^\dagger)^2e^{2i\phi}\right)\right). \tag{6.31}$$

A squeezed vacuum state is defined to be $S(r,\phi)|\,0\,\rangle$. It has the following expansion in the number basis:

$$S(r,\phi)|\,0\,\rangle = \frac{1}{\sqrt{\cosh r}} \sum_n \left(-\frac{1}{\sqrt{2}} e^{2i\phi} \tanh r\right)^n \sqrt{\frac{(2n-1)!!}{n!}} |\,2n\,\rangle. \tag{6.32}$$

Notable here is that the squeezed vacuum has support only on the even photon number states. Like coherent states, squeezed vacua with $\phi = 0$ are minimum uncertainty states, but for any $r \neq 0$, their quadratures differ in their fluctuations,

$$\langle\,0\,|S(r,0)^\dagger (\Delta x)^2 S(r,0)|\,0\,\rangle \neq \langle\,0\,|S(r,0)^\dagger (\Delta p)^2 S(r,0)|\,0\,\rangle. \tag{6.33}$$

Assume, as in the example above, that we are to estimate a shift in the x-quadrature, but now our input state is the squeezed vacuum, squeezed in the x quadrature, that is with $\phi = 0$. With this, the state after the quadrature shift is $|\,\psi(\alpha_0)\,\rangle = D(\alpha_0, a)S(r,0)|\,0\,\rangle$. From Eq. (6.19) and with the help of [Cav], we get

$$\mathcal{F}_Q/4 = 2\langle\,0\,|S(r,0)^\dagger (\Delta p)^2 S(r,0)|\,0\,\rangle = \cosh 2r + \sinh 2r = e^{2r}. \tag{6.34}$$

Again the position measurement is a good candidate for the the optimal measurement. We have:

$$\langle x|\,\psi(\alpha_0)\rangle = \frac{1}{(\pi e^{-2r})^{1/4}} \exp\left(-\frac{(x-\sqrt{2}\alpha_0)^2}{2e^{-2r}}\right) \tag{6.35}$$

To get the classical Fisher information, we can square this expression and plug it into Eq. (6.9) for the classical Fisher information. This is straightforward and immediately shows that this classical Fisher information is the same as $\mathcal{F}_Q$. Another way of showing the equality between the quantum and classical Fisher informations in this case is to use a condition for equality found in [BC94]:

$$\mathrm{Im}\{\langle\psi(\alpha_0)|\,x\rangle\,\langle x|\,\partial_{\alpha_0}\psi(\alpha_0)\rangle\} = 0. \tag{6.36}$$

This condition is satisfied, since Eq. (6.35) is real, and this again shows the optimality of the position measurement.

The mean photon number $\bar{n}$ of a squeezed state is $\bar{n} = \sinh^2 r$. This shows us that the squeezed vacuum state achieves a better resolution for the position as more and more photons are used.

6.4 Quantum noise and the shot-noise limit

Earlier we mentioned that coherent states resemble most closely the classical description of radiation. In particular we use coherent states to describe the light emitted from a laser. The most notable feature of coherent state, which distinguishes them from classical radiation, is the absence of a precisely defined energy. As seen in Eq. (6.25), coherent states are superpositions of photon number states. While number states $|\, n \,\rangle$ have a defined energy, coherent states can only be assigned a *mean* photon number $\bar{n}$ and therefore a mean energy. For coherent state $|\, \alpha \,\rangle$ we have

$$\bar{n} = \langle\, \alpha \,|\hat{n}|\, \alpha \,\rangle = |\alpha|^2 \,. \tag{6.37}$$

To measure the intensity of a coherent light source, one needs to count photons. Assuming a perfect photon counter, coherent states give rise to a Poisson distribution of counts with mean $|\alpha|^2$. It is this statistical behavior of photon detection that gives rise to what is known as the shot-noise limit.

Using only classical resources in an estimation scheme, such as coherent states, the shot-noise limit restricts the scaling of the estimate's error $\delta\phi_{SN}$ to

$$\delta\phi_{SN} \geq \frac{1}{\sqrt{\bar{n}}}. \tag{6.38}$$

Some refer to this as the standard quantum limit. We will not use this term when referring to the shot-noise limit, as we are primarily concerned with interferometry and the term 'standard quantum limit' has a different meaning here.

A second source of quantum noise that affects interferometer is fluctuating forces on the mirrors of the interferometer: The radiation pressure of photons exerts a force the mirror, whose positions get shifted accordingly. The statistical nature of the photon flux therefore causes an uncertainty in the positions of the mirror. This is known as radiation-pressure noise

$$\delta\phi_{RP} \propto \frac{\sqrt{\bar{n}}}{M}. \tag{6.39}$$

Contrary to the shot-noise, this form of quantum noise increases with the mean photon number $\bar{n}$ injected into the interferometer.

The overall quantum noise $\delta\phi$ now is given by [BR04]

$$\delta\phi^2 = \delta\phi_{RP}^2 + \delta\phi_{SN}^2. \tag{6.40}$$

For a given set of parameters for an interferometer, there is an optimal value $\bar{n}_{opt}$ for the mean photon number, thus balancing shot noise and radiation-pressure error.

The standard quantum limit of interferometry is precisely the quantum noise at this $\bar{n}_{opt}$. It turns out that for most currently realistic interferometers, with properties such as a length L and mass M of the mirrors, etc., the radiation pressure is negligible. Even for high power interferometers this is true [Cav81], as it is comparatively easy to scale the mirror masses up. In Chap. 7, the interferometers we consider have 'fixed' mirrors, that is we assume the mirror masses are large enough that we can neglect radiation-pressure noise. We will focus instead on techniques that promise to beat the shot-noise limit, the relevant quantum noise for typical, currently realistic interferometric setups.

6.5 Nonclassical states relevant for interferometry

In this section I will be talking about two nonclassical states that are of particular interest for interferometry: Squeezed states and NOON-states.

6.5.1 Squeezed state interferometry

In 1981 Caves [Cav81] gave a detailed analysis of the quantum mechanical noise in an interferometer. In a standard laser interferometer, a laser, modeled by a coherent state, enters one port of the interferometer, while the second port is unused, that is the second port is in the vacuum state. The detection scheme involves photon counting in both of the output ports. The mean photon number difference now is proportional to the relative phase shift between the two arms of the interferometer. Caves showed that one can think of the noise of the photon counting error being due

to fluctuations entering the unused port of the interferometer, that is the fluctuations of the vacuum state.

He demonstrated that the photon counting error can be reduced by squeezing the vacuum state in the second input mode, at the expense of increasing the radiation pressure error. As the radiation pressure error is negligible for typical, currently feasible interferometers, squeezing the vacuum in the second input port is a viable option to increase the sensitivity of the interferometer.

Other ways of using squeezed states in an interferometric setup have been investigated. In particular it was shown that squeezing the states in both ports is advantageous [BS84, KS96]. These schemes however are more complicated and lack the most attractive point of Caves' original proposal: One does not have to change the original setup for the laser interferometer; instead one simply has to add components in front of the second port. In particular one can rely on strong commercially available lasers as the main power source.

Pezze and Smerzi showed [PS08], that Caves' original proposal is not only capable of beating the shot noise limit, but can also achieve a 'Heisenberg scaling' of the phase sensitivity:

$$\delta\phi \propto \frac{1}{\bar{n}}. \tag{6.41}$$

However, a more complicated detection scheme is necessary to do so.

For linear parameter estimation, that is processes whose Hamiltonian coupling to the phase shift is at linear in the photon number, this $1/\bar{n}$ scaling is considered the be the best scaling possible and therefore is often referred to as the 'Heisenberg limit' [GLM06]. While this Heisenberg limit is easy to prove for the case of a fixed photon number, there is some controversy as to whether such a limit exists when one imposes only a mean photon number constraint.

6.5.2 NOON states

I will provide a simple way to prove this 'Heisenberg limit' for the case of a fixed photon number interferometer, by finding the state that maximizes the quantum Fisher information. The relative phase shift of ϕ in a two-mode interferometer is modeled by the unitary operator $U = \exp(i\phi N_d/2)$. Here we have introduced the photon number difference operator $N_d = a^\dagger a - b^\dagger b$ of the two modes a and b. If $|\,\psi\,\rangle$ describes the quantum state of the light after the first beamsplitter of the interferometer, the state after the phase shift is $U|\,\psi\,\rangle$. Plugging this into Eq. (6.19) yields

$$\mathcal{F}_Q = \langle\,\psi\,|N_d^2|\,\psi\,\rangle - \langle\,\psi\,|N_d|\,\psi\,\rangle^2. \tag{6.42}$$

Any two-mode (pure) state with a fixed photon number N can be written as

$$|\,\psi\,\rangle = \sum_n^N c_n e^{i\varphi_n}|\,n, N-n\,\rangle, \tag{6.43}$$

where the c_n are positive numbers. To simplify this discussion notice that we can always set all the φ_n to zero without changing $\mathcal{F}_Q$ of the state. Moreover, given a state in terms of the coefficients c_n, we can always construct a state with new coefficients c_n' obeying

$$c_n' = c_{N-n}' = \sqrt{c_n^2 + c_{N-n}^2}. \tag{6.44}$$

This new state will always have a quantum Fisher information at least as large as the old state, as the first term in Eq. (6.42) remains unchanged while the second term vanishes. This new state will have the form

$$c_0'\,(|\,0, N\,\rangle + |\,N, 0\,\rangle) + c_1'\,(|\,1, N-1\,\rangle + |\,N-1, 1\,\rangle) + \dots\ . \tag{6.45}$$

Writing out the first term of Eq. (6.42),

$$\langle\,\psi\,|N_d^2|\,\psi\,\rangle = \langle\,\psi\,|(a^\dagger a a^\dagger a + b^\dagger b b^\dagger b - 2a^\dagger a b^\dagger b)|\,\psi\,\rangle, \tag{6.46}$$

it is not too hard to see that the first term in Eq. (6.45) is giving the biggest contribution to this expression; hence the optimal state has the largest possible

$c_0' = 1/\sqrt{2}$. We have shown that the N-photon state that maximizes the Fisher information for a relative phase shift between two modes is

$$|\,\psi_{N00N}\,\rangle = \frac{1}{\sqrt{2}}(|\,N,0\,\rangle + |\,0,N\,\rangle) \tag{6.47}$$

The quantum Fisher information of this state is $\mathcal{F}_Q = N^2$; i.e., the state exhibits a $1/N$ scaling of the estimate error. As this state minimizes the CRLB, this demonstrates a 'Heisenberg-limit' for interferometry with a fixed number of photons.

Due to the form of Eq. (6.47) this state is known as the NOON state [BIWH96, LKD02]. After the relative phase shift the state becomes:

$$U|\,\psi_{N00N}\,\rangle = \frac{1}{\sqrt{2}}(|\,N,0\,\rangle + e^{-i\phi N}|\,0,N\,\rangle), \tag{6.48}$$

where we exploited the global phase freedom. From this expression we can see that this state is invariant under, and therefore insensitive to phase shifts of $2\pi/N$. Another way to say this is that any interference pattern produced by a N00N state will have a periodicity of $2\pi/N$. Hence a single N00N state is only suited to detect phase shifts $\phi < 2\pi/N$, as there is no way to determine which interference fringe one is seeing in an experiment.

While the usual method if subtracting the photocounts of two detectors is not applicable here, it has been shown that a parity detection will achieve the sensitivity promised by a quantum Fisher information calculation. This detection scheme simply looks at the parity of the photon count in one of the output modes to infer the value of the relative phase shift [BIWH96, GBC02]. Moreover, Berry *et al.* [BHB$^+$09] have shown that using an adaptive estimation scheme, the problem of the fringe ambiguity can be dealt with while maintaining the N00N state favorable scaling.

Despite these theoretical benefits of N00N states, their practical infeasibility makes them not a viable option for realistic interferometry The Hong-Ou-Mandel effect produces N00N states deterministically, but unfortunately only with $N = 2$. Other mechanisms to produce N00N states have not been able to produce N00N states with $N > 5$. Moreover, these methods rely on post-selection and therefore are highly inefficient [NOO$^+$07, AAS10]. Considering this, the quadratic enhancement of

sensitivity promised by the usage of N00N states becomes irrelevant when comparing to a commercially available interferometer running on 10^{16} photons per detector integration time.

6.6 Beating the Heisenberg limit

While the Heisenberg limit is straightforward to prove for an interferometer running on a fixed amount of photons simply by finding the optimal state explicitly, there is no proof that limits the sensitivity of phase estimates to $\Delta\phi \geq 1/\bar{n}$ when the constraint is only on a mean number $\bar{n}$ of photons. Still terms like 'Heisenberg limit' or 'Heisenberg scaling' are commonly used in the context of mean photon number constraints, as it is generally believed that this is the best scaling possible for linear parameter estimation schemes employing nonclassical resources.

There are several proposals, however, that challenge this belief by presenting states seemingly beating this Heisenberg limit. A careful analysis of these schemes reveals a sub-Heisenberg scaling. I will nevertheless describe two of them more closely in this section, as they are instructive in showing up a pitfall one can encounter when using various figure of merit, such as the quantum Fisher information, as the only quantifier for the estimation performance.

The following two proposals reportedly beating the Heisenberg limit are not talking about an interferometric setup; instead they simply consider a phase shift in a single mode of the electro-magnetic field. It is, however, straightforward to extend their arguments and apply them to an interferometric setup.

The first proposal I will talk about was put forward in 1989 by Shapiro, Shepard, and Wong [SSW89]. The state they were considering, after the phase shift, is

$$|\,\psi\,\rangle = A\sum_{n=0}^{M}\frac{e^{in\theta}}{n+1}|\,n\,\rangle, \tag{6.49}$$

where A is a normalization constant, M is a cutoff in the photon-number basis, which the mean photon number $\bar{n}$ depends on. The measurement scheme in their proposal

was an unphysical idealized phase measurement, described by the Susskind-Glogower phase operator [SG64]. Instead of Fisher information, the figure of merit for the sensitivity of their scheme is the so called *reciprocal peak likelihood*, that is the inverse of the maximum of the likelihood function. While for a Gaussian likelihood function, this quantity is directly related to the mean squared error, in general the idea is that the higher the peak of the likelihood function $p(x|\theta)$ the more distinguishable it will be from another likelihood function $p(x|\theta')$. SSW's claim was that their state leads to a phase sensitivity of

$$\Delta\theta \propto \frac{1}{\bar{n}^2}, \tag{6.50}$$

that is a quadratic improvement over Heisenberg scaling.

A careful analysis by Lane, Braunstein and Caves [BLC92, LBC93], however, showed that SSW's proposal is not able to beat the Heisenberg limit. Paradoxically, they also showed that Fisher information grows exponentially in the mean photon number for the SSW state, which on the first glance suggests an even better scaling than reported by in [SSW89].

This paradox arises if one does not pay close attention to subtleties associated with the achievability of the CRLB: When one talks about Heisenberg scaling, one usually refers to the scaling of the phase error $\Delta\theta$ in terms of a photon number resource available. Typically this total photon number N_{tot} needs to be divided up into a mean photon number $\bar{n}$ per probe state and a number n_{rep} of repetitions of the experiment in order to get enough data that the distribution of the estimator is nearly Gaussian and the conditions for the optimality of MLE are satisfied. That is, our resource can be written as

$$N_{tot} = n_{rep}\,\bar{n}. \tag{6.51}$$

For a specific estimator, i.e. MLE, and a given resource N_{tot} there will be an optimal number of repetitions $n^*_{rep}(N_{tot})$, and mean photon number per probe state $\bar{n}^*(N_{tot}$ that minimizes the error of the estimator [Bra92b, LBC93].

If the form of the input state is fixed and one is interested in how the performance of this state scales with N_{tot} one needs to find $\bar{n}^*(N_{tot})$. In the case of the SSW state,

Lane *et al.* [LBC93] showed that this dependence of $\bar{n}^*$ on N_{tot} is so unfavorable that the effect of the exponential scaling of the Fisher information with $\bar{n}$ is completely negated.

The second example I will talk about was proposed by Rivas and Luiz [RL12]. To keep the discussion simple, I will describe the R-L state in the form discussed by Tsang [Tsa12]. This form is a bit more general while highlighting at the same time the essential features better than R-L's original proposal. Consider the superposition of the vacuum $|\,0\,\rangle$ with a small amount of another state $|\,\xi\,\rangle$, which has a mean photon number $\bar{n}_\xi$ and large photon number variance $\gamma\,\bar{n}_\xi^2$, where γ is a positive constant. The overall state is

$$|\,\psi_{RL}\,\rangle = \sqrt{1-\epsilon}|\,0\,\rangle + \sqrt{\epsilon}\,|\,\xi\,\rangle, \tag{6.52}$$

where $\epsilon \ll 1$. We assume $\langle\xi|\,0\rangle = 0$ [if $|\,\xi\,\rangle$ has a $|\,0\,\rangle$ component, this can be absorbed into the first term of Eq. (6.52)]. As in Eq. (6.42), the quantum Fisher information for a phase shift of a pure state in a single mode is given by the variance of the photon number, and this gives [Tsa12]

$$\mathcal{F}_Q = (\Delta n)^2 = \left(\frac{1+\gamma}{\epsilon} - 1\right)\bar{n}. \tag{6.53}$$

Now we can see that in this case the Fisher information can be made arbitrarily large even for a fixed $\bar{n}$. Instead of doing a comprehensive sensitivity analysis like Lane *et al.* did for the previous example, Tsang simply derived another lower bound for the sensitivity in terms of the mean photon number [Tsa12]. He showed that in this case his Ziv-Zakai type bound is tighter than the CRLB and, more importantly, ruled out a Heisenberg scaling of the phase sensitivity of the R-L state. Tsang's analysis effectively shows something like what happens with the SSW state must also go wrong if one tries to use MLE on the R-L state.

To deal with the tricky point of achievability when using Fisher information as a measure of performance for an estimation scheme in the following chapters, we will provide an efficient estimator whenever possible. That is, in addition to calculating the CRLB we will provide an estimator whose error matches that of the CRLB after a few runs of the experiment.

In this chapter we presented some concepts and tools used in the analysis of the performance of parameter estimation schemes in general and interferometry in particular. My intention here was not to be particularly thorough, but to provide the information necessary to follow the next chapters. A more complete account of these concepts can be found in [DDJK14].

Chapter 7

Quantum-enhanced interferometry

This chapter will be devoted to the exploration of optimal input states for the estimation of a relative phase in an interferometer.

As shown in Chap. 6, for a fixed input energy the question of the optimal input state is readily answered with N00N states. More precisely, for an actual interferometer, the optimal input state would be the precursor of the N00N state, that is the input state $|\,\psi_{in}\,\rangle$ that gives rise to the N00N state after the first beam splitter of the interferometer.

$$|\,\psi_{in}\,\rangle = B^{\dagger}(|\,N,0\,\rangle + |\,0,N\,\rangle)/\sqrt{2}. \tag{7.1}$$

While this state indeed delivers optimal performance, given a fixed input energy, we argue that it is not of practical relevance because it is very hard to produce with current technology and is therefore only available with quite low photon numbers. Consequently, the phase resolution obtained from using these optimal states cannot compete with the resolution obtained from a classical interferometer operating at or near the shot-noise limit with a strong, commercially available laser.

This does not mean, however, that nonclassical states are useless for metrology. The use of squeezed states to enhance the sensitivity of the GEO 600 and LIGO interferometers is testimony to the efficacy of squeezed light in a situation where the lasers powering the interferometer have been made as powerful as design constraints

allow. In this work, we will try to turn the focus away from states that are simply of conceptual interest and instead we will look at states that deliver optimal performance under constraints motivated by practical concerns for currently feasible experiments.

7.1 Quantum-enhanced interferometry with laser light

Here we will answer a question motivated by the use of nonclassical states to enhance the sensitivity of high power interferometers such as GEO 600 and LIGO: when an interferometer is powered by a laser producing coherent-state light, what is the best state to put into an interferometer's secondary input port? The answer is not surprising: squeezed vacuum.

The setting for our analysis is depicted in Fig. 7.1. Specifically, we consider a situation where laser light, described by a coherent state $|\alpha\rangle = D(a,\alpha)|0\rangle$ of a mode a, is fed into the primary input port of a 50:50 beam splitter. The secondary input port is illuminated by mode b, which is in an arbitrary pure state $|\chi\rangle$. The beam splitter performs the unitary transformation

$$B = e^{-iJ\pi/4}\,, \qquad J \equiv a^\dagger b + b^\dagger a\,. \tag{7.2}$$

The two optical paths after the beam splitter experience phase shifts φ_1 and φ_2; the phase-shift unitary operator is

$$U = e^{i(\varphi_1\, a^\dagger a + \varphi_2\, b^\dagger b)} = e^{iN_s\phi_s/2} e^{iN_d\phi_d/2}\,. \tag{7.3}$$

In the second form we introduce the sum and difference phase shifts, $\phi_s = \varphi_1 + \varphi_2$ and $\phi_d = \varphi_1 - \varphi_2$, and the corresponding sum and difference number operators; $N_s = a^\dagger a + b^\dagger b$ is the total number operator for the two modes, and $N_d = a^\dagger a - b^\dagger b$ is the number-difference operator. We assume that there are no losses in this configuration. The two-mode state after the phase shifters is

$$|\psi\rangle = UB|\psi_{\rm in}\rangle\,, \tag{7.4}$$

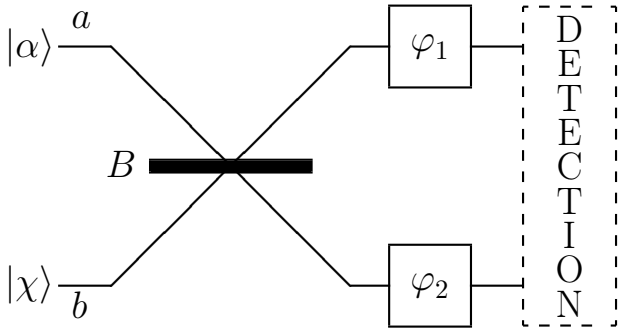

Figure 7.1: Measurement of a differential phase shift. An (upper) mode a in a coherent state $|\alpha\rangle$ and a (lower) mode b in an arbitrary pure state $|\chi\rangle$ are incident on a 50:50 beam splitter, which performs the unitary transformation B of Eq. (7.2). After the beam splitter, phase shifts φ_1 and φ_2 are imposed in the two arms; the action of the phase shifters is contained in the unitary operator U of Eq. (7.3). Finally, a measurement is made to detect the phase shifts. When the measurement is pushed beyond a second 50:50 beam splitter, the result is a Mach-Zehnder interferometer, which is sensitive only to the differential phase shift $\phi_d = \varphi_1 - \varphi_2$.

where $|\psi_{\rm in}\rangle$ is the state before the beam splitter.

As promised in the previous chapter, we will use the quantum Fisher information (QFI) to investigate the optimal resolution for estimating the phase shift ϕ_d. As the phase ϕ_s is an unknown parameter, we start out with the matrix form of the QFI and the quantum Cramér-Rao bound (QCRB). In particular, let $\phi_s^{\rm est}$ and $\phi_d^{\rm est}$ denote estimators of the sum and difference phase shifts, and introduce the covariance matrix of the estimators,

$$\Sigma = \begin{pmatrix} \langle(\Delta\phi_s^{\rm est})^2\rangle & \langle\Delta\phi_s^{\rm est}\Delta\phi_d^{\rm est}\rangle \\ \langle\Delta\phi_d^{\rm est}\Delta\phi_s^{\rm est}\rangle & \langle(\Delta\phi_d^{\rm est})^2\rangle \end{pmatrix} , \tag{7.5}$$

where here and throughout $\Delta O \equiv O - \langle O\rangle$ denotes the deviation of a quantity from its mean. The QCRB is the matrix inequality

$$\Sigma \geq \mathcal{F}^{-1} , \tag{7.6}$$

where $\mathcal{F}$ is the (real, symmetric) Fisher-information matrix [Hel76, Hol11]. The matrix QCRB implies that $\mathrm{tr}\Sigma \geq \mathrm{tr}\mathcal{F}^{-1}$ and $\det\Sigma \geq \det\mathcal{F}^{-1}$; for more than one parameter, the matrix QCRB cannot generally be saturated [Fuj06].

For pure states, the Fisher-information matrix is given by [JDDan12]

$$\mathcal{F}_{jk} = 4\left(\langle\partial_j\psi|\partial_k\psi\rangle - \langle\partial_j\psi|\psi\rangle\langle\psi|\partial_k\psi\rangle\right), \tag{7.7}$$

where j and k take on the values s and d and thus the derivatives are with respect to ϕ_s and ϕ_d. We retain both ϕ_s and ϕ_d in our analysis for the present, but eventually specialize to estimation of the differential phase shift alone. This would be the case if the final measurement were moved behind a second 50:50 beam splitter, giving a standard (Mach-Zehnder) interferometric configuration.

There are important practical reasons for considering the configuration of Fig. 7.1. The first is that in a typical phase measurement, the easiest way to improve sensitivity is to buy more photons. The cheapest coherent source being a laser, the relevant model is that of a laser producing an input coherent state with the largest possible amplitude. To avoid the phase noise of the laser, either intrinsic or excess, one splits the laser light at a 50:50 beam splitter. Phase shifts are imposed in the two arms, and then in a Mach-Zehnder configuration, the light in the two arms is recombined at a second 50:50 beam splitter, after which differenced photodetection or differenced homodyne detection is used to detect the differential phase shift. This interferometric technique is insensitive to the common-mode phase shift ϕ_s in the two arms, which is just another way of saying that it is insensitive to the laser noise. Yet another way of putting this is that each arm serves as a phase reference for the other.

The Mach-Zehnder interferometric configuration gives shot-noise-limited sensitivity when the secondary port is illuminated by vacuum. To go beyond the shot-noise limit, one must replace the vacuum coming into the secondary port with some other, nonclassical quantum state of light; this inevitably makes the light in the two arms of the interferometer entangled, this modal entanglement having been made by the input beam splitter.[1] A major advantage of the setting in Fig. 7.1 is that the main power production is separated from the generation of nonclassical light, which only

[1] Equation (7.9) shows that to beat the overall shot-noise limit requires nonclassical light into the secondary input port, since classical, coherent-state light has $\langle(\Delta p)^2\rangle = 1/2$. In App. B, we show that any nonclassical state in the secondary input port leads to modal entanglement between the two arms after the beam splitter. A more general proof can be found in [JLC13].

has to get a phase reference from the laser. Many analyses of phase sensitivity start by asking what entangled state in the two arms gives the best sensitivity, but this approach generally requires an entangled state that cannot be made by beamsplitting a product state and thus gives up the practical advantage of separating the main power production from the production of nonclassical light.

In accordance with this discussion, the intended mode of operation of our interferometer is to have the coherent state carry many more photons than the light input to the secondary port. Since it does not hinder our analysis, however, we allow for the opposite possibility and all intermediate ones in our analysis.

An analysis similar in spirit to ours has investigated the best performance of an interferometer, given a constraint on the total mean number of photons, when the primary input port is illuminated with many more photons than the secondary input port [DDanBS13]. The results show that a coherent state input to the primary port and squeezed vacuum into the other port comes very close to achieving a bound on the Fisher information that applies to all input states, both product and nonproduct states. This result holds when the photon loss exceeds a certain level, given in terms of the total mean number of photons, and thus is complementary to our result.

In our setting, the Fisher matrix for an arbitrary input state $|\psi_{\rm in}\rangle$ becomes

$$\mathcal{F}_{ss} = \langle\psi_{\rm in}|B^\dagger N_s^2 B|\psi_{\rm in}\rangle - \langle\psi_{\rm in}|B^\dagger N_s B|\psi_{\rm in}\rangle^2 \,, \tag{7.8}$$

$$\mathcal{F}_{dd} = \langle\psi_{\rm in}|B^\dagger N_d^2 B|\psi_{\rm in}\rangle - \langle\psi_{\rm in}|B^\dagger N_d B|\psi_{\rm in}\rangle^2 \,, \tag{7.9}$$

$$\mathcal{F}_{sd} = \mathcal{F}_{ds} = \langle\psi_{\rm in}|B^\dagger N_s N_d B|\psi_{\rm in}\rangle - \langle\psi_{\rm in}|B^\dagger N_s B|\psi_{\rm in}\rangle\langle\psi_{\rm in}|B^\dagger N_d B|\psi_{\rm in}\rangle \,. \tag{7.10}$$

We can use $B^\dagger a B = (a - ib)/\sqrt{2}$ and $B^\dagger b B = (b - ia)/\sqrt{2}$ to get $B^\dagger N_s B = N_s$ and

$$B^\dagger N_d B = -i(a^\dagger b - b^\dagger a) \equiv K \,. \tag{7.11}$$

The Fisher matrix is thus the covariance matrix of N_s and K with respect to the initial state. Notice that $J_z = N_d/2$, $J_x = J/2$, and $J_y = K/2$ make up the three components of an angular momentum and provide a convenient way of analyzing interferometry [YMK86].

For the product input that is our main concern,

$$|\psi_{\rm in}\rangle = |\alpha\rangle \otimes |\chi\rangle, \tag{7.12}$$

the Fisher matrix becomes, with $N_2 = b^\dagger b$,

$$\mathcal{F}_{ss} = |\alpha|^2 + \langle\chi|(\Delta N_2)^2|\chi\rangle \ , \tag{7.13}$$

$$\mathcal{F}_{dd} = |\alpha|^2\langle\chi|(\Delta b\Delta b^\dagger + \Delta b^\dagger\Delta b)|\chi\rangle \\ - \alpha^{*2}\langle\chi|(\Delta b)^2|\chi\rangle - \alpha^2\langle\chi|(\Delta b^\dagger)^2|\chi\rangle + \langle\chi|N_2|\chi\rangle \ , \tag{7.14}$$

$$\mathcal{F}_{sd} = \mathcal{F}_{ds} = -i\alpha^*\langle\chi|N_2(\Delta b)|\chi\rangle + i\alpha\langle\chi|(\Delta b^\dagger)N_2|\chi\rangle - i\alpha^*\langle\chi|b|\chi\rangle \ . \tag{7.15}$$

Partly because the matrix QCRB cannot generally be saturated [Fuj06], but chiefly because we are mainly interested in measurements of the differential phase shift, we specialize now to single-parameter estimation of ϕ_d, for which the QCRB reduces to

$$\langle(\Delta\phi_d^{\rm est})^2\rangle = \Sigma_{dd} \geq \frac{1}{\mathcal{F}_{dd}} \ . \tag{7.16}$$

It is known that there is a quantum measurement that achieves the single-parameter QCRB [BC94], i.e., has the required Fisher information, and it is also known that the resulting QCRB can be attained asymptotically in many trials by maximum-likelihood estimation.

What we do now is to maximize $\mathcal{F}_{dd}$ over all initial states $|\chi\rangle$ of mode b subject to a constraint of fixed mean photon number $\bar{N} = \langle\chi|N_2|\chi\rangle$. The optimal state turns out to be squeezed vacuum with the requisite mean photon number. We then use results of Pezzé and Smerzi [PS08] to indicate how the ultimate sensitivity can be achieved in a Mach-Zehnder interferometer in which one does direct photon detection of the two outputs.

To get started on maximizing $\mathcal{F}_{dd}$, we assume, without loss of generality, that α is real, and we write $\mathcal{F}_{dd}$ in terms of moments of the (Hermitian) quadrature components, x and p, of $b = (x + ip)/\sqrt{2}$:

$$\mathcal{F}_{dd} = 2\alpha^2\langle(\Delta p)^2\rangle + \bar{N} \ . \tag{7.17}$$

Here and for the remainder of the chapter, all expectation values are taken with respect to the initial state (7.12). The first term in Eq. (7.17), $2\alpha^2\langle(\Delta p)^2\rangle$, is due to interference between the coherent state and the phase quadrature p of the light coming into the secondary port; if $\alpha^2 \gg \bar{N}$, this term dominates and gives the shot-noise limit when mode b is in vacuum and improvements beyond shot noise when $\langle(\Delta p)^2\rangle > 1/2$. If $\alpha = 0$, the contribution from $\bar{N}$ in Eq. (7.17) dominates and expresses the shot-noise limit for illumination only through the secondary port.

We now maximize the variance of p, subject to a constraint on the mean number of photons. Writing

$$\begin{aligned} 2\bar{N} + 1 &= \langle p^2\rangle + \langle x^2\rangle \\ &= \langle p\rangle^2 + \langle x\rangle^2 + \langle(\Delta p)^2\rangle + \langle(\Delta x)^2\rangle \,, \end{aligned} \tag{7.18}$$

we see that

$$\langle(\Delta p)^2\rangle + \langle(\Delta x)^2\rangle \leq 2\bar{N} + 1 \,, \tag{7.19}$$

with equality if and only if $\langle x\rangle = \langle p\rangle = 0$. We also have

$$\begin{aligned} &\big(\langle(\Delta p)^2\rangle - \langle(\Delta x)^2\rangle\big)^2 \\ &\qquad = \big(\langle(\Delta p)^2\rangle + \langle(\Delta x)^2\rangle\big)^2 - 4\langle(\Delta x)^2\rangle\langle(\Delta p)^2\rangle \\ &\qquad \leq -1 + \big(\langle(\Delta p)^2\rangle + \langle(\Delta x)^2\rangle\big)^2 \\ &\qquad \leq 4\bar{N}(\bar{N}+1) \,, \end{aligned} \tag{7.20}$$

with equality in the first inequality if and only if $|\chi\rangle$ is a minimum-uncertainty state, i.e., $\langle(\Delta x)^2\rangle\langle(\Delta p)^2\rangle = 1/4$. Combining Eqs. (7.19) and (7.20) bounds $\langle(\Delta p)^2\rangle$ and hence gives a bound on the Fisher information,

$$\mathcal{F}_{dd} \leq \alpha^2\left(2\bar{N} + 2\sqrt{\bar{N}(\bar{N}+1)} + 1\right) + \bar{N} \equiv \mathcal{F}_{\max} \,, \tag{7.21}$$

with equality if and only if $|\chi\rangle$ is a zero-mean minimum-uncertainty state, i.e., the squeezed vacuum state $e^{r(a^2 - a^{\dagger 2})/2}|0\rangle$, with $\bar{N} = \sinh^2 r$. In terms of the squeeze parameter r, the bound on the Fisher information takes the simple form $\mathcal{F}_{\max} = \alpha^2 e^{2r} + \sinh^2 r$.

It is useful to manipulate the bound (7.21) in the following way:

$$\mathcal{F}_{\max} = 4\alpha^2\bar{N} + R = N_{\rm tot}^2 - (\alpha^2 - \bar{N})^2 + R\,. \tag{7.22}$$

Here $N_{\rm tot} = \alpha^2 + \bar{N}$ is the total mean number of photons into both input ports, and the remainder term is given by

$$R = \bar{N} + \alpha^2\left(2\sqrt{\bar{N}(\bar{N}+1)} - 2\bar{N} + 1\right)\,. \tag{7.23}$$

Applying $\bar{N} \leq \sqrt{\bar{N}(\bar{N}+1)} \leq \bar{N} + \frac{1}{2}$, we have $N_{\rm tot} = \alpha^2 + \bar{N} \leq R \leq 2\alpha^2 + \bar{N} = N_{\rm tot} + \alpha^2$. When $N_{\rm tot}$ is large, the remainder term is negligible compared to $N_{\rm tot}^2$. Moreover, when $\alpha^2 = \bar{N}$, we have $\mathcal{F}_{\max} = N_{\rm tot}^2 + R$, which gives the Heisenberg limit on phase sensitivity plus a small correction that satisfies $N_{\rm tot} \leq R \leq 3N_{\rm tot}/2$. The apparent violation of the Heisenberg limit comes from not having a fixed total number of photons. That this configuration using coherent and squeezed light can achieve the Heisenberg limit was shown in [PS08].

The case of primary practical interest has $\alpha^2 \gg \bar{N} = \sinh^2 r$, in which case the maximal Fisher information reduces to $\mathcal{F}_{\max} = \alpha^2 e^{2r}$. This corresponds to the standard picture of reduced fluctuations in the quadrature that produces differential phase fluctuations in the interferometer, and it gives the standard phase sensitivity, $1/\sqrt{\mathcal{F}_{\max}} = e^{-r}/\alpha$, for a squeezed-state interferometer. Indeed, the Fisher bound can be achieved by recombining the two optical paths at a second 50:50 beam splitter to create an interferometer and performing direct detection of the two outputs. The estimator can be taken to be the standard linear estimator that inverts the fringe pattern of the differenced photocount to estimate the differential phase shift.

Though it might be surprising, squeezed vacuum remains the optimal state into the secondary port even when the secondary port is allowed as many or more photons as the coherent-state input. We can appeal to the results of Pezzé and Smerzi [PS08] to show that the Mach-Zehnder configuration, with coherent-state and squeezed-vacuum inputs and direct detection at the output, can achieve the QCRB (7.21) for all values of the ratio $\alpha^2/\bar{N}$. Pezzé and Smerzi showed that for this configuration, the *classical* Fisher information of the probability for the output sum and and

difference photocounts, $P(n_s, n_d|\phi_d) = P(n_d|n_s, \phi_d)P(n_s)$, is equal to $\mathcal{F}_{\max}$ [2]. When $\alpha^2 \lesssim \bar{N} = \sinh^2 r$, however, the interferometer is running partially or mainly on the phase dependence of the squeezed vacuum noise, and the standard linear estimator mentioned above does not deliver optimal sensitivity [PS08, BS84]. Indeed, one can use the convexity of the Fisher information [Coh68] to show that any estimator that uses only the differenced photocount n_d, ignoring the sum photocount n_s, does worse than keeping both,[3] even though n_s is insensitive to the differential phase shift ϕ_d. Instead of using an estimator to verify that the classical Fisher bound—and, hence, from our analysis, the QCRB—can be achieved, Pezzé and Smerzi simulated a Bayesian analysis that indicates the classical bound can be achieved for all ratios $\alpha^2/\bar{N}$.

We note that squeezed vacuum is not the state that maximizes the entanglement of the two optical paths after the input beam splitter. A number state $|\bar{N}\rangle$ in the second mode leads to a larger value of the marginal entropy of the two paths [ACR05].

7.2 Quantum-enhanced interferometry

In this section we will relax one of the assumptions above. In particular we will not require the input state in the first mode to be a coherent state anymore. We only assume that the state input to the beam splitter is a product state $|\,\psi_{\rm in}\,\rangle = |\,\xi\,\rangle \otimes |\,\chi\,\rangle$, where $|\,\xi\,\rangle$ is the state of mode a and $|\,\chi\,\rangle$ is the state of mode b.

Our restriction to product states input to the initial beam splitter is natural—indeed, it is the only sensible assumption—in the case of an interferometric setup.

[2] Pezzé and Smerzi [PS08] did not include this calculation, so we present it in App. (B). There we show a more general result: the classical Fisher information for a Mach-Zehnder configuration with direct detection at the output is given by the quantum Fisher information $\mathcal{F}_{dd}$ of Eq. (7.17) for any input state $|\psi_{\rm in}\rangle$ that has real expansion coefficients in the number basis.

[3] Since the transformations in the interferometer preserve total photon number, the probability of measuring a total photon number n_s and a differenced photon number n_d at the output factors into $P(n_s, n_d|\phi_d) = P(n_d|n_s, \phi_d)P(n_s)$. Applying the convexity property [Coh68] iteratively leads to $\mathcal{F}_{n_s,n_d}(\phi_d) > \mathcal{F}_{n_d}(\phi_d)$.

Product inputs do generally lead to modal entanglement, i.e., entanglement between the two arms, after the initial beam splitter. In an interferometric setup, one is relying on the beam splitter to create modal entanglement from product inputs. If, in contrast, one allowed arbitrary, entangled states $|\,\Xi\,\rangle$ of the two modes to be incident on the beam splitter, one could dispense with the initial beam splitter, since one could just as well input any entangled state $B|\,\Xi\,\rangle$ directly into the two arms approaching the phase shifters.

Again, the relevant part of the Quantum Fisher matrix to determine the sensitivity of a relative phase shift $\mathcal{F}_{dd}$ from Eq. (7.9), is the variance of N_d of the state $B|\psi_{\text{in}}\rangle$ after the first beam splitter:

$$\begin{aligned}\mathcal{F}_{dd} &= \langle\psi_{\text{in}}|B^\dagger N_d^2 B|\psi_{\text{in}}\rangle - \langle\psi_{\text{in}}|B^\dagger N_d B|\psi_{\text{in}}\rangle^2 \\ &= -\langle(a^\dagger b - b^\dagger a)^2\rangle + \langle a^\dagger b - b^\dagger a\rangle^2 \\ &= -\langle a^\dagger a^\dagger bb\rangle - \langle aab^\dagger b^\dagger\rangle + \langle a^\dagger abb^\dagger\rangle + \langle aa^\dagger b^\dagger b\rangle \\ &\quad + \langle a^\dagger b\rangle^2 + \langle ab^\dagger\rangle^2 - 2\langle a^\dagger b\rangle\langle ab^\dagger\rangle\,.\end{aligned} \tag{7.24}$$

We still use the convention that expectation values are to be taken with respect to the input state $|\,\psi_{\text{in}}\,\rangle$: $\langle O\rangle = \langle\psi_{\text{in}}|O|\psi_{\text{in}}\rangle$. Notice however that now this is a state of both modes a and b, whilst before expectation values were only taken with respect to the state in mode b.

To find the optimal performance, we maximize $\mathcal{F}_{dd}$ over all product input states, subject to whatever additional constraints we impose on the input; i.e., we find the input product state that maximizes the variance of N_d after the first beam splitter. The expression (7.24) is valid for arbitrary inputs; specializing to product inputs gives

$$\begin{aligned}\mathcal{F} = 2N_aN_b + N_a + N_b - \langle a^\dagger a^\dagger\rangle\langle bb\rangle - \langle aa\rangle\langle b^\dagger b^\dagger\rangle \\ - 2|\langle a\rangle|^2\,|\langle b\rangle|^2 + \langle a^\dagger\rangle^2\langle b\rangle^2 + \langle a\rangle^2\langle b^\dagger\rangle^2\,,\end{aligned} \tag{7.25}$$

where $N_a = \langle a^\dagger a\rangle$ and $N_b = \langle b^\dagger b\rangle$ are the mean photon numbers in the two input modes.

Liu *et al.* [LJW13] have considered a setup similar to the one we consider here, a Mach-Zehnder interferometer with a product-state input. They focused on the Fisher

information for an arbitrary state in mode a and a state that is a superposition of even or odd photon numbers in mode b.

In the remainder of this section, we first find the optimal input state for a fixed photon number and then find the optimal state for a constraint on mean photon number.

7.2.1 Fixed photon number

If we fix the total photon number $N = N_a + N_b$, all product states have the form $|n\rangle \otimes |N-n\rangle$. Under these circumstances, only the first three terms in Eq. (7.25) contribute to the Fisher information. Finding the maximum reduces to finding the n that maximizes $2n(N-n)$; the maximum is achieved at $n = N/2$ for N even and at $n = (N \pm 1)/2$ for N odd, the two signs corresponding to an exchange of the input modes. The maximal Fisher information is

$$\mathcal{F}_{\max} = \begin{cases} \dfrac{N(N+2)}{2}\,, & N \text{ even}, \\ \dfrac{N(N+2)-1}{2}\,, & N \text{ odd}. \end{cases} \tag{7.26}$$

The optimal state,

$$|\psi_{\text{in}}\rangle_N = \begin{cases} |N/2\rangle \otimes |N/2\rangle\,, & N \text{ even}, \\ |(N \pm 1)/2\rangle \otimes |(N \mp 1)/2\rangle\,, & N \text{ odd}. \end{cases} \tag{7.27}$$

is the *twin-Fock state* for N even [HB93] and its closest equivalent for N odd. For brevity, we use "twin-Fock state" to refer to both the even and odd input states in the following; when we need to distinguish even and odd N, we refer to the former as "identical twins" and the latter as "fraternal twins."

The optimal state gives rise to a Quantum Cramér-Rao Bound (QCRB) for the variance of the phase estimate given by

$$(\Delta\phi_d^{\text{est}})^2 \geq \frac{1}{\mathcal{F}_{\max}} = \begin{cases} \dfrac{2}{N(N+2)}\,, & N \text{ even}, \\ \dfrac{2}{N(N+2)-1}\,, & N \text{ odd}, \end{cases} \tag{7.28}$$

which shows an asymptotic Heisenberg scaling. An input twin-Fock state leads to modal entanglement between the two arms after the beam splitter [JLC13].

Holland and Burnett [HB93] introduced the twin-Fock state (for N even) and considered the Heisenberg scaling of its phase sensitivity. Measurements that achieved the Heisenberg scaling were demonstrated in [KPH$^+$98, CGB03, GM10]. Robustness of the identical-twin-Fock state against various errors was investigated in [KPH$^+$98] and [MH09], and sub-shot-noise precision for interferometry with identical-twin-Fock states was demonstrated experimentally in [SLG$^+$08].

Benatti *et al.* [BFM10] considered the Fisher information for detecting a differential phase shift in an interferometric setting, with the constraint that there be no entanglement between the two inputs to the interferometer. They showed that the identical-twin-Fock state has the Fisher information expressed by the N-even case of Eq. (7.26).

If we were to remove the restriction of having a product input to the interferometer, the optimal input would be the state that maximizes the variance of N_d after the first beam splitter, i.e. the modally entangled input state $|\Xi\rangle$ that becomes a *N00N state*, $B|\Xi\rangle = (|N,0\rangle + |0,N\rangle)/\sqrt{2}$, after the beam splitter [Ger00, BKA$^+$00, Dow08]. The N00N state has a Fisher information $\mathcal{F} = N^2$; this is generally larger than the Fisher information (7.26) of the twin-Fock input, because one is optimizing over a larger set of input states to the initial beam splitter. For $N = 1$ and $N = 2$, however, the twin-Fock product input does produce a N00N-like state on the other side of the beam splitter: the $N = 1$ fraternal-twin-Fock input, $|1\rangle \otimes |0\rangle$, leads to the N00N-like state $B|1,0\rangle = (|1,0\rangle - i|0,1\rangle)/\sqrt{2}$ after the beam splitter, and the $N = 2$ identical-twin-Fock input, $|1\rangle \otimes |1\rangle$, leads to the N00N state $B|1,1\rangle = -i(|2,0\rangle + |0,2\rangle)/\sqrt{2}$. Thus the twin-Fock inputs for $N = 1$ and $N = 2$ have the same Fisher information as the N00N state; for $N > 2$, however, the N00N-state Fisher exceeds that of the twin-Fock input and is a factor of 2 larger asymptotically for large N.

7.2.2 Fixed mean photon number

We now move on to a constraint on the mean photon number $N_a + N_b$; in this section N denotes this total mean photon number. The proof for the optimal input state consists of two steps. The first step finds the optimal states $|\,\xi\,\rangle$ and $|\,\chi\,\rangle$ under the assumption that both N_a and N_b have fixed values. It turns out that the form of the optimal states is independent of how the total mean photon number N is divided up between N_a and N_b. In the second step, we show that the optimal split of resources is an equal division, $N_a = N_b = N/2$.

We begin the first step by noticing that the quantity on the second line of Eq. (7.25) is either negative or zero. We ignore this quantity for the moment. As we see shortly, the product input state that maximizes the top line has $\langle a\rangle = \langle b\rangle = 0$ and, therefore, also maximizes the quantum Fisher information $\mathcal{F}$. Furthermore, in this first step, N_a and N_b are assumed to be fixed, so maximizing the top line reduces to maximizing

$$-\langle a^\dagger a^\dagger\rangle\langle bb\rangle - \langle aa\rangle\langle b^\dagger b^\dagger\rangle. \tag{7.29}$$

We can always choose the phase of mode a, i.e., multiply a by a phase factor, to make $\langle aa\rangle$ real and positive, i.e., $\langle aa\rangle = \langle a^\dagger a^\dagger\rangle \geq 0$. With this choice we need to maximize

$$-\langle aa\rangle(\langle bb\rangle + \langle b^\dagger b^\dagger\rangle) = \langle aa\rangle(\langle p^2\rangle - \langle x^2\rangle)\,, \tag{7.30}$$

where in the second form we introduce the quadrature components x and p for mode b, i.e., $b = (x + ip)/\sqrt{2}$.

The proof continues along the lines of the proof in the previous section (Eq. (7.20)):

$$\begin{aligned}
\big(\langle p^2\rangle - \langle x^2\rangle\big)^2 &= \big(\langle p^2\rangle + \langle x^2\rangle\big)^2 - 4\langle x^2\rangle\langle p^2\rangle \\
&= (2N_b + 1)^2 - 4\langle x^2\rangle\langle p^2\rangle \\
&\leq (2N_b + 1)^2 - 4\langle(\Delta x)^2\rangle\langle(\Delta p)^2\rangle \\
&\leq (2N_b + 1)^2 - 1 \\
&= 4N_b(N_b + 1)\,.
\end{aligned} \tag{7.31}$$

The first inequality is saturated if and only if $\langle x\rangle = \langle p\rangle = 0$; equality is achieved in the second inequality if and only if the input state $|\,\chi\,\rangle$ of mode b is a minimum-

uncertainty state. This situation is identical to that in Sec. (7.1): the optimal choice for $|\,\chi\,\rangle$ is squeezed vacuum with x the squeezed quadrature and p the anti-squeezed quadrature.

What is left now is to maximize

$$2\sqrt{N_b(N_b+1)}\langle aa\rangle = \sqrt{N_b(N_b+1)}(\langle aa\rangle + \langle a^\dagger a^\dagger\rangle) \tag{7.32}$$

over the input states $|\,\xi\,\rangle$ of mode a for which $\langle aa\rangle$ is real and positive. This is the same maximization we just performed for mode b, except for a sign change, whose effect is to exchange the squeezed and anti-squeezed quadratures. The optimal state $|\,\xi\,\rangle$ is squeezed vacuum with p as the squeezed quadrature and x as the anti-squeezed quadrature.

Summarizing, the optimal input state is $S_a(-r)|0\rangle \otimes S_b(r')|0\rangle$, where r and r' are real and positive and $S_c(\gamma) = \exp[\frac{1}{2}(\gamma c^2 - \gamma^* c^{\dagger 2})]$ is the squeeze operator for a field mode c. The values of the squeeze parameters are determined by $N_a = \sinh^2 r$ and $N_b = \sinh^2 r'$. Notice that, as promised, the optimal state has $\langle a\rangle = \langle b\rangle = 0$ and thus maximizes the Fisher information (7.25); the maximum value is

$$\mathcal{F} = 2N_aN_b + N_a + N_b + 2\sqrt{N_a(N_a+1)N_b(N_b+1)}\,. \tag{7.33}$$

The second step of the proof is now trivial. For a constraint on the total mean photon number $N = N_a + N_b$, it is straightforward to see that Eq. (7.33) is maximized by splitting the photons equally between the two modes, i.e., $N_a = N_b = N/2$. The resulting optimal input state has $r = r'$,

$$|\psi_{\rm in}\rangle_{\rm opt} = S_a(-r)|0\rangle \otimes S_b(r)|0\rangle\,, \tag{7.34}$$

and the maximal Fisher information and corresponding QCRB are

$$\mathcal{F} = N(N+2)\,, \quad (\Delta\phi_d^{\rm est})^2 \geq \frac{1}{\mathcal{F}_{\rm max}} = \frac{1}{N(N+2)}\,. \tag{7.35}$$

This again exhibits Heisenberg scaling and, without the factor of 2 that appears in the fixed-photon-number result (7.26), achieves the $1/N^2$ Heisenberg limit.

A question that naturally arises is that of the optimal measurement. In App. B we show, building on work of Pezze and Smerzi [PS08], that the classical Fisher information of photon counting after a second 50:50 beam splitter is the same as the quantum Fisher information, provided the coefficients of the expansion of the input state in the number basis are real. This requirement is met by the optimal state (7.34).

Unlike the situation where there is a strong mean field, however, the interferometer with dual squeezed-vacuum inputs runs on modulated noise, so the mean of the differenced photocount after a second 50:50 beam splitter gives no information about the phase. One strategy for extracting the phase information is to look directly at the fluctuations by squaring the differenced photocount and thus effectively measuring N_d^2 [KPH+98, STL+14]; one can show[4] that the sensitivity at the optimal operating point achieves the QCRB (7.35).

Notice now that since $BaB^\dagger = (a+ib)/\sqrt{2}$ and $BbB^\dagger = (b+ia)/\sqrt{2}$, we have $B(a^2-b^2)B^\dagger = a^2 - b^2$. Thus the beam splitter leaves unchanged the product of squeeze operators in the optimal input state (7.34),

$$BS_a(-r)S_b(r)B^\dagger = S_a(-r)S_b(r)\,, \tag{7.36}$$

and this in turn means that the optimal input state is an eigenstate of the beam splitter,

$$B|\psi_{\rm in}\rangle_{\rm opt} = BS_a(-r)S_b(r)|0,0\rangle = |\psi_{\rm in}\rangle_{\rm opt}\,. \tag{7.37}$$

Thus the state after the 50:50 beam splitter is the same product of squeezed vacua as before the beam splitter[5]. The Heisenberg limit is thus achieved without any

[4]The squared error in estimating ϕ_d from $\langle N_d^2\rangle$ is given by $(\Delta\phi_d^{\rm est})^2 = {\rm Var}(N_d^2)/|\partial\langle N_d^2\rangle/\partial_{\phi_d}| = 1/N(N+2) + (\cos^2\phi_d/4\sin^2\phi_d)[7 + 2/N + N/(N+2)]$, from which we can see that at the optimal operating point, $\phi_d \simeq \pi/2$, this procedure achieves the QCRB.

[5]Since $B(a^2+b^2)B^\dagger = 2iab$, we have $BS_a(r)S_b(r)B^\dagger = e^{ir(ab+a^\dagger b^\dagger)}$. This means that dual, equally squeezed vacua into the two input ports, $S_a(r)|0\rangle \otimes Sb(r)|0\rangle$, become a two-mode squeezed state $e^{ir(ab+a^\dagger b^\dagger)}|0\rangle\otimes|0\rangle$ after the beam splitter. This well-known situation is not, however, the one that gives maximal Fisher information.

entanglement between the arms of the interferometer. In fact, Jiang, Lang, and Caves [JLC13] showed that the state $|\psi_{\text{in}}\rangle_{\text{opt}}$ is the only nonclassical product state, i.e., not a coherent state, that produces no modal entanglement after a beam splitter. These results indicate that, as in [SQ15], modal entanglement is not a crucial resource for quantum-enhanced interferometry.

Caves pointed out that using squeezed states in an interferometer allows one to achieve sensitivities below the shot-noise limit [Cav81]; this original scheme, often simply dubbed "squeezed-state interferometry," involves injecting squeezed vacuum into the secondary input port of an interferometer. That squeezing the light into the primary input port, in addition to inputting squeezed light into the secondary port, is advantageous was first shown by Bondurant and Shapiro [BS84] and further investigated by Kim and Sanders [KS96]. All these papers, however, included a mean field in at least one of the input modes. Paris argued [Par99] that if one considers arbitrary squeezed-coherent states as interferometer inputs, putting all the available power into the squeezing, instead of into a mean field, yields better fringe visibility. Under a Gaussian constraint, Refs. [PFB$^+$12] and [DDJK14] showed that a state that maximizes the Fisher information for a detecting a differential phase shift after a beam splitter is dual squeezed vacua; relative to these last results, the contribution of this section is to remove the assumption of Gaussianity, replacing it with a restriction to product inputs.

A problem with using Fisher information to find optimal states under a mean-number constraint is that one can come up with states that seemingly violate the Heisenberg limit. This was noted for single-mode schemes by Shapiro [SSW89] and later by Rivas [RL12]. For the former case, Braunstein and co-workers showed that under a precise asymptotic analysis, no violation of the Heisenberg limit occurs [Bra92a, BLC92, LBC93]. For the latter case, it was shown that the Fisher information does not provide a tight bound, which makes a Fisher analysis uninformative [Tsa12, GM12]. If we were to allow arbitrary (entangled) states $|\Xi\rangle$ as inputs in our scheme, we would run into the same problem[6]. Requiring product inputs removes

[6]Since the Fisher information is the variance of N_d after the beam splitter, it is given

this pathology of the Fisher information, therefore providing additional motivation for our product constraint.

We have analyzed a real-world scheme for measuring differential phase shifts. First we looked at the case where a coherent state illuminates one side of a 50:50 beam splitter and an arbitrary quantum state of light the other. We showed that given a constraint on the total mean number of photons, the optimal state to put into the secondary input port is a squeezed vacuum state, regardless of the relative mean photon numbers of the two inputs. Then we looked at the more general case where both sides of the 50:50 beam splitter are illuminated by arbitrary quantum states of light, with the only requirement that the joint state is of product form. We demonstrated that in this case the optimal choice is to use squeezed vacua in both modes, again independent of the relative mean photon number. For a joint mean number constraint the optimal choice is to divide the photons evenly.

by $\mathrm{Var}(N_d) = \mathrm{Var}(N_a) + \mathrm{Var}(N_b) - 2\,\mathrm{Covar}(N_a, N_b)$. If mode b is in vacuum, we have $\mathrm{Var}(N_d) = \mathrm{Var}(N_a)$, and this is the quantity that was made arbitrarily large under a mean-number constraint in [RL12].

Chapter 8

Quantum-enhanced interferometry with losses

The main motivation for the work discussed in Chap. 7 was the infeasibility of optimal states previously discussed for quantum-enhanced interferometry. We emphasized the importance of additional constraints, motivated by practical considerations, when discussing the advantage promised by the use of nonclassical states of light in a real world interferometer. Contrary to these efforts of stressing realistic concerns, however, the work in Chap. 7 assumed an idealized interferometer, that is, an interferometer that is not subjected to any form of noise. This chapter remedies this defect by analyzing interferometers subjected to photon loss.

The dominant source of noise in linear optical setups is loss of photons. For an interferometric setting, several authors have shown independently [FI08, KDD10, KSD11], that photon losses will invariably degrade the sought after 'Heisenberg scaling', $\mathcal{F} \propto N^2$, of the quantum Fisher information to a shot-noise type scaling of $\mathcal{F} \propto N$ as $N \to \infty$. While this seems devastating at the first glance, it does not mean that nonclassical states of light are useless for realistic interferometry: While the scaling of the quantum Fisher information asymptotically limits to the same scaling one would get employing only classical resources, nonclassical states can still beat the shot noise limit by a constant factor that depends on the specifics of the setup.

Moreover, given a particular real world setup, the photon numbers involved might be below the point at which the asymptotic scaling is relevant.

Nevertheless, the above at least indicates that it is critical to include losses in an analysis of the performance of an interferometric setup. In this chapter we extend the analysis performed in Chap. 7 by taking linear losses into account. These losses can be readily modeled by introducing two beam splitters into the setup we were considering before, as we illustrate in Fig. 8.1. These beam splitters, which are assumed to be identical, with transmissivity η, allow photons to leak out of the primary system modes, which have annihilation operators a and b, into two ancillary 'loss-modes', E_1 and E_2, which are traced out subsequently. The average fraction of photons lost at each loss beamsplitter is $1-\eta$. The rest of the setup is analogous to the one discussed in the previous chapter. Notice that we place the auxiliary beam splitters after the phase shifts. In App. B we show that for the following discussion it does not matter where these beam splitters are placed.

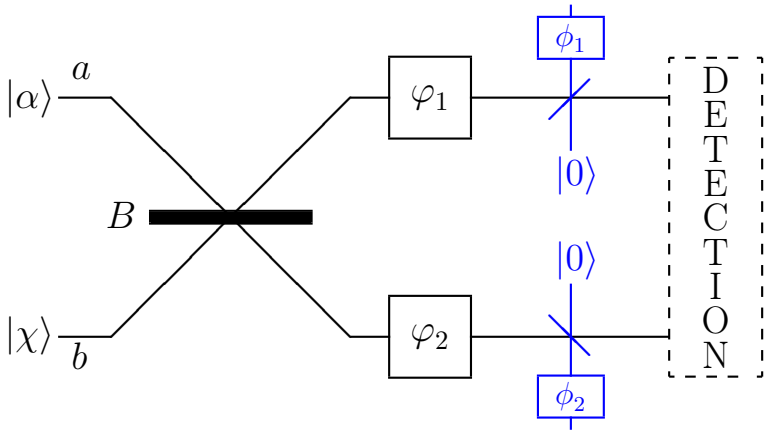

Figure 8.1: Modified interferometric setting for the analysis including linear losses. While most of this setup is identical to Fig. 7.1, two auxiliary beamsplitters, shown in blue, both with transmissivity η, are introduced to model linear losses in the two arms, here assumed to be identical in the two arms. The phase shifts ϕ_1 and ϕ_2 in the ancillary modes after the interaction with the main system do not change any physical quantity we are considering; they should be thought of as a mathematical trick to obtain a better bound for the quantum Fisher information.

Including losses means that an initial pure state is taken to a mixed state. Thus we cannot make use of Eq. (6.19) to calculate the quantum Fisher information. As there is no explicit expression for the Fisher information for mixed states, it becomes impossible to solve the optimization problem of finding the most sensitive input state analytically. To circumvent this problem we will use the following strategy.

First we will adapt to our setting an upper bound $\mathcal{C}_Q$ on the quantum Fisher information, introduced by Escher *et al.* [EdMFD11a, EdMFD11b]. Then we will explicitly calculate the quantum Fisher information $\mathcal{F}_Q$ for the input state we found to be optimal in the lossless case. A subsequent comparison of $\mathcal{C}_Q$ with $\mathcal{F}_Q$ will show that in the cases of practical interest, the results of Chap. 7 still hold.

In [EdMFD11a, EdMFD11b] an upper bound $\mathcal{C}_Q(\rho_0, \{\Pi_l(x)\})$ on the quantum Fisher information $\mathcal{F}_Q(\rho_0)$ was introduced. This bound depends on the Kraus operators $\{\Pi_l(x)\}$ used to describe the system's evolution; the Kraus operators depend on the parameter x to be estimated. While picking any set of Kraus operators will yield an upper bound, Escher *et al.* showed that there always exists a set of Kraus operators for which this bound is tight, i.e., yields the quantum Fisher information. This bound is given by [EdMFD11a]

$$\mathcal{C}_Q(\rho_0, \{\Pi_l(x)\}) = 4\left[\langle H_1(x)\rangle_0 - \langle H_2(x)\rangle_0^2\right], \tag{8.1}$$

where the expectation values are taken with respect to the state $\rho_0 = B\rho_{\text{in}}B^\dagger$ after the beamsplitter. The operators H_1 and H_2 are given in terms of the Kraus operators,

$$H_1(x) = \sum_l \frac{d\,\Pi_l^\dagger(x)}{dx}\frac{d\,\Pi_l(x)}{dx}, \tag{8.2}$$

$$H_2(x) = i\sum_l \frac{d\,\Pi_l^\dagger(x)}{dx}\Pi_l(x). \tag{8.3}$$

As in the Supplemental Material for [EdMFD11a], we choose our Kraus operators to be those defined in the number basis of the ancillary modes:

$$\Pi_{k,l} = \langle k_{E_1}, l_{E_2}|U|0_{E_1}, 0_{E_2}\rangle \tag{8.4}$$

$$= \sqrt{\frac{(1-\eta)^k}{k!}}\sqrt{\frac{(1-\eta)^l}{l!}}e^{i(\varphi_1 n_a+\phi_1 k)}\eta^{n_a/2}a^k\, e^{i(\varphi_2 n_b+\phi_2 l)}\eta^{n_b/2}b^l. \tag{8.5}$$

Here $n_a = a^\dagger a$ and $n_b = b^\dagger b$ are the number operators for the two primary modes. The joint system-environment unitary U corresponds the setting depicted in Fig. 8.1: It contains the phase shifts in modes a and b, respectively, followed by mixing at the beam splitters of a with ancillary mode E_1 and of b with ancillary mode E_2, and finally by phase shifts ϕ_1 and ϕ_2 of the two ancillary modes. The beam splitters have transmissivity η, and the modes E_1 and E_2 start out in the vacuum state.

We choose the additional ancillary phase shifts to be $\phi_1 = -(\gamma+1)\varphi_1$ and $\phi_2 = -(\gamma+1)\varphi_2$. The introduction of the additional parameter γ is inspired by [EdMFD11a], where it is interpreted as coming from the 'position' of the beamsplitter in the following sense. If there are no additional ancillary phase shifts and the loss beamsplitters are positioned after the phase shifters for modes a and b, as depicted in Fig. 8.1, then $\gamma = -1$. If, instead, the loss beamsplitters are positioned before the primary phase shifters, then $\gamma = 0$. To obtain a tighter bound for C_Q one optimizes over γ; the optimal value, it turns out, is in general not within the interval $[-1, 0]$. Hence, we offer a different interpretation of this parameter as merely a mathematical trick to obtain the best bound.

After some manipulation, which is summarized in App. B, we find

$$C_{Q,\gamma} = [1-(1+\gamma)(1-\eta)]^2 \langle \Delta N_d^2 \rangle + (1+\gamma)^2 \eta(1-\eta)\bar{N}_t, \tag{8.6}$$

where $\bar{N} = \bar{N}_a + \bar{N}_b$ is the mean total photon number entering the interferometer. The parameter γ can take any value without changing any physical quantity of the system. To get the best upper bound, we minimize Eq. (8.6) over all γ. As we discuss in App. B, the optimal choice for γ is

$$\gamma_{opt} = \frac{\langle \Delta N_d^2 \rangle}{(1-\eta)\langle \Delta N_d^2 \rangle + \eta \bar{N}_t} - 1. \tag{8.7}$$

Plugging this into Eq. (8.6), we find the optimized bound to be

$$C_{Q,\gamma_{opt}} = \frac{\eta \bar{N}_t \langle \Delta N_d^2 \rangle}{(1-\eta)\langle \Delta N_d^2 \rangle + \eta \bar{N}_t}. \tag{8.8}$$

Notice that, as mentioned above, $\bar{N}_t$ is considered to be constant when we are looking for optimal states, as this is the resource we are using. We can show that the

expression in Eq. (8.8) is a monotonically increasing function of the Fisher information for pure states, i.e., $\langle \Delta N_d^2 \rangle$ (see Chap. B). This means that to obtain an upper bound on the Fisher information with losses for all possible input states that obey our constraints, we can simply put into Eq. (8.8) the state we found to be optimal for the lossless case. We will do this for two cases we considered in the lossless setting. First we will focus on the case where the input in mode a is the coherent state $|\,\alpha\,\rangle$, as in Fig. 8.1.

8.1 Lossy interferometry with laser light

To get an upper bound for the quantum Fisher information with losses for the case where a coherent state $|\,\alpha\,\rangle$ is fed into mode a of the interferometer, we simply take Eq. (7.17) and plug it into Eq. (8.8) to get

$$C_{Q,\gamma_{opt}} = \frac{\eta(\bar{N}_b + \bar{N}_a)(2\bar{N}_a\langle \Delta p^2 \rangle + \bar{N}_b)}{(1-\eta)(2\bar{N}_a\langle \Delta p^2 \rangle + \bar{N}_b) + \eta(\bar{N}_a + \bar{N}_b)} = C_Q, \tag{8.9}$$

where

$$\langle \Delta p^2 \rangle = \bar{N}_b + \sqrt{\bar{N}_b(\bar{N}_b + 1)} + 1/2. \tag{8.10}$$

While we used a specific state (the optimal state for the lossless scenario) to obtain Eq. (8.9), we emphasize that the upper bound on the quantum Fisher information in Eq. (8.9) is an upper bound for all possible input states satisfying our constraints. More specifically, $C_{Q,\gamma_{opt}}$ is an upper bound on the quantum Fisher information for the states that are truly optimal in the lossy setting.

To proceed, we will now calculate the exact expression for the Fisher information of the state $|\,\alpha\,\rangle \otimes S_b(r)|0\rangle$. In general, there is no straightforward way to do this, as one needs to find the symmetric logarithmic derivative operator L, which unfortunately is only defined implicitly (see Chap. 6). Monras [Mon13] and Jiang [Jia14] independently derived manifestly equivalent explicit expressions for the Fisher information of arbitrary Gaussian states. As the state we are interested in is Gaussian, we will

make use of Jiang's work in the following.[1]

The Fisher information for an arbitrary Gaussian state is given by

$$F_Q = \frac{1}{2}\mathrm{tr}(\dot{\Gamma}\Phi) + 2\dot{d}^T\Gamma^{-1}\dot{d}\,. \tag{8.11}$$

Here d is the vector of first moments, i.e., $d^i = \mathrm{tr}[R^i\rho]$, and Γ is the covariance matrix, i.e., $\Gamma^{ij} = 2\,\mathrm{tr}[(R^i - d^i)\circ(R^j - d^j)\rho]$, with $R = (X_1, X_2, P_1, P_2)^T$ being the vector of canonical positions and momenta. The circle $\circ$ denotes the symmetric product (anticommutator), and thus this covariance matrix is twice the conventional one; this gives it the the nice feature that it is the unit matrix for the vacuum. In Eq. (8.11), the dot denotes a derivative of Γ with respect to the parameter we want to estimate, in our case, ϕ_d. The matrix Φ is defined implicitly by Γ and $\dot{\Gamma}$ through

$$\dot{\Gamma} = \Gamma\Phi\Gamma + J\Phi J\,, \tag{8.12}$$

where J is the fundamental antisymmetric symplectic matrix,

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\,. \tag{8.13}$$

If we do a symplectic basis change to the basis where Γ is diagonal (so-called standard form), i.e., $S\Gamma S^T = \Gamma_s$, we can calculate the symplecticly invariant quantity $\frac{1}{2}\mathrm{tr}(\dot{\Gamma}\Phi) = \frac{1}{2}\mathrm{tr}(\dot{\Gamma}_s\Phi_s)$, as Φ_s now has an explicit form in terms of Γ_s and $\dot{\Gamma}_s$ [Jia14]. Following this procedure, we are able to evaluate the quantum Fisher information for the state we are considering under losses:

$$\begin{aligned} F_Q &= \eta\bar{N}_b + \frac{\eta\bar{N}_a}{1-\eta+\eta/2\langle\Delta p^2\rangle} \\ &= \eta\bar{N}_b + \frac{\bar{N}_a\,\eta\left(1 + 2\,\bar{N}_b\,\eta + 2\eta\sqrt{\bar{N}_b(1+\bar{N}_b)}\right)}{1 + 4\,\bar{N}_b\,(1-\eta)\,\eta}. \end{aligned} \tag{8.14}$$

[1] While the expression for the Fisher information in [Mon13] is more compact, it involves calculating the (pseudo) inverse of a $(2N)^2 \times (2N)^2$ Matrix. Using Mathematica on a personal laptop, the case $N = 2$ already proved infeasible to do symbolically. The way presented in [Jia14] seems less computationally intense and was therefore preferred in this work.

Here $\langle \Delta p^2 \rangle$ is as in Eq. (8.10). Now we are left with comparing the two quantities C_Q and F_Q. Notice that $C_Q = F_Q$ in the absence of losses, i.e., $\eta = 1$, and also when all the photons are lost, i.e., $\eta = 0$.

To begin the comparison, we we write the two quantities in the forms

$$C_Q = \frac{\bar{N}_a + \bar{N}_b}{\dfrac{1-\eta}{\eta} + \dfrac{\bar{N}_a + \bar{N}_b}{2\bar{N}_a \langle \Delta p^2 \rangle + \bar{N}_b}}, \tag{8.15}$$

$$F_Q = \frac{\bar{N}_a + \bar{N}_b - \eta \bar{N}_b (1 - 1/2\langle \Delta p^2 \rangle)}{\dfrac{1-\eta}{\eta} + \dfrac{1}{2\langle \Delta p^2 \rangle}}. \tag{8.16}$$

The case of primary interest to us is when mode a carries many more photons than mode b, i.e.,

$$\bar{N}_a \gg \bar{N}_b. \tag{8.17}$$

This again is motivated by our initial considerations and ties in very well with the constraint we already had imposed on our setup: We fixed the input state in mode a to be a coherent state, as the main power source of any practical interferometer will be a commercially available laser; since the lasers used in high-precision interferometry will carry many more photons (per averaging time) than any nonclassical state currently feasible in the laboratory, the assumption of $\bar{N}_a \gg \bar{N}_b$ is natural. Under this assumption, it is clear that both C_Q and F_Q are close to

$$I_Q = \frac{\bar{N}_a + \bar{N}_b}{\dfrac{1-\eta}{\eta} + \dfrac{1}{2\langle \Delta p^2 \rangle}}. \tag{8.18}$$

Moreover, for all values of $\bar{N}_a$ and $\bar{N}_b$, both C_Q and F_Q are smaller than I_Q. What we do now is to investigate how much bigger and how much smaller under the assumption (8.17).

We now write the two quantities in the forms

$$\frac{1}{C_Q} = \frac{1}{I_Q} + \frac{1}{2\bar{N}_a\langle\Delta p^2\rangle + \bar{N}_b)} - \frac{1}{2\langle\Delta p^2\rangle(\bar{N}_a + \bar{N}_b)}$$
$$\simeq \frac{1}{I_Q}\left[1 + \frac{\bar{N}_b}{\bar{N}_a}\frac{I_Q}{2\bar{N}_a\langle\Delta p^2\rangle}\left(1 - \frac{1}{2\langle\Delta p^2\rangle}\right)\right], \tag{8.19}$$

$$F_Q = I_Q\left[1 - \frac{\eta\bar{N}_b}{\bar{N}_a + \bar{N}_b}\left(1 - \frac{1}{2\langle\Delta p^2\rangle}\right)\right]. \tag{8.20}$$

In the approximate expression for $1/C_Q$, we keep only the leading-order terms in $\bar{N}_b/\bar{N}_a$. We can quantify the difference between the two quantities in terms of the relative difference

$$\frac{C_Q - F_Q}{C_Q} \simeq (1-\eta)\frac{\bar{N}_b}{\bar{N}_a}\frac{I_Q}{\bar{N}_a}\left(1 - \frac{1}{2\langle\Delta p^2\rangle}\right)^2$$
$$\simeq \frac{\bar{N}_b}{\bar{N}_a}\left(1 - \frac{1}{2\langle\Delta p^2\rangle}\right)^2 \frac{\eta}{1 + \dfrac{\eta}{1-\eta}\dfrac{1}{2\langle\Delta p^2\rangle}}. \tag{8.21}$$

This form makes clear that the upper bound on the Fisher information coincides nearly exactly with the Fisher information for putting a squeezed vacuum state into the second mode, under the operating conditions of a typical high-power interferometer. Indeed, once one realizes that both C_Q and F_Q differ negligibly from I_Q when $\bar{N}_a \gg \bar{N}_b$, one might as well regard I_Q as the bound on performance and compare F_Q to it. Then the relevant relative difference is

$$\frac{I_Q - F_Q}{I_Q} = \frac{\eta\bar{N}_b}{\bar{N}_a + \bar{N}_b}\left(1 - \frac{1}{2\langle\Delta p^2\rangle}\right) \simeq \eta\frac{\bar{N}_b}{\bar{N}_a}\left(1 - \frac{1}{2\langle\Delta p^2\rangle}\right). \tag{8.22}$$

The first expression is exact. The second applies when $\bar{N}_a \gg \bar{N}_b$ and shows that in the high-power limit, we may effectively regard I_Q as the Fisher information of this sort of interferometer. This tells us that increasing the squeezing improves the resolution as long as $2\langle\Delta p^2\rangle \lesssim \eta/(1-\eta)$. Of course, to increase the squeezing to this critical value requires that the losses be large enough that $\bar{N}_a \gg \bar{N}_b = \sinh^2 r \simeq \langle\Delta p^2\rangle/2 \simeq \eta/(1-\eta)$, but this will generally be true for any conceivable losses in a high-power interferometer. Once squeezing exceeds the critical value by any substantial amount, the Fisher information limits to

$$I_Q \simeq \frac{\eta}{1-\eta}\bar{N}_a, \tag{8.23}$$

which has shot-noise scaling, but with an enhancement by the factor $\eta/(1-\eta)$ relative to the shot-noise limit. This is the same form as the bounds on sensitivity discussed in [KDD10, KSD11, EdMFD11a].

This concludes our demonstration that under the assumption (8.17), the squeezed vacuum state is so close to being the optimal choice to use in practical interferometry that it is not worth considering any other possibility. To be useful in practice, however, one needs to know what the optimal measurement scheme and estimator are. A natural candidate is the standard linear estimator when doing a differenced photon count, as it worked in the lossless case for the regime of high laser power we are now interested in.

The relevant quantity to determine the precision of this estimator is given by the error propagation formula. In our case the variance of the estimator is $\langle \Delta N_d^2 \rangle / \left| \partial_{\phi_d} \langle N_d \rangle \right|^2$. In general, this quantity will be dependent on the parameter one wants to estimate, here ϕ_d. Hence there will be an optimal operating point for the interferometer. In this case the optimal point is $\phi_d = \pi/2$. In terms of the coherent state amplitude α and the squeezing parameter r, we have

$$\frac{\langle \Delta N_d^2 \rangle}{\left| \partial_{\phi_d} \langle N_d \rangle \right|^2} = \frac{\alpha^2 e^{-2r} + \sinh^2 r + \dfrac{1-\eta}{\eta}\alpha^2 + \dfrac{1-\eta}{\eta}\sinh^2 r}{(\alpha^2 - \sinh^2)^2}; \tag{8.24}$$

Making again the assumption of high laser power, i.e., $\alpha^2 \gg \sinh^2 r$, this becomes

$$\frac{e^{-2r} + \dfrac{1-\eta}{\eta}}{\alpha^2} \simeq \frac{1}{I_Q}. \tag{8.25}$$

Comparing this to $1/F_Q$, we see again that we have found the optimal estimation procedure.

An analysis in the same spirit has been carried out in [DDanBS13], to demonstrate the optimality of LIGO. The assumptions in this work, however, differ from the ones we have made and hence provide an independent confirmation of the usefulness of squeezed-light interferometry.

8.2 Lossy interferometry

Analogous to Sec. 7.2 in Chap. 7, we will now repeat this analysis with one assumption relaxed: Instead of fixing the input state in mode a to be a coherent state, in the following we only assume the input state to be of a product form.

Again we make use of the upper bound (8.8). Given Eq. (7.35), the upper bound on the quantum Fisher information of all states that obey the constraints we are considering is

$$C_{Q,\gamma_{opt}} = \frac{\eta\bar{N}_t(\bar{N}_t+2)}{(1-\eta)(\bar{N}_t+2)+\eta} = \frac{\eta^2\bar{N}_t(\bar{N}_t+2)}{\eta(1-\eta)\bar{N}_t+\eta(2-\eta)} = C_Q. \tag{8.26}$$

To calculate the exact Fisher information in the lossy interferometer for the two vacuum squeezed states in Eq. (7.34), we again make use of Jiang's formalism [Jia14]. It turns out that the calculation in this case is easier than the one we had to do when we were considering a coherent state in the first input port: the second term in Eq. (8.11) is zero because all the first moments are zero, and the first term is not so hard to calculate because the symplectic eigenvalues are degenerate, meaning that we do not have to perform a full symplectic diagonalisation procedure [Jia14]. The result for the quantum Fisher information is

$$F_Q = \frac{\eta^2\bar{N}_t(\bar{N}_t+2)}{\eta(1-\eta)\bar{N}_t+1}. \tag{8.27}$$

Comparing this to Eq. (8.26), we see that the two quantities differ only in their denominator. They are equal when $\eta = 0$ or $\eta = 1$.

If we let $\eta = 1-\epsilon$, we have $\eta(2-\eta) = 1-\epsilon^2$, so when losses are small, i.e., $\epsilon \ll 1$, C_Q and F_Q differ only at second order in ϵ. This was to be expected, as it is a general feature of the bound (8.8) that in the regime of low losses, $\mathcal{C}_Q$ is nearly equal to F_Q. More interesting is the case when the losses are not negligible. Here we need the input power to be strong enough that $\eta(1-\eta)\bar{N}_t \gg 1$, in which case C_Q and F_Q are both nearly equal to $(\bar{N}_t+2)\eta/(1-\eta)$. Thus, under these circumstances, the two squeezed vacua are essentially the optimal input product states for interferometry, even in the presence of losses.

In this chapter the we extended the results of the previous chapter to the case where the interferometer is subject to photon losses. We have shown that when the interferometer is powered mainly by laser light, squeezed vacuum is still the optimal choice for the second input port no matter how large the losses are. When allowing an arbitrary product state to enter the interferometer, squeezed vacua in both input ports are optimal if their mean photon number is large enough.

Chapter 9

Conclusion

In the first half of this thesis, we considered several entropic measures of nonclassical correlations. All of these measures purport to quantify the degree of nonclassicality in bipartite quantum states. Just as important as the degree of nonclassicality, however, is the boundary these measures set between quantum and classical states. For the measures employing an unconditioned measurement, a joint state is classical if and only if it is diagonal in a product basis. Both forms of *discord*, however, consider a joint state to be classical if it is diagonal in a conditional product basis pointing from A to B; clearly this boundary between quantum and classical is a consequence of a conditioned measurement strategy. All the measures differ in the "amount" of nonclassicality they assign to states they deem not classical.

The boundary between quantum and classical is important by itself, first, because it can usually be extended to multipartite systems even when the measure of nonclassical correlations is not so easily extended and, second, because it serves as the basis for interesting questions about quantum-information processing. As mentioned in the Introduction, Eastin [Eas10] has recently investigated whether *concordant* computations can be simulated classically. A concordant computation is one such that after every elementary gate, the state of the whole computer is diagonal in a product basis. The entire computation is just a matter of changing the product basis, yet Eastin finds it difficult in general to simulate such computations efficiently.

This discussion raises at least two other questions, the first of which concerns the boundary between quantum and classical. The boundaries induced by the nonclassical-correlation measures considered here are natural—naïve is perhaps a better word—in that the classicality of a state is defined in terms of properties of its eigenvectors. This is quite different from the distinction between separable and entangled states [Wer89], which pays no attention to the properties of the joint state's eigenvectors; a state is separable if and only if it has an ensemble decomposition—not an eigendecomposition—in terms of product states. The boundaries for the measures discussed in this paper do require a classical state to have unentangled eigenvectors, but they are more restrictive than saying that a state is classical if its eigenvectors are product states. There are orthonormal bases of product states that are neither product bases nor conditional product bases [BDF$^+$99]. The boundaries considered here, imposed by the measurement strategies, clearly have to do with wanting the product states in a joint eigenbasis to persist into the marginal eigenbasis of one or both subsystems. A natural question is wheher there is some other way of setting the quantum-classical boundary so that a joint state is classical if and only if it has unentangled eigenvectors?

Any measure of nonclassical correlations assigns a number to a joint quantum state; the second question has to do with what this "amount" means. The demon-based measures have an operational interpretation as the work deficit suffered by local demons that are required to work only with the subsystems. Recently, there have been two closely related proposals for operational interpretations of quantum discord in terms of state-merging protocols [CAB$^+$11, MD11]. Quantum discord has also been connected to the entanglement loss when mixed states are created from entangled states followed by entanglement distillation from those mixed states [CdOF11]. Such operational interpretations are essential to understand the meaning of a measure of nonclassical correlations. Whenever a measure of nonclassical correlations is proposed, the amount of nonclassicality must ultimately gain meaning through an operational interpretation.

The second half of this dissertation discussed a topic from a different branch

of quantum information theory, motivated by more practical considerations. Ever since it was shown that squeezed states can enhance the sensitivity of a laser interferometer [Cav81], schemes have been proposed trying to improve on this idea by presenting a setup that is even more efficient in the use of the resource at hand. It is of particular interest to find the nonclassical input state to the interferometer that performs best under various constraints. The *N00N* state was hallowed by many as the 'optimal' state to distinguish a small phase shift between two modes of light. In fact the *N00N* state is only optimal under the constraint that the interferometer deals with a fixed number of photons. While the photon number is certainly a crucial resource, the relevant quantity corresponding to the energy consumed in 'one run' of the experiment is the *mean* photon number needed to create the current input state. The only reason to impose a constraint of fixed photon number in an optical setup is to ease the analysis of finding the optimal state. This, however, defies the very purpose of the undertaking: any state that performs better than the optimal state for a fixed photon number will be more relevant for estimation purposes regardless of whether it is provably optimal or not. In this dissertation we took a different approach. While we also imposed additional constraints in our search for an optimal state, our constraints are well motivated by practical considerations.

First, we looked at the situation where an interferometer is mainly powered by strong laser light, modeled by a coherent state entering the interferometer's first input port. We showed that given a constraint on the total mean number of photons, the optimal state to put into the secondary input port is a squeezed vacuum state, regardless of the relative mean photon numbers of the two inputs. A question that begs for further attention here is whether one could come up with a simple optimal estimator when the squeezed light carries as many or more photons than the coherent input. Extending our analysis to include photon losses, we demonstrated that such a setup, currently being implemented at LIGO, is indeed optimal for practical purposes, confirming the findings of Ref. [DDanBS13] from a different viewpoint.

The second practical situation we explored is that where the input to the interferometer is of product form. Product states are typically easier to create than more

general states, as one does not have to entangle the two optical modes. Under this assumption of a product input, we demonstrated that the optimal choice for the input state is to use squeezed vacua in both input ports of the interferometer. It turns out that this state is not only easier to produce than a *N00N* state, it even performs slightly better than a *N00N* state, in terms of mean photon number. Extending our analysis, we also showed that our state is far more robust to photon losses than a *N00N* state. Moreover, our findings indicate that modal entanglement does not play a crucial role in quantum-enhanced interferometry.

An outstanding question is still that of the optimal state under nothing but a mean photon number constraint. While Fisher information cannot be used to find the optimal state in this context, a question of considerable interest is whether there exist other techniques allowing for the analysis of this general setting where the input is not required to be of product form. It is interesting to note that a two-mode squeezed vacuum as the input to the interferometer, though highly entangled, yields the same sensitivity that we found [STL+14]: the first beam splitter of the interferometer unentangles the state producing again single-mode-squeezed vacua in both modes with a different phase relation relative to our input state. A simpler objective than finding the state with the optimal performance perhaps would be to come up with one single state that performs better than the dual squeezed vacua that we found. Until such a state is found, we advocate the use of squeezing in interferometers as a more reasonable approach to obtain a quantum enhancement than focusing on the improvement of techniques using *N00N* or other exotic states.

Appendix A

Supplemental material to the correlation-measures chapters

A.1 The POVM inequality

A quantum state written in its eigenbasis,

$$\rho = \sum_{\alpha} \lambda_{\alpha} |e_{\alpha}\rangle\langle e_{\alpha}| \ , \tag{A.1}$$

is subjected to a POVM with elements E_j. This gives outcome probabilities

$$p_j = \mathrm{tr}(E_j \rho) = \sum_{\alpha} \lambda_{\alpha} q_{j|\alpha} \ , \tag{A.2}$$

where $q_{j|\alpha} = \langle e_{\alpha}|E_j|e_{\alpha}\rangle$ is the probability for outcome j given state $|e_{\alpha}\rangle\langle e_{\alpha}|$. We define

$$\mu_j \equiv \mathrm{tr} E_j = \sum_{\alpha} q_{j|\alpha} \ , \tag{A.3}$$

which implies that $q_{j|\alpha}/\mu_j$ is a normalized probability distribution on α.

Now let $f(x) = -x \log x$ and proceed through the following steps:

$$\begin{aligned} H(p_j) &= \sum_j f(p_j) \\ &= \sum_j f\left(\sum_\alpha \frac{q_{j|\alpha}}{\mu_j} \lambda_\alpha \mu_j \right) \\ &\geq \sum_{\alpha,j} \frac{q_{j|\alpha}}{\mu_j} f(\lambda_\alpha \mu_j) \qquad \text{(A.4)} \\ &= -\sum_{\alpha,j} q_{j|\alpha} \lambda_\alpha \log(\lambda_\alpha \mu_j) \\ &= -\sum_{\alpha,j} q_{j|\alpha} \lambda_\alpha \log \lambda_\alpha - \sum_{\alpha,j} \lambda_\alpha q_{j|\alpha} \log \mu_j \\ &= -\sum_\alpha \lambda_\alpha \log \lambda_\alpha - \sum_j p_j \log \mu_j \qquad \text{(A.5)} \\ &= S(\rho) - \sum_j p_j \log(\mathrm{tr} E_j) \; . \qquad \text{(A.6)} \end{aligned}$$

The inequality (A.4) uses that $f(x)$ is a concave function and that $q_{j|\alpha}/\mu_j$ is a probability distribution over α. The step leading to Eq. (A.5) uses that $q_{j|\alpha}$ is a normalized distribution over j.

A.2 Nonnegativity and ordering of the WPM measure and discord

In this Appendix we show that the WPM measure is bounded below by quantum discord and that the quantum discord is nonnegative. We proceed by assuming that systems A and B are subjected to measurements described by POVMs with rank-one POVM elements E_a and F_b, as in measurement strategy (b). For convenience, we use the fact that any set of POVM elements can be extended to rank-one orthogonal projection operators in a space of higher dimension, an extension called the Naimark extension [Nai40]. Formally, we have an orthonormal basis $|e_a\rangle$ in the higher-dimensional space such that $E_a = P_A |e_a\rangle\langle e_a| P_A$, where P_A projects onto the original Hilbert space of system A, which is the space where the states of system A

live; similarly, there is an orthonormal basis $|f_b\rangle$ such that $F_b = P_B|f_b\rangle\langle f_b|P_B$, where P_B projects onto the original Hilbert space of system B.

We can write the joint probability for outcomes a and b as

$$p_{ab} = \mathrm{tr}(E_a \otimes F_b \rho_{AB}) = \mathrm{tr}\Big(P_A|e_a\rangle\langle e_a|P_A \otimes P_B|f_b\rangle\langle f_b|P_B \rho_{AB}\Big) = \langle e_a, f_b|\rho_{AB}|e_a, f_b\rangle \ , \tag{A.7}$$

where the last equality follows because ρ_{AB} lives in the original Hilbert space of A and B, so we can discard the projectors into that space. Other results we need below include

$$\rho_{B|a} = \frac{\mathrm{tr}_A(E_a \rho_{AB})}{p_a} = \frac{\langle e_a|\rho_{AB}|e_a\rangle}{p_a} \ , \tag{A.8}$$

where

$$p_a = \mathrm{tr}(E_a \rho_A) = \langle e_a|\rho_A|e_a\rangle \ . \tag{A.9}$$

We now extend the joint state ρ_{AB} to a space with two additional systems, C and D. We let C have dimension equal to the number of outcomes a, with an orthonormal basis $|g_a\rangle$, we let D have dimension equal to the number of outcomes b, with orthonormal basis $|h_b\rangle$. The extended state,

$$\rho'_{ABCD} = \sum_{a,a',b,b'} |e_a, f_b\rangle\langle e_a, f_b|\rho_{AB}|e_{a'}, f_{b'}\rangle\langle e_{a'}, f_{b'}| \otimes |g_a\rangle\langle g_{a'}| \otimes |h_b\rangle\langle h_{b'}| \ , \tag{A.10}$$

can be regarded as a state where systems C and D record the measurement outcomes in their orthonormal bases. Notice that the entropy of the extended state is

$$S'(A, B, C, D) = S(A, B) \ . \tag{A.11}$$

The proof follows from two applications of the strong-subadditivity property of von Neumann entropy [NC00]. Various marginals of the extended state and their von

Neumann entropies enter into the proof:

$$\rho'_{ABD} = \sum_{a,b,b'} p_a |e_a\rangle\langle e_a| \otimes |f_b\rangle\langle f_b|\rho_{B|a}|f_{b'}\rangle\langle f_{b'}| \otimes |h_b\rangle\langle h_{b'}| \;, \tag{A.12}$$

$$\rho'_{AC} = \sum_{a,a'} |e_a\rangle\langle e_a|\rho_A|e_{a'}\rangle\langle e_{a'}| \otimes |g_a\rangle\langle g_{a'}| \;, \tag{A.13}$$

$$\rho'_{BD} = \sum_{b,b'} |f_b\rangle\langle f_b|\rho_B|f_{b'}\rangle\langle f_{b'}| \otimes |h_b\rangle\langle h_{b'}| \;, \tag{A.14}$$

$$\rho'_{AB} = \sum_{a,b} p_{ab} |e_a, f_b\rangle\langle e_a, f_b| \;, \qquad \rho'_A = \sum_a p_a |e_a\rangle\langle e_a| \;, \tag{A.15}$$

$$\rho'_B = \sum_a p_b |f_b\rangle\langle f_b| \;, \tag{A.16}$$

$$\rho'_C = \sum_a p_a |g_a\rangle\langle g_a| \;, \qquad \rho'_D = \sum_a p_b |h_b\rangle\langle h_b| \;. \tag{A.17}$$

These have von Neumann entropies

$$S'(A,B,D) = H(A) + \sum_a p_a S(B|a) \;, \tag{A.18}$$

$$S'(A,C) = S(A) \;, \qquad S'(B,D) = S(B) \;, \tag{A.19}$$

$$S'(A,B) = H(A,B) \;, \qquad S'(A) = S'(C) = H(A) \;, \tag{A.20}$$

$$S'(B) = S'(D) = H(B) \;. \tag{A.21}$$

The proof now comes in a rush. Recalling Eq. (4.37), we use the above results to write

$$\begin{aligned} [S(A:B) - H(A:B)] - [H_{\{E_a\}}(B|A) - S(B|A)] & \\ = -S'(A,B,D) - S'(B) + S'(A,B) + S'(B,D) & \\ = S'(A|B) - S'(A|B,D) \geq 0 \;. & \end{aligned} \tag{A.22}$$

The inequality is precisely the expression of strong subadditivity for systems A, B, and D. It shows that

$$\mathcal{M}_{1b}^{(\mathrm{WPM})} \geq \mathcal{M}_{2c}^{(\mathrm{discord})} = \mathcal{D}(A \to B) \;. \tag{A.23}$$

Concentrating now on discord, we write

$$\begin{aligned} H_{\{E_a\}}(B|A) - S(B|A) &= -S'(A,B,C,D) - S'(A) + S'(A,B,D) + S'(A,C) \\ &= S'(B,D|A) - S'(B,D|A,C) \geq 0 \;, \end{aligned} \tag{A.24}$$

where again the inequality is strong subadditivity, this time for systems BD, A, and C. This inequality shows that discord is always nonnegative.

The equality conditions for strong subadditivity [HJPW04] can be applied to inequalities (A.22) and (A.24). From the latter inequality, we learn that $\mathcal{M}_{2c}^{(\text{discord})} = \mathcal{D}(A \to B)$ is zero if and only if ρ_{AB} is diagonal in a conditional product basis pointing from A to B. Datta [Dat10], in the proof of his Theorem 2, has shown how to use the equality conditions for strong subadditivity to obtain this necessary and sufficient condition for zero discord. From (A.22), we learn that $\mathcal{M}_{1b}^{(\text{WPM})} = \mathcal{M}_{2c}^{(\text{discord})} = \mathcal{D}(A \to B)$ if and only ρ_{AB} is diagonal in a conditional product basis pointing from B to A. Combining these two results, we have that the WPM measure is zero if and only if ρ_{AB} is diagonal in a product basis.

A.3 Projective measurements vs. POVMs for WPM and discord

In this Appendix, we elaborate an example given by WPM [WPM09], which exhibits a joint state that requires the use of rank-one POVMs, not just orthogonal-projection-valued measurements, to maximize the classical mutual information in evaluating the WPM measure. We extend these results to show that for the same states, rank-one POVMs are required for evaluating the discord.

Consider a joint state

$$\rho_{AB} = \sum_{j=1}^{d_B} p_j \rho_j \otimes |e_j\rangle\langle e_j| = \sum_{j=1}^{d_B} p_j \rho_j \otimes P_j \ , \tag{A.25}$$

where the states $|e_j\rangle$ make up an orthonormal basis for system B. The marginal states are given by

$$\rho_A = \sum_j p_j \rho_j \ , \qquad \rho_B = \sum_j p_j P_j \ , \tag{A.26}$$

and this gives

$$S(B) = H(p_j) \ , \tag{A.27}$$

$$S(A,B) = H(p_j) + \sum_j p_j S(\rho_j) \ , \tag{A.28}$$

$$S(A|B) = S(A,B) - S(B) = \sum_j p_j S(\rho_j) \ , \tag{A.29}$$

$$S(B|A) = S(A,B) - S(A) = H(p_j) + \sum_j p_j S(\rho_j) - S(A) \ , \tag{A.30}$$

$$S(A:B) = S(B) - S(B|A) = S(A) - \sum_j p_j S(\rho_j) \ . \tag{A.31}$$

The quantum mutual information is the Holevo quantity for the ensemble of states ρ_j with probabilities p_j. This is not surprising because ρ_{AB} describes a situation where B sends a message to A: the message has the letters j, with probabilities p_j; B keeps a record of the message in the orthogonal states $|e_j\rangle$ and encodes the letters in the states ρ_j.

The state (A.25) has zero discord when communication goes from B to A, i.e., $\mathcal{D}(B \to A) = 0$, because ρ_{AB} is diagonal in a conditional product basis pointing from B to A. Generally, it has nonzero discord, $\mathcal{D}(A \to B)$, when communication goes from A to B. The results of Appendix A.2 show that the WPM measure equals $\mathcal{D}(A \to B)$ for such states. We return to discord below; for now, we focus on the WPM measure. Given any unconditioned, local measurements on A and B, the joint probability for results a and b is

$$p_{ab} = \sum_{j=1}^{d_B} p_j p_{a|j} p_{b|j} \ , \tag{A.32}$$

where $p_{a|j} = \mathrm{tr}(E_a \rho_j)$ and $p_{b|j} = \mathrm{tr}(F_b P_j)$. We can think of p_{ab} as the marginal of a joint distribution for a, b, and j:

$$p_{abj} = p_j p_{a|j} p_{b|j} \ . \tag{A.33}$$

That A and B are conditionally independent means that $p_{a|bj} = p_{a|j}$, which implies that $H(A|J) = H(A|B,J)$. Thus we have

$$H(A:J) - H(A:B) = H(A|B) - H(A|J) = H(A|B) - H(A|B,J) \geq 0 \ , \tag{A.34}$$

where the final inequality follows from classical strong subadditivity, which says that additional conditioning cannot increase the entropy. Measuring B in the eigenbasis $|e_j\rangle$ gives $H(A:B) = H(A:J)$, so we can conclude that the maximum mutual information is attained on this measurement. The WPM measure reduces to a form that only requires a maximization over the measurement on A:

$$\mathcal{M}_{1b}^{(\mathrm{WPM})} = S(A:B) - \max_{(\mathrm{b})} H(A:B) = S(A:B) - \max_{\{E_a\}} H(A:J) \; . \tag{A.35}$$

We now proceed to specialize the joint state ρ_{AB} in four ways. First, we assume that system A is a qubit and that the ensemble states,

$$\rho_j = \frac{1}{2}(I_A + \boldsymbol{\sigma} \cdot \boldsymbol{n}_j) \; , \tag{A.36}$$

are pure; i.e., the vectors $\boldsymbol{n}_j$ are unit vectors. The measurement on A is described by rank-one POVM elements

$$E_a = q_a(I_A + \boldsymbol{\sigma} \cdot \boldsymbol{m}_a) \; , \tag{A.37}$$

where the vectors $\boldsymbol{m}_a$ are unit vectors. The completeness relation for the POVM implies that the quantities q_a make up a normalized probability distribution and that

$$\sum_a q_a \boldsymbol{m}_a = 0 \; . \tag{A.38}$$

The probability for result a, given state ρ_j, is

$$p_{a|j} = \mathrm{tr}(E_a \rho_j) = q_a(1 + \boldsymbol{n}_j \cdot \boldsymbol{m}_a) \; , \tag{A.39}$$

and the joint probability for results a and j is

$$p_{aj} = \mathrm{tr}(\rho_{AB} E_a \otimes P_j) = p_j \mathrm{tr}(E_a \rho_j) = p_j q_a (1 + \boldsymbol{n}_j \cdot \boldsymbol{m}_a) \; . \tag{A.40}$$

The second specialization is to assume that the states ρ_j are distributed so that

$$\sum_{j=1}^{d_B} p_j \boldsymbol{n}_j = 0 \qquad \Longleftrightarrow \qquad \rho_A = \sum_{j=1}^{d_B} p_j \rho_j = \frac{1}{2} I_A \; . \tag{A.41}$$

With this assumption we have that the probability for result a is $p_a = q_a$ and thus that $S(B|A) = H(p_j) - S(A) = H(p_j) - 1$ and $S(A:B) = S(A) = 1$. The classical mutual information is

$$H(A:J) = H(A) - H(A|J) = \sum_{j,a} p_{aj} \log(p_{a|j}/p_a) = \sum_a q_a F(\boldsymbol{m}_a) \ , \tag{A.42}$$

where we define the function

$$F(\boldsymbol{m}) \equiv \sum_{j=1}^{d_B} p_j (1 + \boldsymbol{n}_j \cdot \boldsymbol{m}) \log(1 + \boldsymbol{n}_j \cdot \boldsymbol{m}) \ . \tag{A.43}$$

The WPM measure is now given by $\mathcal{M}_{1b}^{(\mathrm{WPM})} = 1 - \widetilde{H}(A:J)$, where

$$\widetilde{H}(A:J) \equiv \max_{\{E_a\}} H(A:J) = \max_{\{q_a, \boldsymbol{m}_a\}} \sum_a q_a F(\boldsymbol{m}_a) \ . \tag{A.44}$$

Before going on to the third specialization, let's consider the quantum discord when conditioning on A. We again make the first two specializations: a joint state of the form (A.25), with A being a qubit, and the states ρ_j being the pure states (A.36), distributed according to Eq. (A.41). We measure the POVM (A.37) on A. The probability for result a is $p_a = q_a$, and the state of B, conditioned on result a, is

$$\rho_{B|a} = \frac{\mathrm{tr}_A(E_a \rho_{AB})}{p_a} = \sum_{j=1}^{d_B} \frac{p_j p_{a|j}}{q_a} P_j = \sum_{j=1}^{d_B} p_j (1 + \boldsymbol{n}_j \cdot \boldsymbol{m}_a) P_j \ , \tag{A.45}$$

which has quantum entropy

$$S(B|a) = -\sum_{j=1}^{d_B} p_j (1 + \boldsymbol{n}_j \cdot \boldsymbol{m}_a) \log\big(p_j (1 + \boldsymbol{n}_j \cdot \boldsymbol{m}_a)\big) \ . \tag{A.46}$$

The conditional classical entropy that goes into the definition (4.38) of discord becomes

$$\widetilde{H}(B|A) = \min_{\{E_a\}} \sum_a p_a S(B|a) = H(p_j) - \max_{\{q_a, \boldsymbol{m}_a\}} \sum_a q_a F(\boldsymbol{m}_a) = H(p_j) - \widetilde{H}(A:J) \ , \tag{A.47}$$

so the discord is given by

$$\mathcal{D}(A \to B) = \mathcal{M}_{2c}^{(\mathrm{discord})} = \widetilde{H}(B|A) - S(B|A) = 1 - \widetilde{H}(A:J) = \mathcal{M}_{1b}^{(\mathrm{WPM})} \ . \tag{A.48}$$

As expected, the discord is the same as the WPM measure for this set of joint states.

Our third specialization is to assume that the ensemble probabilities are all equal, i.e., $p_j = 1/d_B$, and the fourth, needed to work out examples, is that the vectors $\boldsymbol{n}_j$ are symmetrically distributed, pointing to the vertices of an equilateral triangle or of a regular polyhedron. Within this configuration, we first maximize the function $F(\boldsymbol{m})$. The high degree of symmetry guarantees that the extrema of $F(\boldsymbol{m})$ occur along the symmetry axes of the triangle or polyhedron. Having determined the maximum value, $F_{\max}$, one knows that this maximum is an upper bound on $\widetilde{H}(A:J)$. The high degree of symmetry further guarantees that one can make up a POVM out of the directions $\boldsymbol{m}_a$ that give the maximum value, with q_a chosen to be the same for all these directions; since this choice achieves the bound, one has $\widetilde{H}(A:J) = F_{\max}$. Moreover, if no two of the directions $\boldsymbol{m}_a$ are oppositely directed, the upper bound cannot be achieved with a projection-valued measurement. One ends up knowing, first, the common value of the WPM measure and quantum discord and, second, that the optimal measurement cannot be described by orthogonal projection operators.

For a triangle (tetrahedron) of states, the maximum value of F is attained on the vectors that are directed opposite to the vectors that define the state. The maximum value is $F_{\max} = \log \frac{3}{2}$ for the triangle and $F_{\max} = \log \frac{4}{3}$ for the tetrahedron. The optimal measurement is the trine (tetrahedron) measurement that uses the triangle (tetrahedron) dual to the state triangle (tetrahedron). Both the WPM measure and discord are equal to $1 - \log \frac{3}{2} = \log \frac{4}{3} = 0.415$ for the triangle and to $1 - \log \frac{4}{3} = \log \frac{3}{2} = 0.585$ for the tetrahedron.

We stress that the examples in this Appendix require that subsystem B have three or more Hilbert-space dimensions. These examples thus do not exclude the possibility that projection-valued-measures suffice for WPM and discord for a pair of qubits.

It is worth noting that if one asks about demon discord, the quantity one needs to evaluate, instead of being the classical conditional entropy (A.47), is the classical

joint entropy (4.49). Making the first two specializations gives a demon discord

$$\mathcal{M}_{3c}^{(\mathrm{dd})} = \min_{\{q_a, \boldsymbol{m}_a\}} \left(H(q_a) - \sum_a q_a F(\boldsymbol{m}_a) \right) . \tag{A.49}$$

The additional contribution from $H(q_a)$ prejudices this minimum toward using a smaller number of outcomes for the measurement on A and thus toward orthogonal-projection-valued measurements. Indeed, for the triangle of states, with a trine measurement made in the dual triangle, the argument of the minimum is equal to 1. For an orthogonal-projection-valued measurement, with $\boldsymbol{m}_1$ pointing toward one vertex of the state triangle and $\boldsymbol{m}_2$ in the opposite direction, the argument is equal to $\frac{4}{3} - \frac{1}{2}\log 3 = 0.541$, which thus becomes the demon discord. For this joint state, the optimal measurement for demon discord is orthogonal-projection-valued.

A.4 Demon-based measures and rank-one POVMs

In this Appendix, we modify the formula for the net classical work to show that the local demons cannot do worse in terms of net classical work by restricting themselves to rank-one POVMs.

We use the general measurement formalism of Sec. 4.2.2, which allows us to do strategies (b) and (c) simultaneously. The state of system A after a measurement yields result a is

$$\rho_{A|a} = \frac{\mathcal{A}_a(\rho_A)}{p_a} , \tag{A.50}$$

and the state of system B after measurements that yield outcomes a and b is given by

$$\rho_{B|ab} = \frac{\mathcal{B}_{b|c(a)}(\rho_{B|a})}{p_{b|a}} . \tag{A.51}$$

As the systems are transformed to the maximally mixed state, the local demons can extract work

$$W^+ = \log(d_A d_B) - \sum_a p_a S(\rho_{A|a}) - \sum_{a,b} p_{ab} S(\rho_{B|ab}) . \tag{A.52}$$

The cost of erasing the measurement record, given communication between the demons, is $W^- = H(A,B)$, giving a net classical work

$$W_c = W^+ - W^- = \log(d_A d_B) - H(A,B) - \sum_a p_a S(\rho_{A|a}) - \sum_{a,b} p_{ab} S(\rho_{B|ab}) \ . \quad \text{(A.53)}$$

The measure of nonclassical correlations requires maximizing W_c over all possible measurements.

After finishing the first round of measurements, instead of extracting work, the local demons can make further measurements in the eigenbases of $\rho_{A|a}$ and $\rho_{B|ab}$. The overall measurement is now described by rank-one POVMs. After these measurements, the subsystems are left in pure states, so the local demons can extract work $W^+ = \log(d_A d_B)$ as the systems are transformed to the maximally mixed state, but they have a more detailed measurement record, so their erasure cost is greater. If we let $\lambda_{\alpha|a}$ be the eigenvalues of $\rho_{A|a}$ and $\lambda_{\beta|ab}$ be the eigenvalues of $\rho_{B|ab}$, then after these measurements, the new erasure cost is

$$W^- = H(A,B) + \sum_a p_a H(\lambda_{\alpha|a}) + \sum_{a,b} p_{ab} H(\lambda_{\beta|ab}) \ . \quad \text{(A.54)}$$

Since the classical entropies of the eigenvalues are the same as the quantum entropies, the net classical work is the same as that given in Eq. (A.53). The reduction in classical work from not using rank-one POVMs has been transferred to an increased erasure cost when making measurements described by rank-one POVMs. We conclude that the demons cannot do worse by restricting themselves to rank-one POVMs, thus justifying our assumption of rank-one POVMs in Sec. 4.2.3.

Appendix B

Supplemental material to the interferometry chapters

B.1 Lossless Interferometry (Supplemental material to chapter 7)

B.1.1 Classical Fisher Information for an interferometric configuration

In this section we guide the reader through the calculation of classical Fisher information for a Mach-Zehnder interferometer with direct detection at the output and show that it is the same as the quantum Fisher information for a broad class of input states. The result of this calculation for the coherent-state–sqeezed-vacuum input was reported by Pezzé and Smerzi [PS08].

By adjusting phases at the second 50:50 beam splitter in the Mach-Zehnder, we can let it be described by the unitary operator $B^\dagger$. With this choice the Mach-Zehnder performs the overall transformation $B^\dagger U B = e^{iN_s\phi_s/2}e^{iK\phi_d/2} = e^{iN_s\phi_s/2}e^{iJ_y\phi_d}$, which includes the common-mode phase shift ϕ_s and a J_y rotation of the input state by angle $-\phi_d$ [YMK86]. The angle ϕ_d is the relative phase shift in the two arms, which

we are trying to estimate. The probability for detecting n_1 photons in the first output mode a_1 and n_2 photons in the second output mode a_2 is

$$P(n_1, n_2|\phi_d) = |\langle n_1, n_2|e^{iN_s\phi_s/2}e^{iJ_y\phi_d}|\psi_{\text{in}}\rangle|^2 = |\langle n_1, n_2|e^{iJ_y\phi_d}|\psi_{\text{in}}\rangle|^2 . \tag{B.1}$$

This is also the probability to detect $n_s = n_1 + n_2$ total photons at the output and a difference $n_d = n_1 - n_2$; written in terms of the eigenstates of sum and difference photon numbers, this probability becomes

$$P(n_s, n_d|\phi_d) = |\langle n_s, n_d|e^{iJ_y\phi_d}|\psi_{\text{in}}\rangle|^2 . \tag{B.2}$$

It is convenient to switch to the angular-momentum basis $|j, m\rangle = |n_s, n_d\rangle$ by identifying $j = n_s/2$ and $m = n_d/2$. The probability becomes

$$P(n_s, n_d|\phi_d) = |\langle j, m|e^{iJ_y\phi_d}|\psi_{\text{in}}\rangle|^2 = \left|\sum_{m'=-j}^{j} d^j_{m,m'}(-\phi_d)\langle j, m'|\psi_{\text{in}}\rangle\right|^2 , \tag{B.3}$$

where we use the fact that J_y conserves total photon number (total angular momentum) and we introduce the Wigner rotation matrices,

$$d^j_{m,m'}(-\phi_d) = \langle j, m|e^{i\phi_d J_y}|j, m'\rangle = \langle j, m'|e^{-i\phi_d J_y}|j, m\rangle . \tag{B.4}$$

Equation (B.3) is the form of the joint probability quoted in Eq. (5) of [PS08].

The case of interest is a product input, $|\psi_{\text{in}}\rangle = |\alpha\rangle \otimes S|0\rangle$, where $S|0\rangle = |0\rangle$ is the optimal squeezed vacuum input to the secondary input port. Without loss of generality, we can assume that α is real. Then the optimal squeezed vacuum state is squeezed along the quadrature axes of the input mode a_2, and the input state takes the form

$$|\psi_{\text{in}}\rangle = |\alpha\rangle \otimes S(r)|0\rangle , \qquad S(r) = e^{r(a^2 - a^{\dagger 2})/2} . \tag{B.5}$$

Under these assumptions, the amplitude $\langle j, m|\psi_{\text{in}}\rangle = \langle n_s, n_d|\psi_{\text{in}}\rangle$ is real. That these amplitudes are real is the only assumption about the input state that we use in the following; our calculation thus applies to all product and nonproduct inputs for which these amplitudes are real. Since the Wigner rotation matrix is real, as displayed in Eq. (B.4), the sum in Eq. (B.3) is also real.

The classical Fisher information for a Mach-Zehnder interferometer with direct detection is

$$F(\phi_d) = \sum_{n_s,n_d} \frac{1}{P(n_s,n_d|\phi_d)} \left(\frac{\partial P(n_s,n_d|\phi_d)}{\partial \phi_d} \right)^2 . \tag{B.6}$$

Omitting indices and decorations for readability when there is no risk of confusion, we can write

$$\frac{\partial P}{\partial \phi_d} = 2\sqrt{P} \sum_{m'=-j}^{j} \langle j,m'|\psi_{\rm in}\rangle \frac{\partial d^j_{m,m'}(-\phi_d)}{\partial \phi_d} , \tag{B.7}$$

where we use the reality of the number-state expansion coefficients of $|\psi_{\rm in}\rangle$, and thus

$$\begin{aligned} F(\phi_d) &= 4 \sum_j \sum_{m=-j}^{j} \left(\sum_{m'=-j}^{j} \langle j,m'|\psi_{\rm in}\rangle \frac{\partial d^j_{m,m'}(-\phi_d)}{\partial \phi_d} \right)^2 \\ &= 4 \sum_j \sum_{m=-j}^{j} \sum_{q,q'=-j}^{j} \langle \psi_{\rm in}|j,q\rangle\langle j,q|J_y e^{-iJ_y\phi_d}|j,m\rangle \\ &\qquad\qquad \times \langle j,m|e^{iJ_y\phi_d} J_y|j,q'\rangle\langle j,q'|\psi_{\rm in}\rangle \\ &= 4 \sum_j \sum_{q,q'=-j}^{j} \langle \psi_{\rm in}|j,q\rangle\langle j,q|J_y^2|j,q'\rangle\langle j,q'|\psi_{\rm in}\rangle . \end{aligned} \tag{B.8}$$

Since J_y does not change the total angular momentum, we can write this as

$$F(\phi_d) = 4 \sum_{j,j'} \sum_{q,q'=-j}^{j} \langle \psi_{\rm in}|j,q\rangle\langle j,q|J_y^2|j',q'\rangle\langle j',q'|\psi_{\rm in}\rangle = \langle \psi_{\rm in}|K^2|\psi_{\rm in}\rangle . \tag{B.9}$$

For an input state that has real expansion coefficients in the number basis, we also have

$$\begin{aligned} \langle \psi_{\rm in}|K|\psi_{\rm in}\rangle &= 2 \sum_{j,m} \langle \psi_{\rm in}|J_y|j,m\rangle\langle j,m|\psi_{\rm in}\rangle \\ &= -i \sum_{j,m} \langle \psi_{\rm in}|(J_+ - J_-)|j,m\rangle\langle j,m|\psi_{\rm in}\rangle \\ &= -i \sum_{j,m} [C_+(j,m)\langle \psi_{\rm in}|j,m+1\rangle - C_-(j,m)\langle \psi_{\rm in}|j,m-1\rangle]\langle j,m|\psi_{\rm in}\rangle \\ &= -i \sum_{j,m} [C_+(j,m) - C_-(j,m+1)]\langle \psi_{\rm in}|j,m+1\rangle\langle j,m|\psi_{\rm in}\rangle \\ &= 0 , \end{aligned} \tag{B.10}$$

where $C_{\pm}(j,m) = \sqrt{j(j+1) - m(m \pm 1)}$. The reality of the expansion coefficients is used to get to the last expression, which vanishes because $C_{+}(j,m) = C_{-}(j,m+1)$. The identity (B.10) allows us to write the classical Fisher information (B.9) as

$$F(\phi_d) = \langle\psi_{\rm in}|K^2|\psi_{\rm in}\rangle - \left(\langle\psi_{\rm in}|K|\psi_{\rm in}\rangle\right)^2 = \mathcal{F}_{dd}\,, \tag{B.11}$$

where $\mathcal{F}_{dd}$ is the quantum Fisher information for the differential phase shift (see Eq. (7.9) in Chap. 7). The equality of the classical and quantum Fisher informations applies to all input states that have real expansion coefficients in the number basis. It also applies to states obtained from such input states by rotating both input modes by the same angle θ, i.e., by applying $e^{iN_s\theta}$ to both modes. Applying this rotation to the state (B.5) removes the assumption that α is real.

While the direct calculation of the classical Fisher information was very instructive, there is a simpler way showing the equality of quantum and classical Fisher information for photon counting in both output ports of the interferometer. As in Eq. (6.36), all we need to show is that

$$\mathrm{Im}\{\langle\partial_\theta\psi(\theta)|\,j,m\rangle\,\langle j,m|\,\psi(\theta)\rangle\} = 0, \tag{B.12}$$

where $|\,\psi(\theta)\,\rangle = e^{iJ_y\phi_d}|\,\psi_{in}\,\rangle$. As discussed earlier,

$$\langle j,m|e^{iJ_y\phi_d}|\psi_{\rm in}\rangle = \sum_{m'=-j}^{j} d^j_{m,m'}(-\phi_d)\langle j,m'|\psi_{\rm in}\rangle \tag{B.13}$$

is a real quantity assuming the expansion coefficients of the input state in the number basis are real. Hence its derivative is real and Eq. (B.12) is satisfied.

B.1.2 Modal entanglement after the initial beam splitter

Here we present simple argument, due to Z. Jiang, to show that only coherent states in the secondary input port yield a product state after the initial beam splitter. A more general proof can be found in [JLC13].

We begin by noting that the product input state can be written as

$$|\psi_{\rm in}\rangle = |\alpha\rangle \otimes |\chi\rangle = D(a_1,\alpha)|0\rangle \otimes D(a_2,\beta)|\chi_0\rangle\,, \tag{B.14}$$

where $D(a,\alpha) = e^{\alpha a^\dagger - \alpha^* a}$ is the displacement operator, $\beta = \langle\chi_0|a_2|\chi_0\rangle$ is the mean amplitude of the state $|\chi\rangle$, and $|\chi_0\rangle = D(a_2,-\beta)|\chi\rangle$ has this mean amplitude removed. The state after the initial beamsplitter is

$$B|\psi_{\text{in}}\rangle = D\Big(a_1,(\alpha-i\beta)/\sqrt{2}\Big)\otimes D\Big(a_2,(\beta-i\alpha)/\sqrt{2}\Big)B|0,\chi_0\rangle\,. \tag{B.15}$$

The two displacement operators act locally in the two arms, so the modal entanglement of $B|\psi_{\text{in}}\rangle$ is the same as the modal entanglement of $B|0,\chi_0\rangle$. Specifically, the displacement of the primary mode before the beamsplitter does not contribute to the entanglement after the beamsplitter [ACR05]; likewise, the displacement of the secondary mode before the beamsplitter does not contribute to the post-beamsplitter entanglement. Showing that $B|\psi_{\text{in}}\rangle$ is a product state only if $|\chi\rangle$ is a coherent state is equivalent to showing that $B|0,\chi_0\rangle$ is a product state only if $|\chi_0\rangle$ is the vacuum state.

If $B|0,\chi_0\rangle$ is a product state, then after the beamsplitter,

$$0 = \langle a_1^\dagger\rangle\langle a_2\rangle = \langle a_1^\dagger a_2\rangle = \langle 0,\chi_0|B^\dagger a_1^\dagger a_2 B|0,\chi_0\rangle\,. \tag{B.16}$$

Using $B^\dagger a_1^\dagger a_2 B = \frac{1}{2}(-ia_1^\dagger a_1 + ia_2^\dagger a_2 + a_1^\dagger a_2 + a_2^\dagger a_1)$, we get

$$0 = \frac{i}{2}\langle\chi_0|a_2^\dagger a_2|\chi_0\rangle\,. \tag{B.17}$$

showing that $|\chi_0\rangle$ is the vacuum state, as promised.

We have shown that any state other than a coherent state into the secondary input port leads to modal entanglement after the beamsplitter.

B.2 Lossy interferometry (Supplemental material to Chap. 8)

In this part of the Appendix we guide the reader through most of the steps in the calculations presented in Chap. 8.

B.2.1 Escher et al.'s bound

First we will elaborate on how we obtained Expression 8.8. Plugging the explicit form for our Kraus operators (Eq. (8.5)) into Eqs. (8.2) and (8.3) we obtain

$$H_1(\phi_d,\eta;\gamma) = \frac{1}{4}\sum_{l,k}\frac{(1-\eta)^l}{l!}\frac{(1-\eta)^k}{k!}a^{\dagger l}b^{\dagger k}\left(N_d-(\gamma l-\gamma k)\right)^2\eta^{N_s}a^l b^k\,, \tag{B.18}$$

$$H_2(\phi_d,\eta;\gamma) = \frac{1}{2}\sum_{l,k}\frac{(1-\eta)^l}{l!}\frac{(1-\eta)^k}{k!}a^{\dagger l}b^{\dagger k}\left(N_d-(\gamma l-\gamma k)\right)\eta^{N_s}a^l b^k\,. \tag{B.19}$$

Examining H_1 and using

$$(N_d-(\gamma l-\gamma k))^2 = (n_a-\gamma l)^2+(n_b-\gamma k)^2-2(n_a-\gamma l)(n_b-\gamma k), \tag{B.20}$$

we find

$$H_1(\phi_d,\eta;\gamma) = \frac{1}{4}H^E_{1,a}+\frac{1}{4}H^E_{1,b}-\frac{1}{2}H^E_{2,a}\cdot H^E_{2,b}\,, \tag{B.21}$$

where $H^E_{n,d}$ refers to H_n in [EdMFD11b] applied to mode d. Given the results derived there we can streamline the presentation here to give

$$\begin{aligned}H_1(\phi_d,\eta;\gamma) &= \frac{1}{4}\left[(1-(1+\gamma)(1-\eta))^2n_a^2+(1+\gamma)^2\eta(1-\eta)n_a\right]\\ &\quad+\frac{1}{4}\left[(1-(1+\gamma)(1-\eta))^2n_b^2+(1+\gamma)^2\eta(1-\eta)n_b\right]\\ &\quad-\frac{1}{2}(1-(1+\gamma)(1-\eta))^2n_an_b\\ &= \frac{1}{4}\left(1-(1+\gamma)(1-\eta)\right)^2(n_a-n_b)^2\\ &\quad+\frac{1}{4}(1+\gamma)^2\eta(1-\eta)(n_a+n_b)\,.\end{aligned} \tag{B.22}$$

Again with the help of [EdMFD11b], we can write H_2 as

$$\begin{aligned}H_2(\phi_d,\eta;\gamma) &= \frac{1}{2}\left[\sum_l\frac{(1-n)^l}{l!}a^{\dagger l}(n_a-\gamma l)\eta^{n_a}a^l\right.\\ &\quad\left.-\sum_k\frac{(1-\eta)^k}{k!}b^{\dagger k}(n_b-\gamma k)\eta^{n_b}b^k\right]\\ &= \frac{1}{2}\left[(1-(1+\gamma)(1-\eta))\right](n_a-n_b).\end{aligned} \tag{B.23}$$

Putting all this together, we find

$$C_{Q,\gamma} = \left[1-(1+\gamma)(1-\eta)\right]^2 \langle \Delta N_d^2 \rangle + (1+\gamma)^2 \eta(1-\eta)\bar{N}_t \,. \tag{B.24}$$

Now we can optimize this bound over the parameter γ:

$$\frac{\partial}{\partial\gamma} C_Q \stackrel{!}{=} 0 \quad \rightarrow \quad \gamma_{opt} = \frac{\langle \Delta N_d^2 \rangle}{(1-\eta)\langle \Delta N_d^2 \rangle + \eta \bar{N}_t} - 1 \,. \tag{B.25}$$

Here we see why an interpretation of γ as the 'position' of the auxiliary beam splitters, as done in [EdMFD11b], is problematic. To achieve Heisenberg scaling, one needs $\langle \Delta N_d^2 \rangle \sim \bar{N}_t^2$, in which case γ_{opt} asymptotes to $1/(1-\eta)$ when $\bar{N}_t$ is large and is, therefore, not confined to the range $[-1, 0]$ as suggested by the interpretation used in [EdMFD11b].

Putting the optimal value for γ into Eq. (B.24) gives

$$C_{Q,\gamma_{opt}} = \frac{\eta \bar{N}_t \langle \Delta N_d^2 \rangle}{(1-\eta)\langle \Delta N_d^2 \rangle + \eta \bar{N}_t} \,. \tag{B.26}$$

To show the monotonicity of this expression as a function of $\langle \Delta N_d^2 \rangle$, we simply calculate its first derivative

$$\frac{\partial}{\partial \langle \Delta N_d^2 \rangle} C_{Q\gamma_{opt}} = \frac{\bar{N}_t^2 \eta^2}{\left(\langle N_d^2 \rangle (1-\eta) + \bar{N}_t \eta\right)^2} \,, \tag{B.27}$$

which is obviously positive. Thus to maximize the bound, one finds the state that maximizes $\langle \Delta N_d^2 \rangle$ subject to whatever constrains one imposes, as in the lossless case. When this maximal value of $\langle \Delta N_d^2 \rangle$ is plugged into Eq. (B.26), it provides a C_Q that is an upper bound on the quantum Fisher information for all possible states satisfying the constraints.

B.2.2 Exact Fisher Information for laser interferometry with losses

Jiang provided a way to calculate explicitly the quantum Fisher information for arbitrary Gaussian states [Jia14]:

$$F_Q = \frac{1}{2}\mathrm{tr}(\dot{\Gamma}\Phi) + 2\dot{d}^T \Gamma^{-1} \dot{d} \,. \tag{B.28}$$

The various quantities in this expression are defined in Chap. 8.

For the state $|\,\alpha\,\rangle \otimes S_b(r)|0\rangle$ as the input to the interferometer, the state after the phase shifts and losses has the following covariance matrix:

$$\Gamma_{11} = \eta\,(\bar{N}_b + \sqrt{\bar{N}_b(\bar{N}_b+1)}\cos(\phi_s+\phi_d) + 1\,, \tag{B.29}$$

$$\Gamma_{22} = \eta\,(\bar{N}_b - \sqrt{\bar{N}_b(\bar{N}_b+1)}\cos(\phi_s-\phi_d) + 1\,, \tag{B.30}$$

$$\Gamma_{33} = \eta\,(\bar{N}_b - \sqrt{\bar{N}_b(\bar{N}_b+1)}\cos(\phi_s+\phi_d) + 1\,, \tag{B.31}$$

$$\Gamma_{44} = \eta\,(\bar{N}_b + \sqrt{\bar{N}_b(\bar{N}_b+1)}\cos(\phi_s-\phi_d) + 1\,, \tag{B.32}$$

$$\Gamma_{12} = -\eta\,\bar{N}_b \sin\phi_d + \eta\sqrt{\bar{N}_b(\bar{N}_b+1)}\sin\phi_s\,, \tag{B.33}$$

$$\Gamma_{13} = -\eta\sqrt{\bar{N}_b(\bar{N}_b+1)}\sin(\phi_s+\phi_d)\,, \tag{B.34}$$

$$\Gamma_{14} = \eta\,\bar{N}_b \cos\phi_d + \eta\sqrt{\bar{N}_b(\bar{N}_b+1)}\cos\phi_s\,, \tag{B.35}$$

$$\Gamma_{23} = -\eta\,\bar{N}_b \cos\phi_d + \eta\sqrt{\bar{N}_b(\bar{N}_b+1)}\cos\phi_s\,, \tag{B.36}$$

$$\Gamma_{24} = \eta\sqrt{\bar{N}_b(\bar{N}_b+1)}\sin(\phi_s-\phi_d)\,, \tag{B.37}$$

$$\Gamma_{34} = -\eta\,\bar{N}_b \sin\phi_d - \eta\sqrt{\bar{N}_b(\bar{N}_b+1)}\sin\phi_s\,. \tag{B.38}$$

If we now do a symplectic basis change to the basis where Γ is diagonal (so-called standard form), i.e., $S\Gamma S^T = \Gamma_s$, we can calculate the symplecticly invariant quantity $\frac{1}{2}\mathrm{tr}(\dot{\Gamma}\Phi) = \frac{1}{2}\mathrm{tr}(\dot{\Gamma}_s\Phi_s)$, with $\dot{\Gamma}_s = S\dot{\Gamma}S^T$. The advantage of doing this change of basis is that the equation for Φ_s can be solved explicitly in terms of Γ_s and $\dot{\Gamma}_s$. To perform this basis change we need to find the matrix S, which is done though a symplectic diagonalization procedure. Details on how to do this can be found in [Jia15]. To make it easier for the reader to verify our calculation, we provide the matrix S for

the state whose covariance matrix is shown above. The columns of S are given by

$$S_1 = \begin{pmatrix} \frac{\sin[\phi_d]}{\sqrt{2}} \\ -\frac{2\sqrt{\bar{N}_b(1+\bar{N}_b)}\eta\sin[\phi_d]+\left(1+2\bar{N}_b\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)\sin[\phi_s]}{2\sqrt{1+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}+2\bar{N}_b\eta\left(2-2\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)}} \\ -\frac{\cos[\phi_d]}{\sqrt{2}} \\ \frac{-2\sqrt{\bar{N}_b(1+\bar{N}_b)}\eta\cos[\phi_d]+\left(1+2\bar{N}_b\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)\cos[\phi_s]}{2\sqrt{1+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}+2\bar{N}_b\eta\left(2-2\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)}} \end{pmatrix}, \tag{B.39}$$

$$S_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{2\sqrt{\bar{N}_b(1+\bar{N}_b)}\eta+\left(1+2\bar{N}_b\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)\cos[\phi_d-\phi_s]}{2\sqrt{1+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}+2\bar{N}_b\eta\left(2-2\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)}} \\ 0 \\ -\frac{\left(1+2\bar{N}_b\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)\sin[\phi_d-\phi_s]}{2\sqrt{1+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}+2\bar{N}_b\eta\left(2-2\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)}} \end{pmatrix}, \tag{B.40}$$

$$S_3 = \begin{pmatrix} \frac{\cos[\phi_d]}{\sqrt{2}} \\ -\frac{2\sqrt{\bar{N}_b(1+\bar{N}_b)}\eta\cos[\phi_d]+\left(1+2\bar{N}_b\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)\cos[\phi_s]}{2\sqrt{1+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}+2\bar{N}_b\eta\left(2-2\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)}} \\ \frac{\sin[\phi_d]}{\sqrt{2}} \\ \frac{2\sqrt{\bar{N}_b(1+\bar{N}_b)}\eta\sin[\phi_d]-\left(1+2\bar{N}_b\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)\sin[\phi_s]}{2\sqrt{1+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}+2\bar{N}_b\eta\left(2-2\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)}} \end{pmatrix}, \tag{B.41}$$

$$S_4 = \begin{pmatrix} 0 \\ \frac{\left(1+2\bar{N}_b\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)\sin[\phi_d-\phi_s]}{2\sqrt{1+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}+2\bar{N}_b\eta\left(2-2\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)}} \\ \frac{1}{\sqrt{2}} \\ \frac{-2\sqrt{\bar{N}_b(1+\bar{N}_b)}\eta+\left(1+2\bar{N}_b\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)\cos[\phi_d-\phi_s]}{2\sqrt{1+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}+2\bar{N}_b\eta\left(2-2\eta+\sqrt{1-4\bar{N}_b(-1+\eta)\eta}\right)}} \end{pmatrix}. \tag{B.42}$$

Adhering to the procedure in [Jia14], we are able to obtain Φ_s and calculate the first term of F_Q in Eq. (B.28):

$$\frac{1}{2}\mathrm{tr}(\dot{\Gamma}\Phi) = \bar{N}_b\eta. \tag{B.43}$$

The second term in Eq. (B.28) is straightforward. The vector of first moments for

our state has the form

$$d = \sqrt{\eta}\,\alpha \begin{pmatrix} \cos\left(\frac{\phi_s+\phi_d}{2}\right) \\ -\sin\left(\frac{\phi_s-\phi_d}{2}\right) \\ -\sin\left(\frac{\phi_s+\phi_d}{2}\right) \\ -\cos\left(\frac{\phi_s-\phi_d}{2}\right) \end{pmatrix}. \tag{B.44}$$

With Γ^{-1} being the inverse of the covariance matrix Γ, the second term in Eq. (B.28) evaluates to

$$2\dot{d}^T\Gamma^{-1}\dot{d} = \frac{N_a\,\eta\left[1+2\,\bar{N}_b\,\eta+2\sqrt{\bar{N}_b(1+\bar{N}_b)}\,\eta\right]}{1+4\,\bar{N}_b\,(1-\eta)\,\eta}. \tag{B.45}$$

B.2.3 Position of the auxiliary beam splitters

As was shown in [EdMFD11b], the case where $\gamma = 0$ is mathematically equivalent to the physical situation where no phase shift is performed in the ancillary modes, but the ancillary beam splitter is moved before the main phase shift in the interferometer. More generally, it is straightforward to show that the situation where the ancillary beam splitters occur before the phase shift of the interferometer is mathematically identical to the situation where they occur after the beam splitters, but with an additional phase shift of $\varphi_{1(2)}$ in ancillary mode 1(2). As we can choose arbitrary unitary transformations in the ancillary modes and get valid Kraus operators for the bound we calculate, it is clearly irrelevant if the ancillary beam splitters are positioned before or after the phase shift of the interferometer. In the following I will show that a similar consideration also lets us move the ancillary beam splitters before the first beam splitter of the interferometer, without changing any of the results.

To start, we will assume our ancillary beam splitters, B_1 and B_2, sit between the interferometer's phase shift and the initial beam splitter B of the interferometer. We can write

$$B_1 = e^{-\mu(a^\dagger c - c^\dagger a)/2}, \quad B_2 = e^{-\mu(b^\dagger d - d^\dagger b)/2}, \tag{B.46}$$

where c and d are the annihilation operators of the two ancillary modes. Notice that here we adopted a different beam splitter convention: To comply with the Krauss

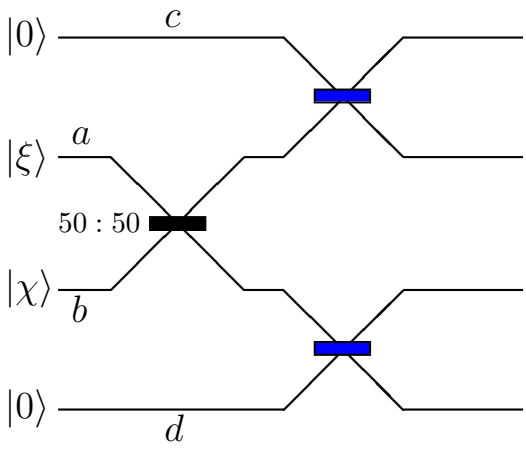

Figure B.1: First half of our interferometric setup. The ancillary modes start out in the vacuum state. The 50:50 beam splitter of the interferometer B is drawn in black, while the ancillary beam splitters, B_1 and B_2, with transmissivity $\cos^2(\mu/2)$, are depicted in blue.

operators derived in [EdMFD11b], B_1 and B_2 are J_y-rotations while all the other beam splitters in this document are J_x-rotations. With this, the first half of our interferometric setup, which involves the unitary sequence B_2B_1B, is depicted in Fig. B.1.

Using $B^\dagger aB = (a - ib)/\sqrt{2}$ and $B^\dagger bB = (b - ia)/\sqrt{2}$, we can write

$$\begin{aligned} B_2B_1B &= BB^\dagger B_2B_1B \\ &= B\exp\Big[-\frac{\mu}{2}\frac{1}{\sqrt{2}}\Big((b^\dagger + ia^\dagger)d - d^\dagger(b - ia) \\ &\qquad\qquad + (a^\dagger + ib^\dagger)c - c^\dagger(a - ib)\Big)\Big] \\ &= B\exp\Big[-\frac{\mu}{2}\frac{1}{\sqrt{2}}\Big(b^\dagger(d + ic) - b(d^\dagger - ic^\dagger) \\ &\qquad\qquad + a^\dagger(c + id) - a(c^\dagger - id^\dagger)\Big)\Big] \\ &= BB_{c,d}B_2B_1B_{c,d}^\dagger\,. \end{aligned} \tag{B.47}$$

Here $B_{c,d}$ denotes a $50:50$ beam splitter between modes c and d,

$$B_{c,d} = e^{-i(c^\dagger d + d^\dagger c)\pi/4}. \tag{B.48}$$

The last line in Eq. (B.47) corresponds to the physical situation depicted in Fig. B.2.

From Fig. (B.2)(b) we can see now that the only difference between having the losses before or after the first beam splitter of the interferometer is a 50:50 beam splitter mixing the ancillary modes after they have interacted with the system. Hence, if we were to model the losses as occurring on the modes before the first beam splitter of the interferometer, we could simply add the the last beam splitter in Fig. (B.2) in order to end up with Kraus operators identical to Eq. (8.5). Then the rest of our analysis of the upper bound in Chap. 8 would be applicable to this case without further changes. Of course, the last beam splitter, acting as it does only on the ancillary modes after they have interacted with primary modes, doesn't make any difference to the actual sensitivity of the interferometer.

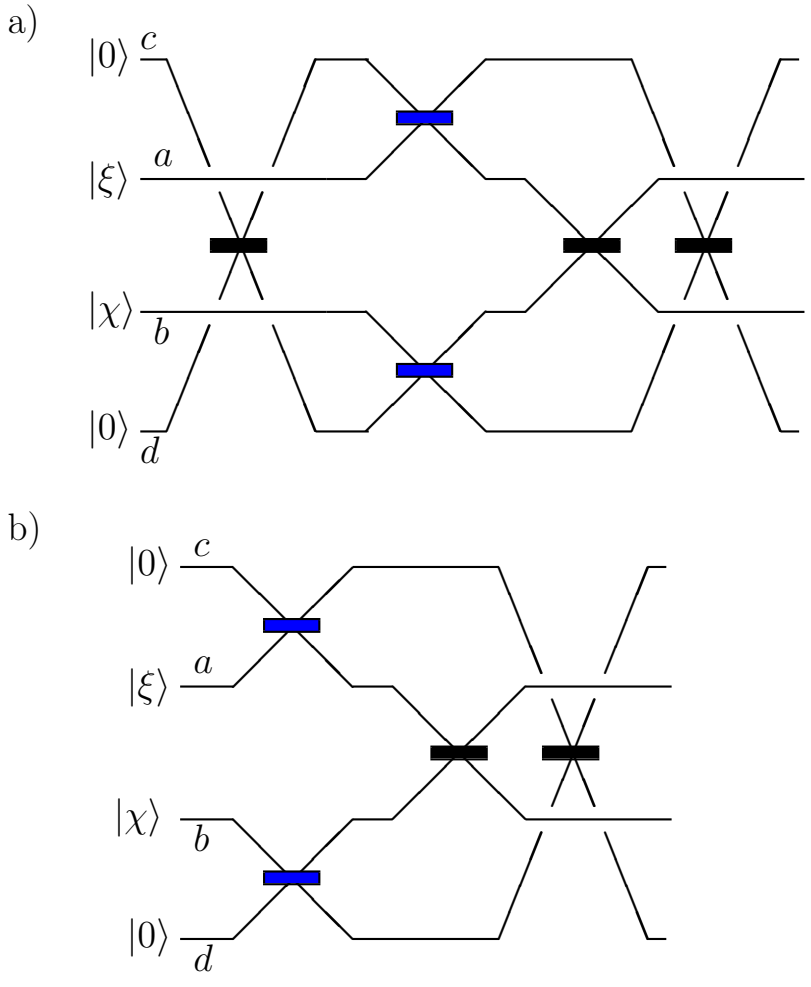

Figure B.2: (a) This physical situation is equivalent to the one pictured in Fig. (B.1). (b) The first beam splitter between modes c and d has been removed as it does not change anything, given that both ancillary modes start out in vacuum.

References

A

[AAA+13] J. Aasi, J. Abadie, B. Abbott, R. Abbott, T. Abbott, M. Abernathy, C. Adams, T. Adams, P. Addesso, R. Adhikari, et al. Enhanced sensitivity of the ligo gravitational wave detector by using squeezed states of light. *Nature Photonics*, 7(8):613–619, 2013. 6, 75

[AAS10] I. Afek, O. Ambar, and Y. Silberberg. High-noon states by mixing quantum and classical light. *Science*, 328(5980):879–881, 2010. 90

[ACR05] J. K. Asbóth, J. Calsamiglia, and H. Ritsch. Computable measure of nonclassicality for light. *Phys. Rev. Lett.*, 94:173602, May 2005. 103, 142

[AD10] G. Adesso and A. Datta. Quantum versus classical correlations in Gaussian states. *Physical Review Letters*, 105(3):030501, 2010. 58

[AJ11] A. Al-Qasimi and D. F. V. James. A comparison of the attempts of quantum discord and quantum entanglement to capture quantum correlations. *Physical Review A*, 83:032101, 2011. 60, 61

B

[BBC+93] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, Mar 1993. 30

[BC94] S. L. Braunstein and C. M. Caves. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.*, 72:3439–3443, May 1994. 8, 80, 85, 100

[BDF+99] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters. Quantum nonlocality without entanglement. *Physical Review A*, 59:1070–1091, 1999. 27, 55, 124

[Ben82] C. Bennett. The thermodynamics of computation - a review. *International Journal of Theoretical Physics*, 21(12):905–940, 1982. 24

[BFM10] F. Benatti, R. Floreanini, and U. Marzolino. Sub-shot-noise quantum metrology with entangled identical particles. *Annals of Physics*, 325(4):924 – 935, 2010. 106

[BHB+09] D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman. How to perform the most accurate possible phase measurements. *Phys. Rev. A*, 80:052114, Nov 2009. 90

[BIWH96] J. J. . Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen. Optimal frequency measurements with maximally correlated states. *Phys. Rev. A*, 54:R4649–R4652, Dec 1996. 6, 90

[BKA+00] A. N. Boto, P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, and J. P. Dowling. Quantum interferometric optical lithography: Exploiting entanglement to beat the diffraction limit. *Phys. Rev. Lett.*, 85:2733–2736, Sep 2000. 6, 106

[BKCD02] R. Blume-Kohout, C. Caves, and I. Deutsch. Climbing mount scalable: Physical resource requirements for a scalable quantum computer. *Foundations of Physics*, 32(11):1641–1670, 2002. 3, 64

[BLC92] S. L. Braunstein, A. S. Lane, and C. M. Caves. Maximum-likelihood analysis of multiple quantum phase measurements. *Physical review letters*, 69(15):2153, 1992. 92, 110

[BR04] H.-A. Bachor and T. C. Ralph. *A Guide to Experiments in Quantum Optics, 2nd.* Wiley, 2004. 87

[Bra92a] S. L. Braunstein. How large a sample is needed for the maximum likelihood estimator to be approximately gaussian? *Journal of Physics A: Mathematical and General*, 25(13):3813, 1992. 110

[Bra92b] S. L. Braunstein. Quantum limits on precision measurements of phase. *Physical review letters*, 69(25):3598, 1992. 92

[BS84] R. S. Bondurant and J. H. Shapiro. Squeezed states in phase-sensing interferometers. *Phys. Rev. D*, 30:2548–2556, Dec 1984. 88, 103, 110

[BT10] A. Brodutch and D. R. Terno. Quantum discord, local operations, and Maxwell's demons. *Physical Review A*, 81(6):062103, 2010. 5, 28, 34, 49, 50, 51, 53

C

[C+11] L. S. Collaboration et al. A gravitational wave observatory operating beyond the quantum shot-noise limit. *Nature Physics*, 7(12):962–965, 2011. 6, 75

[CAB+11] D. Cavalcanti, L. Aolita, S. Boixo, K. Modi, M. Piani, and A. Winter. Operational interpretations of quantum discord. *Physical Review A*, 83:032324, 2011. 124

[Cav] C. M. Caves. Operator formalism and quasidistributions for creation an annihilation operators. *Available under: http://info.phys.unm.edu/∼caves/reports/a.pdf*. 84, 85

[Cav81] C. M. Caves. Quantum-mechanical noise in an interferometer. *Phys. Rev. D*, 23:1693–1708, Apr 1981. 6, 75, 87, 110, 125

[CdOF11] M. F. Cornelio, M. C. de Oliveira, and F. F. Fanchini. Entanglement irreversibility from quantum discord and quantum deficit. *Phys. Rev. Lett.*, 107:020502, Jul 2011. 124

[CGB03] R. A. Campos, C. C. Gerry, and A. Benmoussa. Optical interferometry at the heisenberg limit with twin fock states and parity measurements. *Phys. Rev. A*, 68:023810, Aug 2003. 106

[Coh68] M. Cohen. The fisher information and convexity (corresp.). *Information Theory, IEEE Transactions on*, 14(4):591–592, 1968. 103

[Col10] J. H. Cole. Understanding entanglement sudden death through multipartite entanglement and quantum correlations. *Journal of Physics A: Mathematical and Theoretical*, 43(13):135301, 2010. 24, 64, 66

[Cra99] H. Cramér. *Mathematical methods of statistics*, volume 9. Princeton university press, 1999. 78

[CT12] T. M. Cover and J. A. Thomas. *Elements of information theory*. John Wiley & Sons, 2012. 14, 75

[CTDL91] C. Cohen-Tannoudji, B. Diu, and F. Laloe. *Quantum Mechanics*. Number v. 1 in Quantum Mechanics. Wiley, 1991. 82

[CZY+11] Q. Chen, C. Zhang, S. Yu, X. X. Yi, and C. H. Oh. Quantum discord of two-qubit x states. *Phys. Rev. A*, 84:042313, Oct 2011. 57

D

[DacVB10] B. Dakić, V. Vedral, and Č. Brukner. Necessary and sufficient condition for nonzero quantum discord. *Phys. Rev. Lett.*, 105:190502, Nov 2010. xvi, 5, 30, 68, 72

[Dat08] A. Datta. *Studies on the Role of Entanglement in Mixed-State Quantum Computation*. PhD thesis, University of New Mexico, Albuquerque, New Mexico, 2008. 34, 53, 56

[Dat10] A. Datta. A condition for the nullity of quantum discord. *arXiv:1003.5256v2 [quant-ph]*, 2010. 32, 131

[DDanBS13] R. Demkowicz-Dobrzański, K. Banaszek, and R. Schnabel. Fundamental quantum interferometry bound for the squeezed-light-enhanced gravitational wave detector geo 600. *Phys. Rev. A*, 88:041802, Oct 2013. 99, 120, 125

[DDDS+09] 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.*, 102:040403, Jan 2009. 6, 8

[DDJK14] R. Demkowicz-Dobrzanski, M. Jarzyna, and J. Kolodynski. Quantum limits in optical interferometry. *arXiv preprint arXiv:1405.7703*, 2014. 75, 94, 110

[DFC05] A. Datta, S. T. Flammia, and C. M. Caves. Entanglement and the power of one qubit. *Phys. Rev. A*, 72:042316, Oct 2005. 5

[DG09] A. Datta and S. Gharibian. Signatures of nonclassicality in mixed-state quantum computation. *Physical Review A*, 79(4):042325–8, 2009. 34

[Dow08] J. P. Dowling. Quantum optical metrology - the lowdown on high-N00N states. *Contemporary Physics*, 49(2):125–143, 2008. 106

[DSC08] A. Datta, A. Shaji, and C. M. Caves. Quantum discord and the power of one qubit. *Phys. Rev. Lett.*, 100:050502, Feb 2008. 5, 34

[DV07] A. Datta and G. Vidal. Role of entanglement and correlations in mixed-state quantum computation. *Physical Review A*, 75(4):042310, 2007. 34

E

[Eas10] B. Eastin. Simulating Concordant Computations. *arXiv:1006.4402 [quant-ph]*, June 2010. 4, 34, 123

[EdMFD11a] B. Escher, R. de Matos Filho, and L. Davidovich. General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology. *Nature Physics*, 7(5):406–411, 2011. 6, 8, 114, 115, 120

[EdMFD11b] B. Escher, R. de Matos Filho, and L. Davidovich. Quantum metrology for noisy systems. *Brazilian Journal of Physics*, 41(4-6):229–247, 2011. 6, 114, 143, 144, 147, 148

F

[FAC+10] A. Ferraro, L. Aolita, D. Cavalcanti, F. M. Cucchietti, and A. Acín. Almost all quantum states have nonclassical correlations. *Phys. Rev. A*, 81:052318, May 2010. 4, 32, 64

[FI08] A. Fujiwara and H. Imai. A fibre bundle over manifolds of quantum channels and its application to quantum statistics. *Journal of Physics A: Mathematical and Theoretical*, 41(25):255304, 2008. 8, 112

[Fis25] R. A. Fisher. Theory of statistical estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 22:700–725, 7 1925. 78

[Fuj06] A. Fujiwara. Strong consistency and asymptotic efficiency for adaptive quantum estimation problems. *Journal of Physics A: Mathematical and General*, 39(40):12489, 2006. 97, 100

G

[GBC02] C. C. Gerry, A. Benmoussa, and R. A. Campos. Nonlinear interferometer as a resource for maximally entangled photonic states: Application to interferometry. *Phys. Rev. A*, 66:013804, Jul 2002. 6, 90

[Ger00] C. C. Gerry. Heisenberg-limit interferometry with four-wave mixers operating in a nonlinear regime. *Phys. Rev. A*, 61:043811, Mar 2000. 6, 106

[GGZ11] F. Galve, G. L. Giorgi, and R. Zambrini. Orthogonal measurements are almost sufficient for quantum discord of two qubits. *EPL (Europhysics Letters)*, 96(4):40005, 2011. 58

[GLM06] V. Giovannetti, S. Lloyd, and L. Maccone. Quantum metrology. *Phys. Rev. Lett.*, 96:010401, Jan 2006. 88

[GM10] C. C. Gerry and J. Mimih. The parity operator in quantum optical metrology. *Contemporary Physics*, 51(6):497–511, 2010. 106

[GM12] V. Giovannetti and L. Maccone. Sub-heisenberg estimation strategies are ineffective. *Phys. Rev. Lett.*, 108:210404, May 2012. 110

[GP10] P. Giorda and M. G. A. Paris. Gaussian quantum discord. *Physical Review Letters*, 105(2):020503, 2010. 58

[GPA10] D. Girolami, M. Paternostro, and G. Adesso. Non-classicality indicators and extremal quantum correlations in two-qubit states. *arXiv:1008.4136v1 [quant-ph]*, 2010. 61

[GPA11] D. Girolami, M. Paternostro, and G. Adesso. Faithful nonclassicality indicators and extremal quantum correlations in two-qubit states. *Journal of Physics A: Mathematical and Theoretical*, 44(35):352002, 2011. 61

[GSYL87] P. Grangier, R. E. Slusher, B. Yurke, and A. LaPorta. Squeezed-light – enhanced polarization interferometer. *Phys. Rev. Lett.*, 59:2153–2156, Nov 1987. 6

H

[HB93] M. J. Holland and K. Burnett. Interferometric detection of optical phase shifts at the heisenberg limit. *Phys. Rev. Lett.*, 71:1355–1358, Aug 1993. 105, 106

[Hel76] C. W. Helstrom. *Quantum detection and estimation theory*. Academic press, 1976. 8, 75, 80, 97

[HH96] R. Horodecki and M. Horodecki. Information-theoretic aspects of inseparability of mixed states. *Phys. Rev. A*, 54:1838–1843, Sep 1996. 65, 66

[HHHH09] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009. 3, 4, 22, 65, 66

[HJPW04] P. Hayden, R. Jozsa, D. Petz, and A. Winter. Structure of states which satisfy strong subadditivity of quantum entropy with equality. *Communications in Mathematical Physics*, 246:359–374, 2004. 53, 131

[Hol11] A. S. Holevo. *Probabilistic and statistical aspects of quantum theory*, volume 1. Springer Science & Business Media, 2011. 8, 75, 80, 97

[HOW05] M. Horodecki, J. Oppenheim, and A. Winter. Partial quantum information. *Nature*, 436(7051):673–676, 2005. 30

[HV01] L. Henderson and V. Vedral. Classical, quantum and total correlations. *Journal of Physics A: Mathematical and General*, 34:6899–6905, 2001. 28, 29, 34, 67

J

[JDDan12] M. Jarzyna and R. Demkowicz-Dobrzański. Quantum interferometry with and without an external phase reference. *Phys. Rev. A*, 85:011801, Jan 2012. 98

[Jia14] Z. Jiang. Quantum fisher information for states in exponential form. *Phys. Rev. A*, 89:032128, Mar 2014. 116, 117, 121, 144, 146

[Jia15] Z. Jiang. Particle Correlations in Bose-Einstein Condensates. *ArXiv e-prints*, February 2015. 145

[JL03] R. Jozsa and N. Linden. On the role of entanglement in quantum-computational speed-up. *Proceedings of the Royal Society A*, 459(2036):2011–2032, 2003. 3, 4, 34

[JLC13] Z. Jiang, M. D. Lang, and C. M. Caves. Mixing nonclassical pure states in a linear-optical network almost always generates modal entanglement. *Phys. Rev. A*, 88:044301, Oct 2013. 98, 106, 110, 141

K

[KDD10] J. Kołodyński and R. Demkowicz-Dobrzański. Phase estimation without *a priori* phase knowledge in the presence of loss. *Phys. Rev. A*, 82:053804, Nov 2010. 112, 120

[KL98] E. Knill and R. Laflamme. Power of one bit of quantum information. *Phys. Rev. Lett.*, 81:5672–5675, Dec 1998. 5

[KPH+98] T. Kim, O. Pfister, M. J. Holland, J. Noh, and J. L. Hall. Influence of decorrelation on heisenberg-limited interferometry with quantum correlated photons. *Phys. Rev. A*, 57:4004–4013, May 1998. 106, 109

[KS96] M. S. Kim and B. C. Sanders. Squeezing and antisqueezing in homodyne measurements. *Phys. Rev. A*, 53:3694–3697, May 1996. 88, 110

[KSD11] S. Knysh, V. N. Smelyanskiy, and G. A. Durkin. Scaling laws for precision in quantum interferometry and the bifurcation landscape of the optimal state. *Phys. Rev. A*, 83:021804, Feb 2011. 6, 8, 112, 120

[KVL+12] A. Khalaidovski, H. Vahlbruch, N. Lastzka, C. Gräf, H. Lück, K. Danzmann, H. Grote, and R. Schnabel. Status of the geo 600 squeezed-light laser. *Journal of Physics: Conference Series*, 363(1):012013, 2012. 75

L

[Lan61] R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3):183–191, July 1961. 24, 35

[LBC93] A. S. Lane, S. L. Braunstein, and C. M. Caves. Maximum-likelihood statistics of multiple quantum phase measurements. *Physical Review A*, 47(3):1667, 1993. 92, 93, 110

[LCS11] M. D. Lang, C. M. Caves, and A. Shaji. Entropic measures of non-classical correlations. *International Journal of Quantum Information*, 09(07n08):1553–1586, 2011. 5, 28

[LJW13] J. Liu, X. Jing, and X. Wang. Phase-matching condition for enhancement of phase sensitivity in quantum metrology. *Phys. Rev. A*, 88:042316, Oct 2013. 104

[LKD02] H. Lee, P. Kok, and J. P. Dowling. A quantum rosetta stone for interferometry. *Journal of Modern Optics*, 49(14-15):2325–2338, 2002. 6, 90

[LR02] H. Leff and A. F. Rex. *Maxwell's Demon 2 Entropy, Classical and Quantum Information, Computing*. CRC Press, 2002. 24

[Luo08a] S. Luo. Using measurement-induced disturbance to characterize correlations as classical or quantum. *Physical Review A*, 77(2):022301, February 2008. 5, 28, 31, 34, 45

[Luo08b] S. Luo. Quantum discord for two-qubit systems. *Phys. Rev. A*, 77:042303, Apr 2008. 66, 67

M

[Max91] J. C. Maxwell. *Theory of Heat*. Longmans, Green & Co., London, 1891. 33

[MBC$^+$12] K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral. The classical-quantum boundary for correlations: Discord and related measures. *Rev. Mod. Phys.*, 84:1655–1707, Nov 2012. 28, 30

[MCSV09] J. Maziero, L. C. Céleri, R. M. Serra, and V. Vedral. Classical and quantum correlations under decoherence. *Phys. Rev. A*, 80:044102, Oct 2009. 64, 66

[MD11] V. Madhok and A. Datta. Interpreting quantum discord through quantum state merging. *Physical Review A*, 83:032323, 2011. 30, 124

[MD13] V. Madhok and A. Datta. Quantum discord as a resource in quantum communication. *International Journal of Modern Physics B*, 27(01n03):1345041, 2013. 30

[MH09] D. Meiser and M. J. Holland. Robustness of heisenberg-limited interferometry with balanced fock states. *New Journal of Physics*, 11(3):033002, 2009. 106

[Mon13] A. Monras. Phase space formalism for quantum estimation of gaussian states. *arXiv preprint arXiv:1303.3682*, 2013. 116, 117

[MPM10] L. Mazzola, J. Piilo, and S. Maniscalco. Sudden transition between classical and quantum decoherence. *Phys. Rev. Lett.*, 104:200401, May 2010. 64, 66, 69, 70

[MPS+10] K. Modi, T. Paterek, W. Son, V. Vedral, and M. Williamson. Unified view of quantum and classical correlations. *Phys. Rev. Lett.*, 104:080501, Feb 2010. 5, 28, 31, 34, 52, 68

[MTG+11] L. Mišta, R. Tatham, D. Girolami, N. Korolkova, and G. Adesso. Measurement-induced disturbances and nonclassical correlations of Gaussian states. *Physical Review A*, 83(4):042325, 2011. 58

[MWF+10] J. Maziero, T. Werlang, F. F. Fanchini, L. C. Céleri, and R. M. Serra. System-reservoir dynamics of quantum and classical correlations. *Phys. Rev. A*, 81:022116, Feb 2010. 64, 66

N

[Nai40] M. A. Naimark (aka Neumark). Spectral functions of a symmetric operator. *Izvestya Akademii Nauk SSSR: Seria Matematicheskaya*, 4(3):277–318, 1940. 128

[NC00] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000. 12, 13, 19, 21, 26, 40, 44, 53, 129

[NOO+07] T. Nagata, R. Okamoto, J. L. O'Brien, K. Sasaki, and S. Takeuchi. Beating the standard quantum limit with four-entangled photons. *Science*, 316(5825):726–729, 2007. 90

O

[OHHH02] J. Oppenheim, M. Horodecki, P. Horodecki, and R. Horodecki. Thermodynamical approach to quantifying quantum correlations. *Phys. Rev. Lett.*, 89:180402, Oct 2002. 5, 28, 34, 49

[OZ01] H. Ollivier and W. H. Zurek. Quantum discord: A measure of the quantumness of correlations. *Physical Review Letters*, 88(1):017901, Dec 2001. 5, 28, 29, 34, 48, 51, 56

P

[Par99] M. G. A. Paris. Entanglement and visibility at the output of a mach-zehnder interferometer. *Phys. Rev. A*, 59:1615–1621, Feb 1999. 110

[PCMH09] M. Piani, M. Christandl, C. E. Mora, and P. Horodecki. Broadcast copies reveal the quantumness of correlations. *Phys. Rev. Lett.*, 102:250503, Jun 2009. 5, 28

[PFB+12] O. Pinel, J. Fade, D. Braun, P. Jian, N. Treps, and C. Fabre. Ultimate sensitivity of precision measurements with intense gaussian quantum light: A multimodal approach. *Phys. Rev. A*, 85:010101, Jan 2012. 110

[PHH08] M. Piani, P. Horodecki, and R. Horodecki. No-local-broadcasting theorem for multipartite quantum correlations. *Physical Review Letters*, 100:090502, 2008. 34, 46, 54

[PS08] L. Pezzé and A. Smerzi. Mach-zehnder interferometry at the heisenberg limit with coherent and squeezed-vacuum light. *Phys. Rev. Lett.*, 100:073601, Feb 2008. 88, 100, 102, 103, 109, 138, 139

R

[RL12] Á. Rivas and A. Luis. Sub-heisenberg estimation of non-random phase shifts. *New Journal of Physics*, 14(9):093052, 2012. 93, 110, 111

[RR02] A. K. Rajagopal and R. W. Randall. Separability and correlations in composite states based on entropy methods. *Physical Review A*, 66:022104, 2002. 34, 45, 51

S

[SG64] L. Susskind and J. Glogower. Quantum mechanical phase and time operator. Technical report, Cornell Univ., Ithaca, NY, 1964. 92

[SLG+08] F. W. Sun, B. H. Liu, Y. X. Gong, Y. F. Huang, Z. Y. Ou, and G. C. Guo. Experimental demonstration of phase measurement precision beating standard quantum limit by projection measurement. *EPL (Europhysics Letters)*, 82(2):24001, 2008. 106

[SQ15] J. Sahota and N. Quesada. Quantum correlations in optical metrology: Heisenberg-limited phase estimation without mode entanglement. *Phys. Rev. A*, 91:013808, Jan 2015. 110

[SSW89] J. H. Shapiro, S. R. Shepard, and N. C. Wong. Ultimate quantum limits on phase measurement. *Physical review letters*, 62(20):2377, 1989. 91, 92, 110

[STL+14] S. S. Szigeti, B. Tonekaboni, W. Y. S. Lau, S. N. Hood, and S. A. Haine. Squeezed-light-enhanced atom interferometry below the standard quantum limit. *Phys. Rev. A*, 90:063630, Dec 2014. 109, 126

[Szi29] L. Szilard. Über die entropieverminderung in einem thermodynamischen system bei eingriffen intelligenter wesen. *Zeitschrift für Physik*, 53(11-12):840–856, 1929. 25

T

[Tsa12] M. Tsang. Ziv-zakai error bounds for quantum parameter estimation. *Physical review letters*, 108(23):230401, 2012. 93, 110

V

[VdN13] M. Van den Nest. Universal quantum computation with little entanglement. *Phys. Rev. Lett.*, 110:060504, Feb 2013. 3

[Vid03] G. Vidal. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.*, 91:147902, Oct 2003. 3

[VKL+10] H. Vahlbruch, A. Khalaidovski, N. Lastzka, C. GrÃd'f, K. Danzmann, and R. Schnabel. The geoâĂĽ600 squeezed light source. *Classical and Quantum Gravity*, 27(8):084027, 2010. 75

[VTB13] H. L. Van Trees and K. L. Bell. *Detection estimation and modulation theory, pt. I.* Wiley, 2013. 75

W

[Wer89] R. F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Physical Review A*, 40:4277–4281, 1989. 124

[Woo98] W. K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80:2245–2248, Mar 1998. 61, 66

[WPM09] S. Wu, U. V. Poulsen, and K. Mølmer. Correlations in local measurements on a quantum state, and complementarity as an explanation of nonclassicality. *Physical Review A*, 80:032319, 2009. 5, 28, 34, 46, 131

X

[XWK87] M. Xiao, L.-A. Wu, and H. J. Kimble. Precision measurement beyond the shot-noise limit. *Phys. Rev. Lett.*, 59:278–281, Jul 1987. 6

Y

[YE09] T. Yu and J. H. Eberly. Sudden death of entanglement. *Science*, 323(5914):598–601, 2009. 24, 69

[YMK86] B. Yurke, S. L. McCall, and J. R. Klauder. Su(2) and su(1,1) interferometers. *Phys. Rev. A*, 33:4033–4054, Jun 1986. 99, 138

Z

[ZHB03] V. V. Zhirnov, J. A. Hutchby, and G. I. Bourianoff. Limits to binary logic switch scaling-a gedanken model. *Proceedings of the IEEE*, 91(11):1934–1939, 2003. 1

[Zur93] W. H. Zurek. Preferred states, predictability, classicality and the environment-induced decoherence. *Progress of Theoretical Physics*, 89(2):281–312, 1993. 23

[Zur00] W. H. Zurek. Einselection and decoherence from an information theory perspective. *Annalen der Physik (Leipzig)*, 9(11-12):855–864, 2000. 28, 29, 34, 48, 51

[Zur03] W. H. Zurek. Quantum discord and Maxwell's demons. *Physical Review A*, 67(1):012320, 2003. 31, 34, 49, 51

[ŻHSL98] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein. Volume of the set of separable states. *Phys. Rev. A*, 58:883–892, Aug 1998. 24