Precision Limits and Uncertainty Relations for Quantum Multiparameter Estimation

Simon K. Yung

A thesis submitted for the degree of
Bachelor of Philosophy with Honours in Physics at
The Australian National University

October 2023

Declaration

This thesis is an account of research undertaken between February 2023 and October 2023 at the Research School of Physics, The Australian National University, Canberra, Australia, and the Institute of Materials Research and Engineering, Agency for Science, Research and Technology, Singapore, Republic of Singapore.

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree at any other university.

Simon K. Yung
October 2023

Acknowledgements

First and foremost, I would like to thank my supervisors Syed Assad and Lorcán Conlon, who made this honours project possible. Thank you for sharing your expertise with me and fostering my interest in quantum metrology; it has been a pleasure working with you.

I would also like to thank Ping Koy Lam for giving me the opportunity to visit Singapore and experience another research environment. My time here has been an extremely valuable experience, both scientifically and personally.

I extend my thanks to the A*STAR Quantum Opto-Mechatronics group and the ANU Quantum Optics group, who have made me feel most welcome and supported me through this year. In particular, I thank Assad, Lorcán, Ruvi Lecamwasam, Aritra Das, and Ping Koy, for proofreading this thesis. I also thank Laura-Lynn Liew for helping to organise my A*STAR visit.

My honours year and this thesis would not have been possible without the encouragement of my family and friends. Thank you for everything that you do to support me.

I acknowledge the financial support of the ANU Research School of Physics Dunbar Physics Honours Scholarship and the A*STAR Singapore International Pre-Graduate Award.

Abstract

Determining optimal measurements for simultaneously estimating multiple parameters of a physical system is an important task for many applications. For estimating parameters in quantum systems, the uncertainty principle fundamentally limits the estimation of incompatible parameters—there is a trade-off in terms of the accuracy with which each parameter can be measured. Limits on the precision of estimation have been formulated, in analogy to classical theory, as lower bounds on a weighted sum of estimation variances, such as the Nagaoka–Hayashi Cramér–Rao bound (NHCRB). Recently, an alternative limit called the Lu–Wang uncertainty relation (LWUR) was proposed based on an existing uncertainty relation. It is a trade-off relation between estimation variances, quantifying the effect of Heisenberg's uncertainty principle on the joint estimation of two parameters. Of interest for any theoretical estimation limit is its attainability: whether or not there exists a measurement saturating the limit.

In this work, I investigate the differences between the limits that the NHCRB and LWUR place on estimation variances. I compare the NHCRB and the minimum sum of variances allowed by the LWUR, using a model problem of estimating orthogonal rotations of a quantum state. I find estimation problems where both approaches place an identical limit on the estimation variances. However, there are also estimation problems where the difference between the limits is infinitely large and the LWUR cannot be attained. In particular, I highlight a flaw in the figure-of-merit of estimation performance used to derive the LWUR—there are estimation problems in which measurements optimising this figure-of-merit do not yield any information about the system's parameters. To circumvent this, I present an uncertainty relation for parameter estimation based on the NHCRB, which can define an attainable uncertainty relation. Additionally, I show that this approach can provide non-trivial trade-off relations for estimating more than two parameters, something that the LWUR cannot do.

List of Acronyms

The following acronyms are used within this thesis. While the list is long, those that are commonly used share terminology, such as the Cramér–Rao bounds (CRBs) or those relating to the symmetric logarithmic derivative SLD. For reference, I provide the defining equation number or relevant section number.

Acronym	Meaning	Equation/Section
CCRB	classical Cramér–Rao bound	Eq. (3.16)
CFIM	classical Fisher information matrix	Eq. (3.13)
HCRB	Holevo Cramér–Rao bound	Eq. (3.27)
IRTR	information regret trade-off relation	Eq. (3.34)
KKT	Karush–Kuhn–Tucker	Sec. 4.1.2
LWB	Lu–Wang estimation bound	Eq. (4.1)
LWUR	Lu–Wang uncertainty relation	Eq. (3.37)
MLE	maximum likelihood estimator	Eq. (3.7)
MSE	mean squared error	Eq. (3.10)
NCRB	Nagaoka Cramér–Rao bound	Eq. (3.30)
NHCRB	Nagaoka–Hayashi Cramér–Rao bound	Eq. (3.32)
PVM	projection-valued measure	Sec. 2.2.1
POVM	positive operator-valued measure	Sec. 2.2.2
SLD	symmetric logarithmic derivative	Eq. (3.21)
SLDCRB	SLD Cramér–Rao bound	Eq. (3.24)
SLDFIM	SLD Fisher information matrix	Eq. (3.22)

Contents

Chapter 1

Introduction

Accurate and precise measurements are fundamental to experimental tests and further development of physical theories [1, 2]. Historically, increasing the precision of measurements enabled us to reach milestones in understanding our universe. For example, the observation of spectral lines influenced the development of the Bohr model and quantum theory, and precise experimental verifications of predictions made by quantum electrodynamics gave credibility to quantum field theories [3]. The obstacles created by noise and measurement bias remain in current scientific pursuits, but can be mitigated thanks to advancements in metrology, the study of measurement.

In the pursuit of attaining the ultimate precision of measurements, apparatus imperfections that introduce sources of noise and bias can be eliminated. However, there will remain limitations on measurement precision due to noise that cannot be explained by classical theory: inherent quantum noise owing to the intrinsically probabilistic nature of quantum mechanics. A complete understanding of measurement therefore requires a quantum theory. This necessity is the basis of quantum metrology, which also provides a framework to utilise quantum-mechanical properties such as entanglement and squeezing to obtain better precision than is classically allowed [4].

The theory of quantum metrology has been used to great effect in a range of applications. Notably, squeezed light—which has sub-classical uncertainty in one physical quantity, such as phase, and correspondingly increased uncertainty in another—was used to increase the sensitivity of the gravitational wave detectors, increasing the expected rate of detection by 50% [5, 6]. The theory has also been applied to resolving points sources of light, such as distant stars, beyond the classical Rayleigh limit, called quantum superresolution [7]. There are also interdisciplinary applications of quantum metrology, such as enhanced microscopy signal-to-noise ratio for biological imaging where light intensity is limited to avoid inflicting damage to the sample [8].

In this thesis, I focus on the simultaneous measurement of multiple properties of a quantum system—called multiparameter estimation—and the limits on the precision of such measurements. In this case, the ultimate limits on precision are influenced not only by noise in the measurement itself but also by Heisenberg's uncertainty principle [9–12]. That is, there are quantities which cannot be simultaneously measured with unlimited precision, and this induces limits on the precision of their joint measurement. Quantum theory imposes this unavoidable limitation but at the same time it also provides avenues for improving measurement precision using entanglement, which has recently been demonstrated experimentally [13].

The basis of quantum multiparameter estimation theory is built on the well-established theory of classical estimation, which mathematically describes the inference of quantities via the sampling of a parameter-dependent probability distribution. The performance of an estimation procedure is quantified by its mean squared error, which is, roughly, the expected spread of the estimated values compared to the true value, and will be rigorously defined in Chapter 3. For estimating multiple parameters, the objective is to minimise the sum of the mean squared error of each parameter.

When parameters encoded in a quantum state are to be estimated, estimation theory must describe both the physical quantum measurement and the classical post-processing of the measurement outcomes. In general, there are many possible quantum measurements that can be performed. This "choice" of measurement complicates the minimisation of the mean squared error and renders brute-force optimisation intractable, even for simple systems. Instead, theorists have derived measurement-independent lower bounds on the sum of mean squared errors that are tractable to compute [14, 15]. Numerous such lower bounds exist, with different degrees of attainability—whether a measurement exists that achieves the lower bound—and assumptions about measurement scenarios. Tighter[1] lower bounds are desirable because they offer more information about the ultimate precision limit. The general attainability of these lower bounds remains an open question.

An alternative limit on multiparameter estimation precision was recently proposed

[1] A *tight* lower bound is one that is attainable. A lower bound X is *tighter* than lower bound Y if $X > Y$, meaning that X is closer to being tight.

based on an uncertainty relation that directly quantifies the trade-off in the mean squared error of each parameter due to Heisenberg's uncertainty principle [16]. The limit is a non-trivial inequality that limits measurement precision in a fundamentally different way to that done by mean squared error sum-lower-bounds, which do not directly provide information about the distribution of the mean squared errors. Additionally, the most attainable lower bounds require optimisation to compute while the trade-off relation does not, rendering it easier to calculate for a given problem. These differences beg several open questions:

(i) Does this alternative approach produce the same lower bound on the sum of the mean squared errors as existing lower bounds?

(ii) Is the uncertainty-relation-based limit attainable?

(iii) If it is not attainable, can a better, more attainable trade-off relation be developed?

These are the questions that I investigated for this thesis. I developed methods that enable the comparison of the different estimation limits, and I searched for answers to the questions using a combination of general theoretic comparisons, simple model estimation problems, and numerical simulations. This combination allowed for a holistic comparison and for conclusions to be drawn about the conditional equivalence of the estimation limits.

The outline of this thesis is as follows: in Chapter 2, I outline the quantum mechanical framework for measurements and discuss uncertainty relations. In Chapter 3, I formalise estimation theory and present the existing precision limits for quantum multiparameter estimation. In Chapter 4, I discuss the methods used in the rest of the thesis to compare estimation limits. In Chapter 5, I present results pertaining to generic quantum estimation problems, with some constraints. In Chapters 6 and 7, I use simple model estimation problems to gain insight into the difference between the precision limits and further investigate multiparameter estimation uncertainty relations. In Chapter 8, I present results based on numerical simulations involving random estimation problems and random measurements. Finally, in Chapter 9, I discuss the results of this thesis and avenues for further work, and provide concluding remarks.

Chapter 2

Quantum Measurements and Uncertainty Relations

Before delving into the details of the topic of this thesis, quantum parameter estimation, I briefly outline the quantum theory necessary for its understanding. Parts of this theoretical basis appear in standard quantum mechanics courses, but I present a summary to clarify the notation, largely following the notation in Ref. [17]. I then discuss measurements of quantum systems, how to describe them mathematically, and how this description can be translated into a physical realisation of the measurement. Finally, I discuss uncertainty relations in quantum mechanics.

2.1 Quantum Systems

Let $\mathcal{H}_d$ denote the d-dimensional complex Hilbert space corresponding to a quantum system, and let $|\psi\rangle \in \mathcal{H}_d$ denote a state of the system. The inner product between a state $|\psi\rangle$ and a conjugate-transpose state $\langle\varphi|$, is denoted by $\langle\varphi|\psi\rangle$. We consider states that are normalised such that $\langle\psi|\psi\rangle = 1$. Operators on $\mathcal{H}_d$ can be represented by $d \times d$ complex matrices and act on a state $|\psi\rangle$ represented by a column vector of size d, whose entries correspond to the decomposition of $|\psi\rangle$ into an orthonormal basis of $\mathcal{H}_d$. Important classes of operators are: Hermitian or self-adjoint operators for which $H^\dagger = H$ (with $\dagger$ denoting the conjugate-transpose) which have real eigenvalues, and unitary operators for which $U^\dagger U = UU^\dagger = \mathbb{1}_d$. Here, $\mathbb{1}_d$ denotes the d-dimensional identity operator.

The density operator formalism allows for the possibility of classical mixtures of quantum states by representing a state of the system with a density operator ρ, i.e., a positive semidefinite Hermitian $d \times d$ matrix with unit trace. This allows for describing, for example, a system prepared probabilistically in one of several states. In this

representation, the pure state $|\psi\rangle$ has density operator $\rho_{\text{pure}} = |\psi\rangle\langle\psi|$. A state which cannot be written in this form (i.e., its density operator is not rank-1[1]) is called a mixed state. The eigenvalues of the density operator correspond to the (classical) probability that the state was prepared in the corresponding eigenstate—hence, the trace condition $\text{Tr}[\rho] = 1$ ensures that the classical probabilities sum to unity.

Throughout this thesis, I will often use a two-dimensional quantum system (called a *qubit*) in examples and as a platform for new research. Qubits are useful due to their relative simplicity compared to systems with higher-dimensional Hilbert spaces and are commonly used for quantum computational tasks [17]. At times, I will use higher-dimensional systems, such as three-dimensional (called *qutrits*) or arbitrary d-dimensional (called *qudits*). These systems have potential practical utility due to their larger number of degrees of freedom and significant work has been devoted to their experimental realisation [18, 19].

Qubits can be realised, and have been used for quantum metrology, in many physical systems, including photonic [20], superconducting electrons [21], and trapped ions [22]. When not specialising to a specific physical system, the basis states of a qubit will often be denoted $|0\rangle$ and $|1\rangle$, called the computational basis. These states are associated to the vectors $(1,0)^\top$ and $(0,1)^\top$, respectively. In the case of a qubit encoded in the polarisation of a photon, I will use the horizontal and vertical polarisations as the basis states, denoted $|1\rangle_{\text{H}}$ and $|1\rangle_{\text{V}}$, respectively. Here, the "1" denotes that it is a state with a single photon. In this specialised context, the state with zero photons, called the vacuum state, is also denoted $|0\rangle$, and is not to be confused with the generic qubit state $|0\rangle$.

A common representation of a qubit is the correspondence between a qubit and a unit sphere, called the Bloch sphere. The correspondence is built by recalling that the state should be normalised, so a valid qubit state is $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ with $|\alpha|^2 + |\beta|^2 = 1$. However, the global phase of α and β is unimportant, so a general qubit state can be parameterised as

$$|\psi\rangle = \cos\frac{\vartheta}{2}\,|0\rangle + e^{i\varphi}\sin\frac{\vartheta}{2}\,|1\rangle\,, \tag{2.1}$$

[1]The rank of a matrix is equal to the number of its non-zero eigenvalues (including degeneracy).

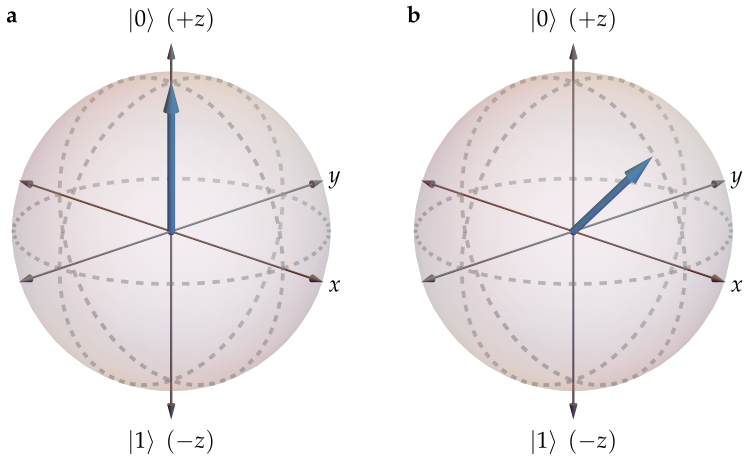

Figure 2.1: Depiction of the Bloch sphere. **a**, the pure state $|0\rangle$. **b**, the mixed state with Bloch representation $(r, \vartheta, \varphi) = (1/\sqrt{2}, \pi/4, \pi/4)$ corresponding to the state with projections onto the x, y, z-axes $(s_x, s_y, s_z) = (1/\sqrt{8}, 1/\sqrt{8}, 1/2)$, respectively.

with real numbers ϑ and φ [17]. These numbers also parameterise the polar angle and azimuthal angle, respectively, of a point on the surface of a sphere, where the north pole (positive z-axis) is identified with the state $|0\rangle$, and the south pole with the state $|1\rangle$, as depicted in Figure 2.1. The x- and y-axes are identified with particular superpositions of the computational basis states. Mixed states lie within the Bloch sphere and are described by an additional radial coordinate $0 \leq r \leq 1$. The qubit state's density matrix is given by

$$\rho = \frac{1}{2}\left(\mathbb{1} + s_x\sigma_x + s_y\sigma_y + s_z\sigma_z\right), \tag{2.2}$$

where σ_i are the Pauli operators

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.3}$$

and s_i are related to the polar coordinates r, ϑ, φ by

$$s_x = r\cos\varphi\sin\vartheta, \quad s_y = r\sin\varphi\sin\vartheta, \quad s_z = r\cos\vartheta. \tag{2.4}$$

In this representation, the basis states $|0\rangle$ and $|1\rangle$ are the eigenstates of the operator σ_z, chosen to represent the two levels of a physical spin-$\frac{1}{2}$ system.

2.1.1 Quantum Operations

The evolution of a pure quantum state can be described by the unitary transformation $|\psi\rangle \rightarrow U|\psi\rangle$. The condition that U is unitary ensures that the normalisation of the state is unchanged, as unitary operators preserve the inner product. The evolution of a state described by a density operator is $\rho \rightarrow U\rho U^\dagger$.

The simple mathematical description of evolution means that to be able to perform any manipulation of a quantum state, it is sufficient to be able to implement any unitary operation[2]. A common approach to realising interesting physics of quantum systems is to implement quantum circuits on quantum computers. At the risk of oversimplification, a quantum circuit consists of a preparation stage, where a quantum state is prepared according to a classical input, an evolution stage, where the unitary transformation is implemented, and finally, a measurement stage, where a classical output is obtained and processed. A universal quantum computer is a device capable of performing all such operations.

The aforementioned measurement stage typically measures the quantum state in the computational basis $\{|0\rangle, |1\rangle, \dots\}$. A measurement in this basis of the state ρ gives rise to an outcome from the set $\{0, 1, \dots\}$ with probability given by the Born rule: $p(n) = \mathrm{Tr}[|n\rangle\langle n|\,\rho]$. In practice, this "outcome" may mean the "click" of the detector with label n. The label of each outcome need not have any physical significance. By repeating the quantum circuit many times, the statistics of the outcomes can be recorded and used to infer properties of the quantum state.

As an example, consider a qubit encoded in the polarisation of a photon: $|\psi\rangle = \alpha|1\rangle_\mathrm{H} + \beta|1\rangle_\mathrm{V}$ with $|\alpha|^2 + |\beta|^2 = 1$. This qubit can be manipulated using linear optical elements: half-wave plates (HWP) and quarter-wave plates (QWP). In this case, measurement in this basis can be performed by using a polarising beam splitter, which spatially separates horizontal and vertical polarisations, and a detector at each output port.

Measurement in the computational basis is not the only measurement that can be performed on a quantum system—in the following section, I discuss general measure-

[2] In the context of implementing measurements, non-unitary dynamics (e.g., loss) are not relevant, so are ignored in this thesis.

ments in quantum mechanics.

2.2 Measurements in Quantum Mechanics

In quantum mechanics courses, we are taught that each measurable quantity corresponds to a self-adjoint (Hermitian) operator [23]. These operators are often called "observables". Being Hermitian, observables have real eigenvalues, which we associate with the possible quantitative outcomes of measuring that physical quantity. It remains to understand how we can mathematically describe and physically implement a measurement, two somewhat distinct tasks.

2.2.1 Projective Measurements

Given an observable A on $\mathcal{H}_d$, a measurement of A can be achieved by a *projective measurement* on its eigenbasis [17]. Owing to Hermiticity, A has an orthonormal eigenbasis $\{|\psi_1\rangle, \ldots, |\psi_d\rangle\}$ with eigenvalues $\lambda_1, \ldots, \lambda_d$ so that $A = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$. A projective measurement can be described by projection operators satisfying $P_i^2 = P_i$ (for example, $P_i = |\varphi_i\rangle\langle\varphi_i|$, or operators which project onto a larger subspace), that collectively satisfy $\sum_i P_i = \mathbb{1}_d$ and form what is called a "projection-valued measure" (PVM) [24]. This summation property is a manifestation of outcome probabilities summing to unity because the probability of a state ρ being detected in each eigenstate is given by the Born rule as $p(i) = \mathrm{Tr}[P_i\rho]$. Additionally, a PVM requires that the operators are orthogonal: $P_iP_j = 0$ if $i \neq j$.

Notionally, a projection operator represents a detector that 'clicks' indicating the state collapsed to the corresponding eigenstate, with probability given by the Born rule. This can be implemented on a universal quantum computer by first implementing the unitary operation which transforms the desired eigenbasis into the computational basis, and then measuring in the computational basis. Of course, measurement in the computational basis is itself an example of a projective measurement.

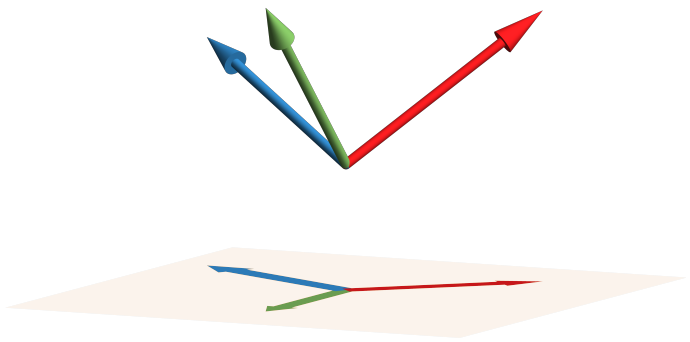

Figure 2.2: Visualisation of Naimark's dilation theorem for rank-1 operators. A set of three vectors in two dimensions that are not mutually orthogonal (lower) can be realised as the projection of a mutually orthogonal set in a higher dimension (upper).

2.2.2 Generalised Measurements

A limitation of the aforementioned projective measurements is that the dimension of the Hilbert space limits the number of outcomes. As it turns out, a more generalised measurement can be described by a "positive operator-valued measure" POVM, in which the orthogonality condition of a PVM is relaxed [17, 24]. A POVM is a set of positive semidefinite self-adjoint operators $\Pi \equiv \{\Pi_i\}$ (called POVM elements) which sum to the identity operator. This way, the number of measurement outcomes is not restricted, which can provide more information when paired with appropriate post-processing. Every PVM is a POVM, but not vice versa. The probability of the outcome associated with each POVM element is given by the Born rule:

$$p(i) = \mathrm{Tr}[\Pi_i \rho]. \tag{2.5}$$

The ability to implement a POVM is guaranteed by Naimark's dilation theorem, which states that every POVM can be implemented as a projective measurement in an enlarged space [25]. That is, the measurement statistics described by the POVM on $\mathcal{H}$ are identical to the statistics of some PVM on $\mathcal{H} \otimes \mathcal{H}_a$, where $\mathcal{H}_a$ is some ancillary space. Intuitively, for rank-1 operators, this can be understood as a consequence of the fact that a set of non-orthogonal vectors can be realised as the projection of a set of mutually orthogonal vectors in a higher dimension, as illustrated in Figure 2.2.

Naimark's dilation theorem is of practical importance to the realisation of a POVM. It also highlights the utility of describing measurements using POVMs: a POVM may describe a more general measurement on the state–ancillary space, which has more degrees of freedom, without considering the ancillary space's particulars. Of course, the ancillary space must be explicitly considered when physically realising a POVM. The implementation of general POVMs has been the topic of much research, with progress made using quantum walks [26–29] and probabilistically implemented projective measurements combined with post-processing and postselection [30–32].

2.2.3 Separable and Collective Measurements

In measuring a property of any system, measurements are performed on identically prepared copies of the system to average over noise and learn about the system's statistics. It is typical to perform measurements on one copy at a time, in analogy to the classical case. I refer to this as a single-copy or separable measurement, because the overall measurement operators can be viewed as tensor-products[3] $M^{\otimes N}$ acting on the tensor-product state $\rho^{\otimes N}$. A separable measurement is depicted in Figure 2.3**a**.

In quantum systems, via entanglement, a measurement can also be performed simultaneously on multiple copies of the system, called a collective measurement and depicted in Figure 2.3**b**. In this case, the state $\rho^{\otimes N}$ is regarded as a single higher-dimensional state, and the higher-dimensional measurement operator cannot be separated into N single-copy measurements. Fundamentally, collective measurement does not have a classical analogue—it is a uniquely quantum concept.

A collective measurement has more degrees of freedom than a separable measurement and can extract information from the system more efficiently [33, 34], though are more experimentally complex. Nevertheless, recent quantum computing advances have made them possible to implement experimentally [13]. Details of the advantage of collective measurements will be discussed in Section 3.2.3.

[3] The notation $X^{\otimes N}$ denotes the N-fold tensor product $X \otimes \cdots \otimes X$.

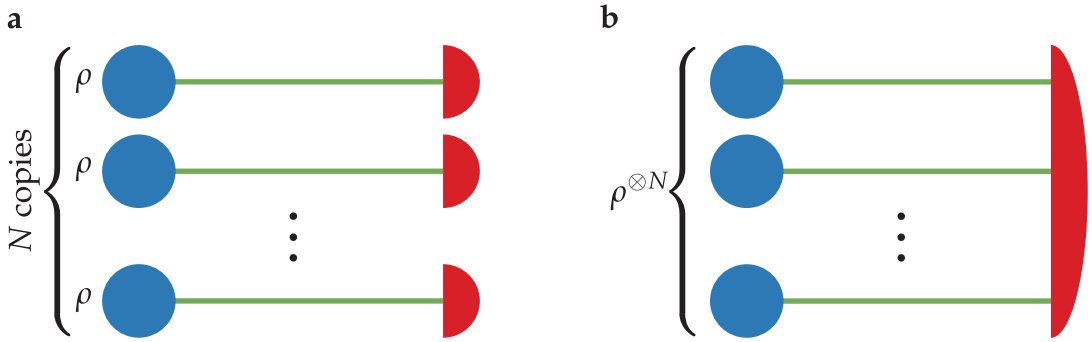

Figure 2.3: Separable and collective measurements. The blue circles represent the quantum state (ρ), the green lines represent a quantum channel, e.g., for encoding parameters, and the red semicircles represent the measurement. **a**, N separable (single-copy) measurements are performed individually on each copy of the state. **b**, a collective measurement is performed on N copies of the state simultaneously.

2.2.4 Examples of Measurement Implementation

In the following, I outline two practical examples of how POVMs can be implemented.

A Simple Example

Consider a polarisation-encoded qubit system with basis states $|1\rangle_{\mathrm{H}}$ and $|1\rangle_{\mathrm{V}}$, and the three-outcome POVM with elements

$$\Pi_1 = \frac{1}{2}|1\rangle\langle 1|_{\mathrm{H}}, \; \Pi_2 = \frac{1}{2}|1\rangle\langle 1|_{\mathrm{V}}, \; \Pi_3 = \frac{\mathbb{1}_2}{2} = \frac{1}{2}(|1\rangle\langle 1|_{\mathrm{H}} + |1\rangle\langle 1|_{\mathrm{V}}). \tag{2.6}$$

Here, Π_1 and Π_2 are proportional to projection operators of the computational basis, but Π_3 is full-rank. By the Born rule (Eq. (2.5)), for any state, one of outcomes 1 and 2 occurs with 50% probability (their probabilities sum to $1/2$ for any state), and outcome 3 occurs with the other 50%. This POVM can be implemented with a balanced beam splitter, a polarising beam splitter, and three photodetectors, as illustrated in Figure 2.4**a**. Whilst it may appear that this implementation does not use an ancillary space, formally, we must consider the vacuum state ($|0\rangle$) at the second input port of the balanced beam splitter. This accounts for the possibility that the photon is not detected at either detector 1 or 2.

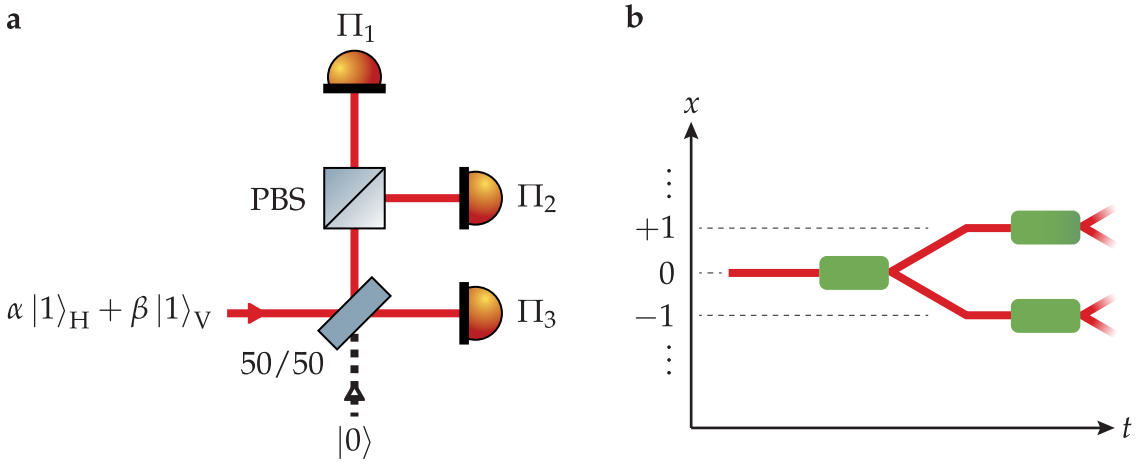

Figure 2.4: **a**, Schematic of an optical implementation of the POVM in Eq. (2.6), involving a balanced beam splitter (50/50), a polarising beam splitter (PBS), and three photodetectors. Diagram made using ComponentLibrary [35]. **b**, Illustration of a discrete quantum walk. The red lines represent the paths the system takes and the green boxes indicate the unitary transformations consisting of the coin operation and conditional translation operation.

Quantum Walks

A quantum walk is the analogue of a classical random walk where the random steps occur at probabilities dictated by a quantum system [36, 37]. Quantum walks have proven to be useful for quantum computational algorithms, but here, I focus on their utility for implementing POVMs. Quantum walks have been used to implement general measurements on qubits [26–28] and qudits [29], and recently to implement a two-copy collective measurement [38]. In the following, I provide a basic description that is sufficient to describe the two-copy collective measurement.

Here, I consider only discrete, one-dimensional quantum walks. Consider a system with a discrete position degree of freedom with possible states $\{|x\rangle \mid x = 0, \pm 1, \dots\}$ (called the "walker"), and an additional "coin" degree of freedom which dictates the random steps. In the simplest case, the coin is a qubit, but higher dimensional versions are also possible [29]. A quantum walk consists of a series of unitary operators that manipulate the coin and then translate the walker (increment x by ± 1) depending on the coin state, as illustrated in Figure 2.4**b**. Due to superposition, the overall state can take many paths, and the probability of detection at the end each path (which depends on the initial coin state) can be designed to match that of a POVM by appropriately choosing the unitary coin operations. For single-copy POVMs, the position degree-of-

freedom forms the ancillary space and any number of outcomes can be implemented by choosing the number of steps to be large enough [26]. Hou *et al.* implemented a two-copy qubit measurement by designing a quantum walk where positions ± 1 of the walker form a second copy of the coin qubit, and all other walker positions form the ancillary space [38].

2.3 Uncertainty Relations

The possibility in quantum mechanics of operators that do not commute ($AB \neq BA$) has significant implications for the simultaneous measurement of multiple quantities of a quantum system, commonly known as "the uncertainty principle". In this section, I review common uncertainty relations and introduce an uncertainty relation that applies to measurements.

2.3.1 Heisenberg's Uncertainty Principle

The earliest notion of the uncertainty principle is attributed to Heisenberg's seminal paper in 1927 [39]. Heisenberg realised, via his γ-ray thought experiment, that knowing the position of a quantum-mechanical particle to a certain precision limits the precision with which its momentum can be determined. Though largely qualitative, Heisenberg's notion of uncertainty principle refers to the unavoidable disturbance in the momentum imparted by the position measurement. This is distinct from uncertainty relations about the preparation of a quantum state, which I discuss next.

2.3.2 Preparation Uncertainty Relations

Heisenberg's uncertainty relation, as it is now commonly taught in quantum mechanics courses, was rigorously proven separately by Kennard [40] and Weyl [41]. For position operator Q and momentum operator P, this uncertainty relation is

$$\sqrt{\mathrm{Var}(Q)}\sqrt{\mathrm{Var}(P)} \geq \frac{\hbar}{2}, \tag{2.7}$$

where for an observable A, its variance is $\mathrm{Var}(A) = \langle A^2 \rangle - \langle A \rangle^2$, and $\langle \cdot \rangle$ denotes the expectation value with respect to the quantum state of interest. This relation is a statement about the precision with which a state's position and momentum can be *prepared*—that is, it is not possible for a particle to simultaneously have an arbitrarily narrow spread in both position and momentum. As such, the term *statistical dispersion princple* was proposed by Ballentine [42], to distinguish from Heisenberg's original notion of uncertainty principle. Experimentally, this position spread can be quantified by measuring the position on several independent, identically prepared states and similarly for the momentum, but not measuring the position and momentum jointly. The position–momentum preparation uncertainty relation was generalised to any pair of observables A and B by Robertson [43] and then strengthed by Schrodinger [44], who obtained

$$\mathrm{Var}(A)\,\mathrm{Var}(B) \geq \frac{1}{4}\langle i[A,B] \rangle^2 + \mathrm{Cor}(A,B)^2, \quad \mathrm{Cor}(A,B) := \frac{1}{2}\langle \{A,B\} \rangle - \langle A \rangle \langle B \rangle, \tag{2.8}$$

where $\mathrm{Cor}(A,B)$ is called the correlation between A and B, $[A,B] = AB - BA$ is the commutator of A and B, and $\{A,B\} = AB + BA$ is the anticommutator of A and B. This says that for any pair of non-commuting observables, a state cannot exist with arbitrarily sharp distributions in both variables.

2.3.3 Uncertainty Relations for Measurement

The shortcoming of the Robertson–Schrödinger uncertainty relation in an experimental context is that it has no bearing on the limitations on the accuracy of measurement devices which are designed to measure both quantities [42]. This was resolved by Ozawa, who developed a joint-measurement uncertainty relation [45], relating the "error" in measuring observable A and B, and their inherent preparation uncertainties. These preparation uncertainties, denoted σ_A, σ_B, are defined as the square root of the variances, e.g., $\sigma_A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}$.

The basis of Ozawa's uncertainty relation is an approximate joint measurement of incompatible observables. Approximate joint measurements utilise the fact that incompatible observables A and B on a Hilbert space $\mathcal{H}$ can be approximated by com-

muting observables $\mathcal{A}$ and $\mathcal{B}$ on a larger Hilbert space $\mathcal{H} \otimes \mathcal{K}$ by the inclusion of an ancillary system in state $|\xi\rangle$ of a Hilbert space $\mathcal{K}$. Because $\mathcal{A}$ and $\mathcal{B}$ commute, the two approximate observables can be jointly measured. The root-mean-squared error in the approximate measurement of $A(B)$ is denoted $\varepsilon_{A(B)}$, and measures how "far" the approximate observable is from the true observable. Ozawa [45, 46] developed a relation limiting the possible values of the errors $\varepsilon_A, \varepsilon_B$ and the preparation uncertainties σ_A, σ_B, which he referred to as a "universally valid uncertainty relation for joint measurement". Branciard [47] improved Ozawa's uncertainty relation, for pure states, to an attainable (tight) trade-off relation for joint measurements. Ozawa [48] later improved Branciard's result, strengthening the uncertainty relation for mixed states. This universally valid uncertainty relation is

$$\varepsilon_A^2 \sigma_B^2 + \varepsilon_B^2 \sigma_A^2 + 2\sqrt{\sigma_A^2 \sigma_B^2 - D_{AB}^2}\, \varepsilon_A \varepsilon_B \geq D_{AB}^2, \tag{2.9}$$

where

$$D_{AB} = \frac{1}{2} \operatorname{Tr} \left| \sqrt{\rho}[A, B] \sqrt{\rho} \right|, \tag{2.10}$$

where $\sqrt{\cdot}$ denotes the matrix square-root and for an operator C, its absolute value is defined as $|C| = \sqrt{C^\dagger C}$. The coefficient D_{AB} can be seen as a generalisation of the first term on the right-hand side of the Robertson–Schrödinger uncertainty relation (Eq. (2.8)). Ultimately, the joint-measurement uncertainty relation reflects that two non-commuting observables cannot be simultaneously measured with arbitrary precision.

Chapter 3

Parameter Estimation and Precision Limits

Before discussing quantum statistical inference and parameter estimation, it will be prudent to outline the classical analogue. Classical statistical inference is implicitly used in experiments designed to measure a quantity, being the framework for processing measurement results into estimates of the quantities of interest. This is especially important when the quantity of interest is not directly measurable and must be inferred from the results of indirect measurement. In this chapter, I describe the general classical and quantum parameter estimation process and discuss lower bounds on estimation performance. Finally, I discuss the connection between uncertainty relations and parameter estimation.

3.1 Classical Parameter Estimation

Parameter estimation refers to the inference of the value of a system's parameter(s) based on sampling of its probability distribution. This general process is depicted in Figure 3.1. Any measurement can be viewed as random sampling from a probability distribution, so, mathematically, we may represent the system by its statistical distribution. This is based on the principle that a system can be independently and identically prepared infinitely many times and that the outcomes will be distributed according to a definite probability distribution [49].

Deviation of the estimate from the true value of the parameter, i.e., *error*, arises due to the statistical nature of measurements. The exact probability distribution of measurement outcomes cannot be known with any finite number of measurements. A central question in the field of parameter estimation is how this error can be reduced and by how much: what is the smallest average error that can be achieved?

Throughout this thesis, I will use the vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)^\top \in \mathbb{R}^n$ to denote

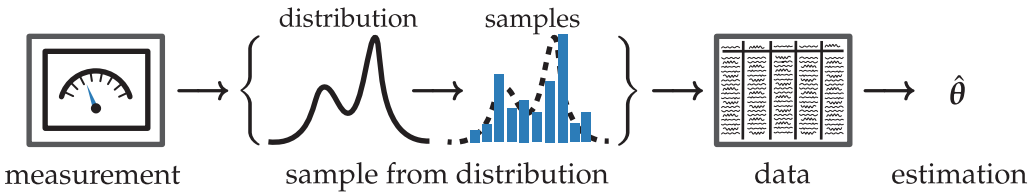

Figure 3.1: Depiction of a parameter estimation process. Measurement of a system can be viewed as random sampling from the system's parameter-dependent statistical distribution. The measured data is processed to obtain an estimate of the system's parameters or, equivalently, the parameters of the probability distribution.

both the parameters of interest and their true values—the distinction should be clear from the context. Without loss of generality, we assume that the parameters are real-valued—a complex-valued parameter can be separated into two real parameters.

3.1.1 Local Parameter Estimation

In this thesis, I restrict attention to *local* parameter estimation instead of *global* parameter estimation. In global estimation, the aim is to optimise the estimation procedure by maximising some suitable metric averaged over all possible values of the parameter. On the other hand, local estimation is concerned only with optimising the process for a particular value of the parameters. In this way, the estimation performance is optimal for sensing small changes in the parameters around a known value. We can reparameterise a local estimation problem to estimating parameters close to zero.

The assumption of local estimation is reasonable when we have a large number N of copies of the system. First, we can use $\sqrt{N}$ of the copies to obtain a rough estimate of the parameters, and then knowing this estimate, we may apply local parameter estimation on the remaining $N - \sqrt{N}$ copies to improve our estimate. For large N, the $\sqrt{N}$ deficit becomes negligible, so most of our resources are used for the local estimation.

3.1.2 Example: Biased Coin Toss (Bernoulli distribution)

I use a simple example to contextualise some important concepts in the rest of this section. Consider a biased coin toss where the probability of landing on heads (H) or

tails (T) is unknown. The coin toss may be described by the probabilities

$$p(\mathrm{H}) = \theta,\ p(\mathrm{T}) = 1 - \theta,\ 0 \leq \theta \leq 1. \tag{3.1}$$

We imagine tossing the coin N times and using the results to estimate θ.

3.1.3 Estimators

An essential step in parameter estimation is converting measurement outcomes into an estimate of the system's parameters. Mathematically, this is done via an estimator function, denoted with a 'hat': $\hat{\boldsymbol{\theta}}$. In an experiment, many measurements can be made on samples of the system, and these results are input to the estimator function. The measurement outcomes may be discrete or continuous and will often not be the parameter itself. That is, the parameters need to be inferred from the measurement outcomes. For example, the mean of a distribution is estimated by averaging samples after assigning values to outcomes, if necessary.

An important property of an estimator is its *bias*, defined as the difference between the expected value of the estimate and the true value of the parameter

$$\mathrm{Bias}(\hat{\boldsymbol{\theta}}) = \mathbb{E}\big[\hat{\boldsymbol{\theta}}\big] - \boldsymbol{\theta}, \tag{3.2}$$

where $\mathbb{E}$ denotes the expectation value with respect to all possible outcomes. Naturally, we would prefer an estimator that has no bias—an unbiased estimator—so that the estimate based on a large sample converges to the true value of the parameter. Under the assumption of locality for parameter estimation, we enforce the weaker condition of locally unbiased estimators: the estimator is unbiased at the true (fixed) value $\boldsymbol{\theta}_0$ and up to first order in $\boldsymbol{\theta}$ around this point:

$$\mathrm{Bias}(\hat{\boldsymbol{\theta}})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = \mathbf{0}, \quad \frac{\partial}{\partial\theta_i}\mathrm{Bias}(\hat{\boldsymbol{\theta}})\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = \underbrace{(0,\ldots,0,1,0,\ldots,0)^\top}_{i^{\text{th}}\text{ coordinate non-zero}},\ i = 1,\ldots,n. \tag{3.3}$$

The locally unbiased conditions are sufficient for local parameter estimation because we do not expect to estimate values of $\boldsymbol{\theta}$ far from $\boldsymbol{\theta}_0$.

For discrete probability distributions, the estimator for the parameter θ_j can be written in terms of estimator coefficients $\hat{\theta}_{jk}$, where the index k labels the outcome k, one of the possible measurement results. The estimator coefficients are then processed with the conditional probabilities of the outcomes $p(k|\boldsymbol{\theta})$ (or, in practice, the observed frequencies) to obtain the estimates

$$\hat{\theta}_j = \sum_k \hat{\theta}_{jk} p(k|\boldsymbol{\theta}). \tag{3.4}$$

The locally unbiased conditions for the estimator coefficients are then

$$\sum_k \hat{\theta}_{jk} p(k|\boldsymbol{\theta}) = \theta_{0j}, \quad \sum_k \hat{\theta}_{jk} \frac{\partial}{\partial \theta_i} p(k|\boldsymbol{\theta}) = \delta_{ij}, \tag{3.5}$$

where δ_{ij} is the Kronecker delta and θ_{0j} is the j^{th} element of $\boldsymbol{\theta}_0$.

Maximum Likelihood Estimator

The maximum likelihood estimator (MLE) is a commonly used estimator function that, as we will see, has desirable asymptotic properties, i.e., it performs well in the limit of a large sample size [50]. In essence, the MLE provides the parameter value that was most likely to produce the observed data sample. Mathematically, it is based on the joint probability distribution of the observed data and the parameter value $\boldsymbol{\theta}$, treated as a function of $\boldsymbol{\theta}$:

$$\ell(\boldsymbol{\theta}) = \log p(X|\boldsymbol{\theta}), \tag{3.6}$$

where $p(X|\boldsymbol{\theta})$ is the probability of the series of outcomes $X = \{x_1, x_2, \ldots, x_N\}$ conditioned on $\boldsymbol{\theta}$. In statistics, $\ell(\boldsymbol{\theta})$ is called the log-likelihood function [50]. The logarithm is included so that the contributions from each (independently sampled) outcome in the series X is summed, because the joint probability of the series is the product of the individual probabilities. The MLE is then defined, as its name suggests, as the value of $\boldsymbol{\theta}$ which maximises the likelihood function:

$$\hat{\boldsymbol{\theta}}_{\text{MLE}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\, \ell(\boldsymbol{\theta}). \tag{3.7}$$

The MLE is asymptotically unbiased—its bias tends to zero as the sample size tends to infinity [49].

Biased Coin Toss Estimator

In this example, the log-likelihood function of obtaining N_{H} heads out of N trials is

$$\ell(\theta) = \log\left(\theta^{N_{\mathrm{H}}}(1-\theta)^{N-N_{\mathrm{H}}}\right) = N_{\mathrm{H}}\log\theta + (N-N_{\mathrm{H}})\log(1-\theta), \tag{3.8}$$

which achieves its maximum at $\theta = N_{\mathrm{H}}/N$. That is, the MLE is $\hat{\theta}_{\mathrm{MLE}} = N_{\mathrm{H}}/N$. This estimator is unbiased:

$$\mathbb{E}_{N_{\mathrm{H}}}[\hat{\theta}_{\mathrm{MLE}}] = \sum_{N_{\mathrm{H}}=0}^{N} p(N_{\mathrm{H}})\frac{N_{\mathrm{H}}}{N} = \sum_{N_{\mathrm{H}}=0}^{N}\binom{N}{N_{\mathrm{H}}}\theta^{N_{\mathrm{H}}}(1-\theta)^{N-N_{\mathrm{H}}}\frac{N_{\mathrm{H}}}{N} = \theta, \tag{3.9}$$

where $p(N_{\mathrm{H}})$ is the probability of N_{H} out of N outcomes being heads (the order of outcomes is unimportant).

3.1.4 Quantifying Estimation Performance

The error of a parameter estimate can be rigorously defined by the mean squared error (MSE) matrix with elements

$$[V(\hat{\boldsymbol{\theta}})]_{ij} = \mathbb{E}\left[\left(\hat{\theta}_i - \theta_i\right)\left(\hat{\theta}_j - \theta_j\right)\right]. \tag{3.10}$$

If the estimator is unbiased, then the MSE matrix coincides with the estimator's covariance matrix

$$[\mathrm{Cov}(\hat{\boldsymbol{\theta}})]_{ij} = \mathbb{E}\left[\hat{\theta}_i\hat{\theta}_j\right] - \mathbb{E}\left[\hat{\theta}_i\right]\mathbb{E}\left[\hat{\theta}_j\right]. \tag{3.11}$$

The diagonal elements of the covariance matrix are each parameter's estimation variance, and the off-diagonal elements measure correlations between the estimates. The correspondence with variances allows for the notion of uncertainty, which is often defined as the square root of the variance. Throughout this thesis, where it is unambiguous, I will use the term variance interchangeably with MSE to describe the diagonal elements of the MSE matrix.

Biased Coin Toss MSE

The variance of the MLE derived above is

$$\mathrm{Var}(\hat{\theta}_{\mathrm{MLE}}) = \sum_{N_{\mathrm{H}}=0}^{N} \binom{N}{N_{\mathrm{H}}} \theta^{N_{\mathrm{H}}} (1-\theta)^{N-N_{\mathrm{H}}} \left(\frac{N_{\mathrm{H}}}{N} - \theta \right)^2 = \frac{\theta(1-\theta)}{N}. \tag{3.12}$$

Intuitively, this makes sense: when the true value $\theta \to 0$ (or $\theta \to 1$) the coin will almost always land on heads (tails) and the estimation variance will be small.

3.1.5 Fisher Information and the Classical Cramér–Rao Bound

An estimator-independent measure of the amount of *information* a probability distribution intrinsically carries about a parameter was introduced by Fisher in 1922 [51]. The quantity is now commonly known as the classical Fisher information and is a useful metric for the possible efficiency of parameter estimation. Let $p(k|\boldsymbol{\theta})$ denote the probability that the outcome of a measurement is k, which depends on the value of $\boldsymbol{\theta}$. Then the classical Fisher information matrix (CFIM) has elements

$$F_{ij} = \mathbb{E}_k \left[\frac{\partial \log p(k|\boldsymbol{\theta})}{\partial \theta_i} \frac{\partial \log p(k|\boldsymbol{\theta})}{\partial \theta_j} \right] = \sum_k \frac{1}{p(k|\boldsymbol{\theta})} \frac{\partial p(k|\boldsymbol{\theta})}{\partial \theta_i} \frac{\partial p(k|\boldsymbol{\theta})}{\partial \theta_j}. \tag{3.13}$$

If the spectrum of measurement outcomes is continuous, the sum is replaced by an integral.

The Fisher information places a limit on the MSE matrix of any unbiased estimator [52–54]. The Cramér–Rao inequality is

$$V(\hat{\boldsymbol{\theta}}) \succeq \frac{1}{N} F^{-1}, \tag{3.14}$$

where N is the number of samples, and for two equal-sized square matrices A and B, the matrix inequality $A \succeq B$ means that $A - B$ is positive semidefinite, i.e., has non-negative eigenvalues. The scaling with N is a consequence of the CFIM being additive for independent measurement repetitions. In the single-parameter case, this is equivalent to $\mathrm{Var}(\hat{\theta}) \geq 1/(NF)$, so the greater the Fisher information, the smaller the estimation variance may be. In the multiparameter case, it is common to transform

the Cramér–Rao inequality into a scalar lower bound on the sum of the variances by taking traces

$$\mathsf{Tr}[V(\hat{\boldsymbol{\theta}})] \geq \frac{1}{N}\mathsf{Tr}[F^{-1}], \tag{3.15}$$

where Tr denotes the trace of the matrix associated to the parameter vector space, to distinguish from Tr for quantum operators. This lower bound, hereafter called the classical Cramér–Rao bound (CCRB), can be generalised to place a lower bound on a weighted sum of estimation variances, allowing some variances to be preferentially minimised by including a positive-definite matrix W into the traces:

$$\mathsf{Tr}[WV(\hat{\boldsymbol{\theta}})] \geq \mathsf{Tr}[WF^{-1}]. \tag{3.16}$$

This gives a scalar quantity to optimise, rather than the covariance matrix itself. Here, I also drop the explicit dependence on the number of samples N; this amounts to considering the MSE per sample, and in the rest of this thesis, the dependence of the MSE on the number of samples will always be the same. The weighted MSE is equivalent to the MSE of the estimation of a different set of parameters $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \ldots, \varphi_m)^\top$ which are related to the original parameters by

$$\varphi_i = \sum_j H_{ij}\theta_j, \tag{3.17}$$

where $H^\top H = W$ [55]. This framework includes, for example, estimating the sum $\sum_i \theta_i$, or indeed any function $f(\boldsymbol{\theta})$ via its first-order Taylor expansion.

In addition to the MLE being asymptotically unbiased, Fisher demonstrated that CCRB is attainable in the asymptotic limit by the MLE [49]. As such, the classical Fisher information indicates the optimal MSE.

Biased Coin Toss CCRB

In this example, the Fisher information is

$$F = \sum_{k=\mathrm{H,T}} \frac{1}{p(k|\theta)} \left(\frac{\partial p(k|\theta)}{\partial \theta}\right)^2 = \frac{1}{\theta}(1)^2 + \frac{1}{1-\theta}(-1)^2 = \frac{1}{\theta(1-\theta)}. \tag{3.18}$$

Thus, the CCRB provides the lower bound on the variance

$$\mathrm{Var}(\hat{\theta}) \geq \frac{1}{NF} = \frac{\theta(1-\theta)}{N}, \tag{3.19}$$

which is equal to the above-derived value for the MLE.

3.2 Quantum Parameter Estimation

Quantum parameter estimation refers to the estimation of classical parameters (real values) encoded on quantum states. The above CCRB is a statement about the information which can be extracted from a probability distribution via random sampling. On the other hand, in quantum parameter estimation this "probability distribution" is concealed in the parameter-dependence of the quantum state and depends on the performed measurement.

The central task of quantum parameter estimation remains maximising the amount of information that is learnt about the parameters, but the distinction is that the random sampling is performed via quantum measurement. The parameters of the state need not correspond to physical observables, and they can be inferred from any measurement whose outcomes depend on the parameters. An example is parameters quantifying the amount of entanglement between two systems [56]. Quantum generalisations of classical estimation theory exist and have been studied extensively. Comprehensive reviews of the field can be found in Refs. [9–12].

As I will further discuss in this section, optimising quantum estimation of multiple parameters is complicated by non-commuting operators and the uncertainty principle. Even if we can determine the optimal measurement to estimate each parameter, the individually optimal measurements may not be compatible and hence cannot be implemented jointly. As such, the optimal measurement scheme must simultaneously consider the mean squared error of all parameters.

Figure 3.2 depicts a typical quantum parameter estimation process. A quantum state, called the probe state, has the parameters encoded onto it by the system of interest and is then measured. A large number of probe states are typically used to build

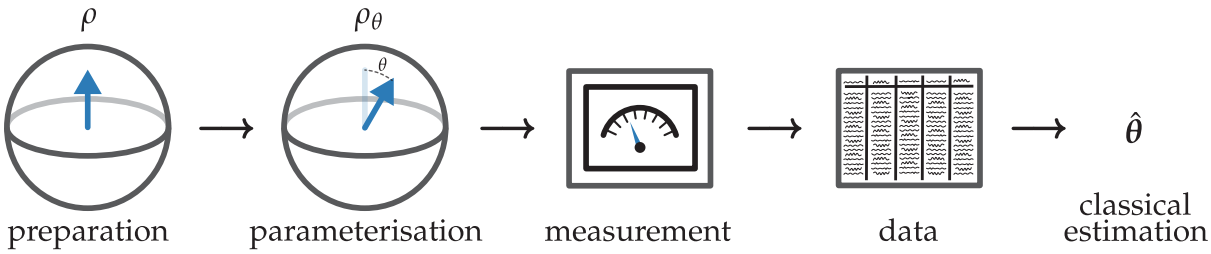

Figure 3.2: Typical quantum parameter estimation procedure. A quantum probe state (ρ) is prepared, and the parameter(s) are encoded onto it. Here, the parameter is the angle of rotation of the probe state in the Bloch sphere. The parameterised state is then measured. Repeating this process produces the statistics of the measurement outcomes, which the estimator converts to an estimate of the parameter(s).

statistics of the measurement. Finally, classical estimation theory is applied to convert the measurement outcomes (i.e., the sampled probability distribution) into an estimate of the parameters.

The process depicted in Figure 3.2 is for a scenario where the parameters of interest are associated with some system external to the quantum states being measured. The alternative, where the parameters are inherent to the quantum states, is also possible. For example, in quantum state tomography the parameters describe the quantum state itself.

Mathematically, for quantum parameter estimation, we consider a differentiable family of density operators $\rho_{\boldsymbol{\theta}}$ parameterised by $\boldsymbol{\theta}$. Owing to the local estimation setting, we can assume, without loss of generality, that the true value of $\boldsymbol{\theta}$ is zero and that we are optimising for estimating values of the parameters near this value. Here, differentiable means that the derivatives $\partial_i \rho := \frac{\partial \rho_{\boldsymbol{\theta}}}{\partial \theta_i}$ exist, so $\rho_{\boldsymbol{\theta}}$ varies smoothly with $\boldsymbol{\theta}$. The trace property of the density operator implies that $\partial_i \rho$ is traceless. In the following, unless otherwise stated, the derivatives of $\rho_{\boldsymbol{\theta}}$ will be evaluated at $\boldsymbol{\theta} = \mathbf{0}$.

3.2.1 Quantum Cramér–Rao Bound

Several measurement-independent quantum generalisations of the CFIM have been developed. The version that is most commonly associated with the term "quantum Fisher information" is based on the symmetric logarithmic derivative (SLD). The SLD operators (L_i) are one choice of generalisation of the logarithmic derivative in Eq. (3.13)

for non-commuting operators and were introduced by Helstrom [57, 58]. The necessity for a choice of logarithmic derivative can be seen by observing that the logarithmic derivative in Eq. (3.13) can be related to the derivative of the probability distribution by

$$\frac{\partial p(k|\boldsymbol{\theta})}{\partial \theta_i} = p(k|\boldsymbol{\theta}) \frac{\partial \log p(k|\boldsymbol{\theta})}{\partial \theta_i}. \tag{3.20}$$

In the quantum case, we replace $p(k|\boldsymbol{\theta})$ by the density operator $\rho_{\boldsymbol{\theta}}$, but there is ambiguity in the order of $\rho_{\boldsymbol{\theta}}$ and the logarithmic derivative if $\rho_{\boldsymbol{\theta}}$ and $\partial_i \rho_{\boldsymbol{\theta}}$ do not commute, which is generally the case. The self-adjoint SLD operators, L_i, are defined implicitly by

$$\frac{\partial \rho_{\boldsymbol{\theta}}}{\partial \theta_i} = \frac{1}{2} \left(\rho_{\boldsymbol{\theta}} L_i + L_i \rho_{\boldsymbol{\theta}} \right). \tag{3.21}$$

In analogy to the CFIM, the symmetric logarithmic derivative Fisher information matrix (SLDFIM), also known as the quantum Fisher information, is the expectation value of products of the SLD operators

$$\mathcal{F}_{ij} = \frac{1}{2} \operatorname{Tr}[\rho_{\boldsymbol{\theta}} \{L_i, L_j\}] = \operatorname{Re} \operatorname{Tr}[\rho_{\boldsymbol{\theta}} L_i L_j], \tag{3.22}$$

where Re denotes the real part and the equivalence of the two expressions can be shown using the self-adjoint property of the operators and the cyclic property of the trace. I will use the term SLDFIM to distinguish from other quantum generalisations of the CCRB (see Appendix A for a brief description of the right logarithmic derivative and its associated Fisher information). The utility of the SLDFIM is that it bounds the CFIM (F) of any measurement[1] [59]

$$F \preceq \mathcal{F}. \tag{3.23}$$

Due to the inequality Eq. (3.23) and the CCRB in Eq. (3.16), the SLDFIM defines a measurement-independent lower bound on the MSE, called the symmetric logarithmic derivative Cramér–Rao bound (SLDCRB)

$$\operatorname{Tr}[W V(\hat{\boldsymbol{\theta}})] \geq \mathcal{C}_{\text{SLD}} := \operatorname{Tr}[W \mathcal{F}^{-1}]. \tag{3.24}$$

[1] The CFIM of a POVM is calculated by Eq. (3.13) using the Born rule (Eq. (2.5)) for the probabilities.

For single-parameter estimation, both the CFIM and SLDFIM are simply numbers and the inequality is tight: the SLDFIM is the maximum of the CFIM over all measurements. In this case, the SLDCRB can be attained by a projective measurement on the eigenbasis of the SLD operator and post-processing with the maximum likelihood estimator [59]. In the next section, I will discuss the SLDCRB's attainability in the multiparameter case.

While the SLD operators cannot generally be determined directly from Eq. (3.21), they can be calculated using the eigenbasis of the quantum state's density operator. Explicitly, writing the density operator as $\rho = \sum_n \rho_n |\psi_n\rangle\langle\psi_n|$, the SLD operators are [59–61]

$$L_i = \sum_{n,m;\ \rho_n+\rho_m\neq 0} \frac{2\langle\psi_m|\,\partial_i\rho\,|\psi_n\rangle}{\rho_n+\rho_m}\,|\psi_m\rangle\langle\psi_n|\,, \tag{3.25}$$

where any terms in the summation where $\rho_n + \rho_m = 0$ are ignored.

3.2.2 The Complication of Multiple Parameters

Many physical problems require the determination of multiple parameters of a system, which motivated the development of quantum multiparameter estimation theory. Some examples include the estimation of vector fields in 3D magnetometry [62], the estimation of the positions of multiple point sources beyond the Rayleigh limit [7], and estimating optical phase shift and loss in interferometry [63].

The generalisation from single-parameter quantum estimation theory to multiple parameters is not straightforward. This is because of the uncertainty principle. Even if we know the optimal measurement and estimation scheme to estimate each system parameter individually, there is no guarantee that these individually optimal measurements will be compatible. In general, they will not be compatible, so they cannot be jointly implemented. As such, quantum multiparameter estimation has an added layer of complexity in that the optimal measurement scheme must simultaneously consider all parameters.

As alluded to in the previous section, the SLDCRB is in general not attainable. Nevertheless, the SLDCRB is attainable in special cases. Namely, if the SLD operators for all parameters commute ($[L_i, L_j] = 0$ for all i, j), then they share a common eigenbasis, and a projective measurement on such a basis will be optimal for estimating all

parameters.

Despite its general unattainability, the SLDCRB has been widely used for multiparameter estimation because it is easy to calculate, as the SLD operators can be computed directly (Eq. (3.25)). However, other tighter estimation bounds exist, which I discuss in the following section. The drawback of these other lower bounds is that they generally do not have a closed-form expression and require non-trivial minimisation.

3.2.3 Tighter Bounds for Multiparameter Estimation

In contrast to the SLDCRB, the remaining estimation lower bounds studied in this thesis all require solving a complex minimisation problem. Aside from the *most informative Cramér–Rao bound* (see below), I focus on alternative lower bounds which can be recast as a semidefinite program. This means that they can be computed efficiently using standard software [64].

In practice, in this thesis, except where general exact expressions exist (in the case of two-parameter estimation of qubit systems), I use a version of publicly available codes [65] written in `MATLAB` to compute the estimation bounds, which use the `YALMIP` toolbox [66] and the `MOSEK` solver [67]. Sometimes, perhaps surprisingly, an exact expression for the lower bound can be deduced from the numerical SDP solution.

Most Informative Cramér–Rao Bound

Nagaoka [68] introduced the most informative Cramér–Rao bound as the minimisation problem

$$\mathcal{C}_{\mathrm{MI}} = \min \left\{ \mathsf{Tr}[WF(\Pi)^{-1}] \mid \Pi \text{ is a POVM} \right\}, \tag{3.26}$$

where $F(\Pi)$ is the CFIM of the POVM Π. This bound is the "most informative" because by definition it is attainable, and thus it is greater than (or equal to) all other Cramér–Rao type lower bounds. However, the minimisation problem is intractable even for small quantum systems, so this bound is impractical for direct application to an estimation problem. In practice, we can attempt to solve the minimisation problem

numerically[2], but there is no guarantee that doing so will result in the true minimum. Nevertheless, the most informative bound is a useful benchmark for the optimal lower bound. Nagaoka demonstrated that for single-parameter estimation, $\mathcal{C}_{\mathrm{MI}} = \mathcal{C}_{\mathrm{SLD}}$, in accordance with the fact that the SLDCRB can be attained for single-parameter estimation.

Holevo Cramér–Rao Bound

A remarkable improvement to the SLDCRB was introduced by Holevo [14, 24]—now called the Holevo Cramér–Rao bound (HCRB)—and is widely used because it places a lower bound on the MSE for any measurement, separable or collective. As such, if a measurement is found to achieve the HCRB, then no other measurement can outperform it. For this reason, it sets the ultimate precision limit for multiparameter estimation.

The HCRB can be expressed as a minimisation over Hermitian matrices $X = (X_1, \ldots, X_n)^\top$ [68]

$$\mathsf{Tr}[WV(\hat{\boldsymbol{\theta}})] \geq \mathcal{C}_{\mathrm{H}} \coloneqq \min_X \left\{ \mathsf{Tr}[W \operatorname{Re} Z_{\boldsymbol{\theta}}[X]] + \mathsf{TrAbs}[W \operatorname{Im} Z_{\boldsymbol{\theta}}[X]] \right\}, \tag{3.27}$$

where the matrix $Z_{\boldsymbol{\theta}}[X]$ has elements $Z_{\boldsymbol{\theta}}[X]_{ik} = \mathrm{Tr}[\rho_{\boldsymbol{\theta}} X_i X_j]$, for a matrix A, $\mathsf{TrAbs}[A]$ is the sum of the absolute values of the eigenvalues of A, and Im denotes the imaginary part. The X matrices can be related to an estimator and a POVM via the relationship

$$X_i = \sum_j \left(\hat{\theta}_{ij} - \theta_i\right) \Pi_j, \tag{3.28}$$

and satisfy locally unbiased conditions

$$\mathrm{Tr}[\rho_{\boldsymbol{\theta}} X_i] = \theta_i, \quad \frac{\partial}{\partial \theta_j} \mathrm{Tr}[\rho_{\boldsymbol{\theta}} X_i] = \delta_{ij}. \tag{3.29}$$

Attainability aside, the X matrices make the minimisation problem tractable, in con-

[2]For example, a POVM with a fixed (chosen) number m of outcomes can be written in terms of m complex vectors $|\psi_i\rangle$. Then, the objective $\mathsf{Tr}[WF(\Pi)^{-1}]$ can be written in terms of $|\psi_i\rangle$ and minimised using standard numerical packages. While such an approach is not necessarily sufficient to determine $\mathcal{C}_{\mathrm{MI}}$, the solution can be compared to other lower bounds and used to investigate attainability.

trast to a problem of minimising over a POVM with an arbitrary number of elements, as is the case for the most informative bound.

The numerical tractability of the HCRB was proven by Albarelli *et al.* [69] who demonstrated that the minimisation problem can be recast as a semidefinite program. Furthermore, Suzuki [70] demonstrated that for two-parameter estimation of qubits, the HCRB has an exact expression, which I provide in Appendix A.

Nagaoka Cramér–Rao Bound

For two-parameter estimation, Nagaoka introduced a new estimation lower bound which is more restrictive than the HCRB because it only applies to separable (single-copy) measurements [68]. Now known as the Nagaoka Cramér–Rao bound (NCRB), this lower bound is also the solution of a minimisation problem over the Hermitian matrices $X = (X_1, X_2)^\top$. The NCRB can be written

$$\mathsf{Tr}[WV(\hat{\boldsymbol{\theta}})] \geq \mathcal{C}_{\mathrm{N}} := \min_{X} \left\{ \mathsf{Tr}[WZ_{\boldsymbol{\theta}}(X)] + \sqrt{\mathsf{Det}[W]}\, \mathsf{TrAbs}[\rho_{\boldsymbol{\theta}}[X_1, X_2]] \right\}, \tag{3.30}$$

where the elements of $Z_{\boldsymbol{\theta}}[X]$ are the same as in the HCRB. That the NCRB applies only for separable measurements follows from the derivation (not shown here) of the lower bound, where the measurement (e.g., a POVM) is explicitly included. Nagaoka demonstrated that the NCRB is a more informative (tighter) bound than the HCRB (i.e., that $\mathcal{C}_{\mathrm{N}} \geq \mathcal{C}_{\mathrm{H}}$), that the NCRB is attainable (tight) for estimation of qubits (i.e., that for $\dim \mathcal{H} = 2$, $\mathcal{C}_{\mathrm{MI}} = \mathcal{C}_{\mathrm{N}}$), and conjectured that it is also tight for higher dimensions [71]. Nagaoka also demonstrated that the NCRB for qubit estimation has an analytic form [68, 71], given by

$$\mathcal{C}_{\mathrm{N}} = \mathsf{Tr}[W\mathcal{F}^{-1}] + 2\sqrt{\mathsf{Det}[W\mathcal{F}^{-1}]}, \tag{3.31}$$

where $\mathcal{F}$ is the SLDFIM. As this is an expression for a tight estimation bound, it is useful for comparison to other estimation bounds, as I will demonstrate in Chapters 5 and 6.

Nagaoka–Hayashi Cramér–Rao Bound

Recently, Conlon *et al.* introduced a new lower bound which is an extension of the NCRB to estimating more than two parameters [72]. Called the Nagaoka–Hayashi Cramér–Rao bound (NHCRB), this new bound can be written as a minimisation over Hermitian matrices $X = (X_1, \ldots, X_n)^\top$, as in the previous two bounds, and additionally a matrix $\mathbb{L}$:

$$\mathrm{Tr}[WV(\hat{\boldsymbol{\theta}})] \geq \mathcal{C}_{\mathrm{NH}} \coloneqq \min_{\mathbb{L},X} \left\{ \mathbb{T}\mathrm{r}[\mathbb{S}_{\boldsymbol{\theta}}\mathbb{L}] \mid \mathbb{L}_{jk} = \mathbb{L}_{kj} \text{ Hermitian}, \mathbb{L} \geq XX^\top \right\}. \tag{3.32}$$

Here, $\mathbb{S}_{\boldsymbol{\theta}} = W \otimes \rho_{\boldsymbol{\theta}}$ is an extension to the density operator, and $\mathbb{L}$ is an $n \times n$ matrix of Hermitian operators $\mathbb{L}_{jk}$. The symbol $\mathbb{T}\mathrm{r}$ is used to denote the trace over both the classical vector space and the quantum Hilbert space.

Conlon *et al.* further demonstrated that the NHCRB (and hence also the NCRB, which the NHCRB reduces to for two-parameter estimation) can be recast as a semidefinite program, and can thus be efficiently computed [72].

The HCRB can be written in a form similar to Eq. (3.32) (i.e., with a matrix of Hermitian operators) except with less stringent constraints [24], so the NHCRB is a tighter estimation bound than the HCRB: $\mathcal{C}_{\mathrm{NH}} \geq \mathcal{C}_{\mathrm{H}}$. This reflects that a collective measurement may offer an improved MSE—the entanglement allows for enhanced measurement efficiency, i.e., more information is extracted using the same number of probe states.

Relationship Between the Cramér–Rao Type Bounds and Their Attainability

I remind the reader that an important aspect of estimation lower bounds is attainability, and this is the measure with which we can compare different bounds. Tighter lower bounds (i.e., closer to being attainable) are better because they provide more information about the attainable measurement precision. As such, a central theme of the investigations presented in this thesis is the tightness of estimation limits.

In Figure 3.3, I provide a summary of the restrictions included in the Cramér–Rao bounds. Here, I outline results regarding the bounds that are notable and relevant to

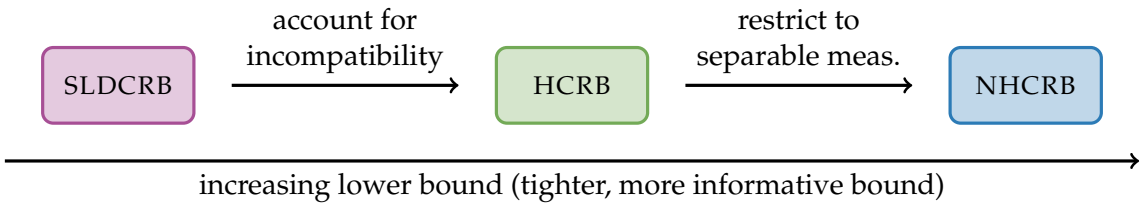

Figure 3.3: Summary of hierarchy and restrictions on calculable Cramér–Rao bounds. From left to right, the lower bound increases as additional restrictions are imposed, meaning that the bounds are tighter and provide more information about the ultimate precision limit. The HCRB includes consideration for measurement compatibility and the NHCRB adds the restriction of separable measurements. The NCRB is a special case of the NHCRB for two-parameter estimation.

this thesis. The hierarchy of the Cramér–Rao type bounds is

$$\mathcal{C}_{\mathrm{MI}} \geq \mathcal{C}_{\mathrm{NH}} \geq \mathcal{C}_{\mathrm{H}} \geq \mathcal{C}_{\mathrm{SLD}}. \tag{3.33}$$

The first inequality is saturated for two-parameter estimation in qubits [68, 71], but was recently shown, by counterexample, to not be saturated in general [73]. As such, the NHCRB is not always an attainable bound. The second inequality is saturated for estimation of pure states [74], and the third is saturated when the weak commutative condition holds: $\mathrm{Tr}[\rho_{\boldsymbol{\theta}}[L_i, L_j]] = 0$, for all i, j [75].

The HCRB is often used because it gives the ultimate limit on the estimation precision, and this limit is asymptotically attainable, meaning that it can be achieved by a collective measurement on N copies as N tends to infinity [76–78]. However, this asymptotic attainability was recently demonstrated to lack practical utility because if the HCRB cannot be saturated with a single-copy measurement, then it cannot be saturated by a collective measurement on any finite number of copies [79].

3.2.4 Lu–Wang Information Regret Trade-off Relation

A notable absence in the aforementioned precision limits for quantum multiparameter estimation is an explicit reference to Heisenberg's uncertainty relation or any of the other uncertainty relations. The uncertainty principle is baked into the definitions of the Holevo, Nagaoka, and Nagaoka–Hayashi bounds, but in a way that is obscured into their minimisation constraints. It begs the question of whether one can arrive at

a limit on the MSE by directly invoking an uncertainty relation. Indeed, this topic has recently been explored by Lu and Wang [16].

Lu and Wang bridged the gap between uncertainty relations and parameter estimation theory by incorporating classical and quantum Fisher information into Ozawa's joint-measurement uncertainty relation (Eq. (2.9)) [16]. They define the "information regret" matrix R as the difference between the classical and quantum Fisher information: $R = \mathcal{F} - F$, where $\mathcal{F}$ is the SLDFIM and $F = F(\Pi)$ is the CFIM of the POVM Π. The information regret provides a metric of the optimality of a measurement because vanishing regret corresponds to saturation of the SLDCRB. Of course, due to the incompatibility of operators, the regret typically does not vanish, so Lu and Wang use Ozawa's universally valid uncertainty relation to investigate the trade-off between components of the regret. Their information-regret trade-off relation (IRTR) is

$$\Delta_j^2 + \Delta_k^2 + 2\sqrt{1 - \tilde{c}_{jk}^2}\,\Delta_j\Delta_k \geq \tilde{c}_{jk}, \tag{3.34}$$

where $\Delta_j = \sqrt{R_{jj}/\mathcal{F}_{jj}}$ is the normalised-square-root regret and

$$\tilde{c}_{jk} = \frac{\mathrm{Tr}\left|\sqrt{\rho_\theta}[L_j, L_k]\sqrt{\rho_\theta}\right|}{2\sqrt{\mathcal{F}_{jj}\mathcal{F}_{kk}}} \tag{3.35}$$

is a normalised version of Ozawa's D_{AB} for the SLD operators (see Eq. (2.10)). Conforming to the literature (e.g., Refs [80, 81]), I will refer to $\tilde{c}_{jk}$ as the incompatibility coefficient and will often drop the subscript when it is unambiguous. Note that $0 \leq \tilde{c} \leq 1$ [16, 48]. Owing to Branciard's result, the IRTR is tight for pure states [16], i.e., that there exists a measurement that achieves equality in Eq. (3.34). For estimating n parameters, there are $\binom{n}{2}$ IRTRs which each consider only two-parameter estimation.

3.2.5 Lu–Wang Uncertainty Relation

Lu and Wang presented a second form of their parameter estimation uncertainty relation regarding the MSE [16]. This makes comparing their result to the Cramér–Rao type bounds possible, even though it is not necessarily straightforward. To derive this second relation, they used the single-parameter CCRB for the diagonal elements of the

MSE V

$$V_{jj} \geq (F^{-1})_{jj} \geq \frac{1}{F_{jj}} \implies F_{jj} \geq \frac{1}{V_{jj}} \quad \text{so that} \quad \Delta_j^2 = 1 - \frac{F_{jj}}{\mathcal{F}_{jj}} \leq 1 - \frac{1}{V_{jj}\mathcal{F}_{jj}}, \tag{3.36}$$

where the first inequality is the CCRB for V_{jj} and the second inequality is a property of positive semidefinite matrices, which is saturated if and only if F is diagonal. Inserting the last expression in Eq. (3.36) into Eq. (3.34), Lu and Wang derived their estimation uncertainty relation

$$\gamma_j + \gamma_k - 2\sqrt{1 - \tilde{c}_{jk}^2}\sqrt{(1 - \gamma_j)\,(1 - \gamma_k)} \leq 2 - \tilde{c}_{jk}^2, \tag{3.37}$$

where $\gamma_j = 1/\left(V_{jj}F_{jj}\right)$. I will refer to this as the Lu–Wang uncertainty relation (LWUR). Because the inequalities in Eq. (3.36) may not be able to be simultaneously saturated for both values of j, the LWUR may not be as tight as the IRTR. The topic of tightness of the LWUR and IRTR will be explored throughout the rest of this thesis, particularly in Chapter 8.

The LWUR and the Cramér–Rao bounds offer different insights into the limitations of the MSE. As an uncertainty relation, the LWUR explicitly defines the trade-off between the two estimation variances but provides no information about the minimum allowed sum of the variances. On the other hand, the Cramér–Rao bounds define limits on the sum of the variances but provide no immediate insight into the distribution of error amongst the different parameters or their trade-off. These differences mean that it is not obvious how the estimation limits relate to each other and how they can be compared. In Chapter 4, I develop methods for the comparison of the estimation limits, and I subsequently investigate the relationship between the limits in the rest of this thesis.

Chapter 4

Methods for Comparing Estimation Limits

As we have seen in Chapter 3, the Cramér–Rao bounds and the Lu–Wang uncertainty relation place limits on estimation variances in different ways; a lower bound on the sum for the former and a non-trivial inequality for the latter. The difference between the precision limits makes them difficult to compare directly. It is instructive to visualise which estimation variances are not forbidden[1] by a certain lower bound by drawing a plot with one axis for each variance. In this way, the stark difference between the limits is evident, as illustrated by the region plots in Figure 4.1.

In this chapter, I develop methods for comparing Cramér–Rao bounds and the Lu–Wang uncertainty relation. First, I define a new estimation lower bound based on the LWUR and discuss how to calculate it. Next, I discuss deriving uncertainty relations from Cramér–Rao bounds. The methods outlined in this chapter will be applied to estimation problems throughout the rest of this thesis.

[1] I use the term "not forbidden" as opposed to "allowed" to avoid confusion with achievable variances in the context of the tightness of bounds.

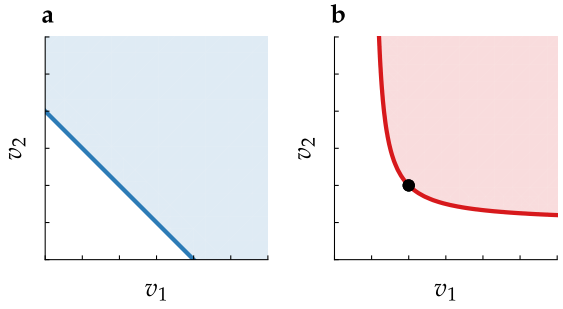

Figure 4.1: Visualising the difference between Cramér–Rao bounds and the LWUR. The shaded regions denote the variances that are not forbidden by the estimation limit. **a**, a Cramér–Rao bound (unweighted) places a limit on $v_1 + v_2$. **b**, the LWUR is a non-trivial inequality and the boundary of its region is curved. The black point denotes the point where $v_1 + v_2$ is minimal and is relevant to the LWB.

4.1 Sum-type Bounds from Uncertainty Relations

In this section, I introduce a new lower bound on the sum of estimation variances based on the LWUR. I then discuss convex optimisation and analytic methods for minimisation problems. Finally, I demonstrate the tractability of the optimisation problem of the new lower bound, and explain how it can be solved numerically.

4.1.1 Lu–Wang Estimation Bound

For two-parameter estimation, I define the Lu–Wang Estimation bound (LWB), $\mathcal{C}_{\mathrm{LW}}$, as the minimum sum of estimation variances allowed by the LWUR. This bound can be formulated as the solution of a minimisation problem

$$\mathcal{C}_{\mathrm{LW}} \coloneqq \min_{v_1, v_2} \left\{ v_1 + v_2 \;\middle|\; \gamma_1 + \gamma_2 - 2\sqrt{1-\tilde{c}^2}\sqrt{(1-\gamma_1)\,(1-\gamma_2)} \leq 2 - \tilde{c}^2 \right\} \tag{4.1}$$

where $\gamma_i = 1/(v_i \mathcal{F}_{ii})$, which according to the single-parameter SLDCRB must be greater than or equal to 1. By construction, $\mathcal{C}_{\mathrm{LW}}$ is a lower bound for the sum of estimation variances and can be compared directly to Cramér–Rao-type bounds. Important to the tractability of this minimisation problem is that the problem is convex, which I will prove later in this section. Convexity ensures that a local minimum of $v_1 + v_2$ is a global minimum, simplifying the bound's computation.

4.1.2 Convex Optimisation Toolkit

Here, I discuss convex optimisation, the optimisation of convex functions over convex sets. I begin with a few definitions and then discuss analytic methods for the solution of optimisation problems. I follow the notation used in the textbook *Convex Optimisation* by Boyd and Vandenberghe [64].

Convex Optimisation Problems

Consider the generic constrained optimisation problem:

$$\min_{\mathbf{x}} \{ f_0(\mathbf{x}) \mid f_i(\mathbf{x}) \leq 0,\ i = 1, \ldots, m;\ h_i(\mathbf{x}) = 0,\ i = 1, \ldots, p \} \tag{4.2}$$

where $\mathbf{x} \in \mathbb{R}^n$. f_0 is called the objective function and f_i (h_i) are called the inequality (equality) constraint functions. The set $\mathcal{D} \subset \mathbb{R}^n$, called the feasible region, denotes the set of $\mathbf{x} \in \mathbb{R}^n$ which satisfy the constraints. A function $g : \mathbb{R}^n \to \mathbb{R}$ is called convex if $g(\alpha\mathbf{x} + \beta\mathbf{y}) \leq \alpha g(\mathbf{x}) + \beta g(\mathbf{y})$, for all $\mathbf{x}, \mathbf{y}$ in its domain and all $\alpha, \beta \in \mathbb{R}$ [64]. Graphically, this is equivalent to the condition that the line connecting any two points on the graph of g does not lie below the graph at any point.

An optimisation problem is called convex if the objective and all inequality constraints (f_0, f_i) are convex functions and all equality constraints (h_i) are affine functions[2]. These conditions ensure that the feasible region of the optimisation problem is convex—the line connecting any two points is contained wholly within the set.

A very useful property of convex optimisation problems is that every local minimum is also a global minimum [64], which simplifies solving the problem. Additionally, there are several methods for solving convex problems, including analytic techniques, which I will outline below. In the investigations presented in this thesis, we desire analytic solutions because they allow us to observe the dependence of the solution on different variables of the system, and compare this across different estimation bounds.

In the following, $p^\star$ denotes the optimal value of $f_0(\mathbf{x})$, and $\mathbf{x}^\star$ denotes the minimiser, i.e., $p^\star = f_0(\mathbf{x}^\star)$.

The Lagrangian for Optimisation Problems

Directly determining the solution of the original, or *primal*, problem Eq. (4.2) analytically is often difficult or even intractable. Several approaches to a solution involve a function called the *Lagrangian*, which incorporates the constraint functions:

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}), \tag{4.3}$$

where λ_i, ν_i are called the Lagrange multipliers [64]. As we will see, this allows applying zero-gradient optimality tests, just as in the unconstrained case. The technique of

[2] An affine function h is one that can be written as $h(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} + b$ for some $\mathbf{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}$, i.e., h is a combination of a linear transformation and a translation.

Lagrange multipliers is not uncommon in physics, but it is often applied to problems with equality constraints only.

The Dual Problem and Slater's Theorem

The basis of the Lagrange multiplier approach is that the Lagrangian can be used to place a lower bound on $p^\star$, the solution of the primal problem. In some instances, this lower bound can be achieved.

Given a Lagrangian $L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu})$, the *Lagrange dual function* is defined

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}). \tag{4.4}$$

It is easy to see that for $\lambda_i \geq 0$, the Lagrange dual sets a lower bound on the primal solution: $g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq p^\star$, because $f_i(\mathbf{x}) \leq 0$ and $h_i(\mathbf{x}) = 0$ for points in the feasible region. For this reason, the set of points $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ with $\lambda_i \geq 0$ is called the dual feasible region. To obtain the greatest lower bound, the Lagrange dual can be maximised in the dual optimisation problem: $d^\star = \max_{\lambda,\nu} \{g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \mid \lambda_i \geq 0\}$, where there are no constraints on $\boldsymbol{\nu}$, and $d^\star$ denotes the optimal value of the dual problem. For problems with simple objective and constraints, the Lagrangian dual function can be computed exactly and then maximised.

The *weak duality theorem* states that $d^\star \leq p^\star$, which follows from the lower-bound property of the Lagrange dual function. Of course, in searching for exact solutions to the original problem, the dual solution is only helpful if $d^\star = p^\star$, called *strong duality*. Fortunately, it follows from *Slater's theorem* that a convex optimisation problem has strong duality if the interior of its feasible region is non-empty[3]. The Lagrange dual function will be used to calculate the LWB for a class of estimation problems in Sec. 5.3.

The Karush–Kuhn–Tucker (KKT) conditions

One alternative to the dual problem approach to solving an optimisation problem is via the Karush–Kuhn–Tucker (KKT) conditions. Assuming that strong duality holds and the additional assumption that the objective and constraint functions are differen-

[3] $\mathcal{D}$ is non-empty if there exists $\tilde{\mathbf{x}}$ such that $f_i(\tilde{\mathbf{x}}) < 0$ (strict inequality) and $h_i(\tilde{\mathbf{x}}) = 0$ for $i > 0$.

tiable, the KKT conditions provide an optimality condition for solutions of the optimisation problem [64]. The conditions are the analog of finding points with zero gradient in the unconstrained optimisation case. Let $(\boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star)$ denote the optimisers of the dual problem. The KKT conditions are a set of necessary conditions for $(\mathbf{x}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star)$ to be optimisers. The conditions are

$$\begin{aligned}
&\text{(I; II)} && f_i(\mathbf{x}^\star) \leq 0, \quad i = 1, \ldots, m; \quad h_i(\mathbf{x}^\star) \leq 0, \quad i = 1, \ldots, p \\
&\text{(III; IV)} && \lambda_i^\star \geq 0, \quad i = 1, \ldots, m; \quad \lambda_i^\star f_i(\mathbf{x}^\star) = 0, \quad i = 1, \ldots, m \\
&\text{(V)} && \boldsymbol{\nabla} f_0(\mathbf{x}^\star) + \sum_{i=1}^{m} \lambda_i^\star \boldsymbol{\nabla} f_i(\mathbf{x}^\star) + \sum_{i=1}^{p} \nu_i \boldsymbol{\nabla} h_i(\mathbf{x}^\star) = 0
\end{aligned} \tag{4.5}$$

where conditions I–III ensure feasibility of the solution. Condition IV is called complementary slackness and reflects that if the optimum occurs where an inequality is strict ($f_i(\mathbf{x}^\star) < 0$ for some i), then that constraint should be ignored in the gradient–optimality consideration (V), so $\lambda_i = 0$. Finally, condition V enforces that $x^\star$ is a stationary point of $L(\mathbf{x}, \boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star)$, which is a necessary condition that follows from strong duality [64].

The above KKT conditions do not require the optimisation problem to be convex. However, if the problem is convex, then in addition to being necessary, the KKT conditions are sufficient for the optimum [64]. Solving the optimisation problem is then equivalent to finding $(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu})$ satisfying the KKT conditions. The KKT conditions will be used to calculate the LWB in an example in Sec. 6.2.2.

4.1.3 Convexity of Lu–Wang Feasible Region

As I have discussed above, convex optimisation problems are significantly simpler than generic optimisation problems. It is thus fortunate that the LWB minimisation problem is convex, which I will now demonstrate. The LWB objective function $(v_1 + v_2)$ is convex because it is a positive linear combination of the trivially convex functions v_1 and v_2. It remains to prove that the set defined by the nonlinear constraint in Eq. (4.1) is convex.

By writing the SLDFIM as $\mathcal{F} = \begin{pmatrix} \alpha & \delta \\ \delta & \beta \end{pmatrix}$, where $\alpha, \beta > 0$, the constraint function in

Eq. (4.1) can be written as

$$\underbrace{\frac{1}{\alpha v_1}+\frac{1}{\beta v_2}}_{g_1} -2\sqrt{1-\tilde{c}^2}\underbrace{\sqrt{\left(1-\frac{1}{\alpha v_1}\right)\left(1-\frac{1}{\beta v_2}\right)}}_{-g_2}+\tilde{c}^2-2\leq 0. \tag{4.6}$$

Here, I have separated the left-hand side into a positive sum of two functions, denoted g_1 and g_2, and a constant. Using the linear combination property of convex functions, to show that Eq. (4.6) is convex it suffices to show that g_1 and g_2 are each convex. Because the functions are of two variables, convexity is equivalent to its Hessian, the matrix of second partial derivatives, being positive semidefinite, provided the function's domain is convex [64], which is the case here.

For g_1, the Hessian is

$$H_{g_1}=\begin{pmatrix}\frac{\partial^2 g_1}{\partial v_1^2} & \frac{\partial^2 g_1}{\partial v_1 \partial v_2}\\ \frac{\partial^2 g_1}{\partial v_2 \partial v_1} & \frac{\partial^2 g_1}{\partial v_2^2}\end{pmatrix}=\begin{pmatrix}\frac{2}{\alpha v_1^3} & 0\\ 0 & \frac{2}{\beta v_2^3}.\end{pmatrix}, \tag{4.7}$$

which has non-negative eigenvalues, so it is positive semidefinite, and g_1 is convex. Similarly, the Hessian of g_2 is

$$H_{g_2}=\begin{pmatrix}\frac{(4\alpha v_1-3)\sqrt{\left(1-\frac{1}{\alpha v_1}\right)\left(1-\frac{1}{\beta v_2}\right)}}{4v_1^2(\alpha v_1-1)^2} & \frac{-1}{4v_1^2v_2^2\alpha\beta\sqrt{\left(1-\frac{1}{\alpha v_1}\right)\left(1-\frac{1}{\beta v_2}\right)}}\\ \frac{-1}{4v_1^2v_2^2\alpha\beta\sqrt{\left(1-\frac{1}{\alpha v_1}\right)\left(1-\frac{1}{\beta v_2}\right)}} & \frac{(4\beta v_2-3)\sqrt{\left(1-\frac{1}{\alpha v_1}\right)\left(1-\frac{1}{\beta v_2}\right)}}{4v_2^2(\beta v_2-1)^2}\end{pmatrix}. \tag{4.8}$$

A matrix M is positive semidefinite if $z^\top Mz\geq 0$ for all non-zero real vectors $z=(z_1,z_2)^\top$. In this case, $z^\top H_{g_2}z$ can be expressed as

$$z^\top H_{g_2}z=\frac{r_1^2+r_2^2+\left(\frac{r_1}{2\sqrt{v_1\alpha-1}}-\frac{r_2}{2\sqrt{v_2\beta-1}}\right)^2}{\sqrt{v_1v_2\alpha\beta(v_1\alpha-1)(v_2\beta-1)}};\quad r_1=\frac{z_1}{v_1}\sqrt{v_2\beta-1},\quad r_2=\frac{z_2}{v_2}\sqrt{v_1\alpha-1} \tag{4.9}$$

The single-parameter SLDCRB conditions are that $v_1\alpha\geq 1$ and $v_2\beta\geq 1$, so the square-root terms in Eq. (4.9) are real, and hence $z^\top H_{g_2}z\geq 0$ and the Hessian of g_2 is positive

semidefinite. It follows that the feasible region of the LWB minimisation problem is convex and that the problem itself is well-posed.

4.1.4 Numerical Computation of the Lu–Wang Estimation Bound

As demonstrated in the previous subsection, the minimisation problem for the LWB is convex. The aforementioned analytic techniques for optimisation can thus be used to compute the bound. However, it is also convenient to be able to find a solution numerically, allowing the LWB to be quickly computed and compared to other estimation bounds.

I wrote code to numerically calculate the LWB for any estimation problem. I numerically calculate the LWB by utilising the constrained optimisation package `fmincon` in `MATLAB` to solve the optimisation problem (4.1). The LWUR constraint is input into `fmincon`, along with the single-parameter SLDCRB conditions which ensure that the LWUR inequality is real-valued.

4.2 Uncertainty Relations from Cramér–Rao Bounds

A Cramér–Rao bound provides a single value characterising the possible estimation performance via a lower bound on a weighted trace of the MSE. It thus does not specify the trade-off between the individual variances: how much must one variance increase to allow another to decrease? However, as we will see, this information can be extracted from a Cramér–Rao bound by considering different weightings of the variances. This allows for a holistic picture of the estimation precisions which are not forbidden by a Cramér–Rao bound, and provides a further avenue to compare the NCRB to the LWUR.

4.2.1 Visualising Weighted Cramér–Rao bounds

The most general Cramér–Rao bound places a lower bound on a weighted sum of the MSE elements

$$\mathsf{Tr}[WV_{\boldsymbol{\theta}}] \geq \mathcal{C}, \tag{4.10}$$

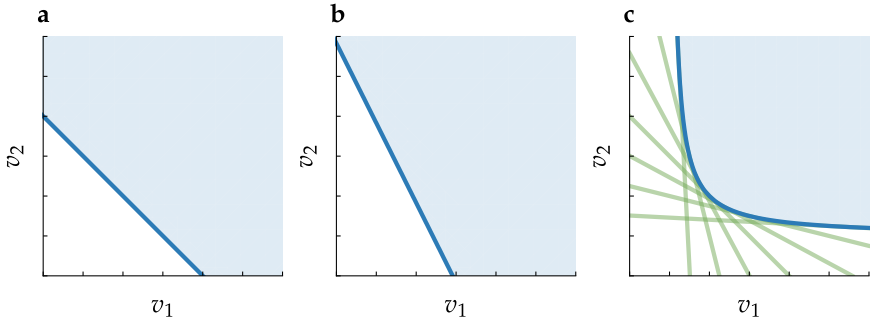

Figure 4.2: Sketch of the two-parameter variance region plot for an unweighted Cramér–Rao-type bound. **a**, Unweighted bound ($W = \mathbb{1}$). **b**, Weighted bound with $W = \left(\begin{smallmatrix} w & 0 \\ 0 & 2-w \end{smallmatrix}\right)$ and $w = 4/3$. **c**, The overall non-forbidden region (blue) is the intersection of the regions from each weighting. The boundary is the envelope of the lines from each weighting (some plotted in green for illustration).

where W is a positive weight matrix [55]. Notionally, placing a lower bound on a weighted sum of the estimation variances allows the variances to have non-uniform cost: some parameters' variances ought to be preferentially optimised.

Figure 4.2**a**,**b** demonstrates the region defined by a Cramér–Rao bound for two-parameter estimation in the unweighted and weighted cases. It is evident that the unweighted region contains points that are not possible (e.g., along one of the axes), but some weighted bound forbids such points. This observation suggests that a tighter region can be derived by combining the regions from all possible weightings, as illustrated in Figure 4.2**c**. Indeed, this method was performed by Kull *et al.* in Ref. [82] for some Cramér–Rao bounds, but not the NCRB. More recently, the same method was applied to the NCRB in a specific estimation problem [13]. In the following subsections, I outline this approach, which will be applied in Chapters 5 and 7.

4.2.2 Combined Variance Regions

The combined variance region is the set of points in variance space that are not forbidden by any weighting of the Cramér–Rao bound. The region defined by each weighting of the bound is convex, so the combined region, which is the intersection, will also be convex. It thus suffices to determine the boundary of this region, which also naturally defines a trade-off between the variances—an uncertainty relation. As illustrated

in Figure 4.2**c**, this boundary is the envelope formed by the lines of different gradients[4]. Note that the envelope's shape depends on the bound's dependence on the weight matrix. For example, if all the lines intersect at the same point, the envelope would have a sharp corner at this point rather than being curved.

If the weighted Cramér–Rao bound is known, then the uncertainty relation can, in principle, be derived. To see this, consider the two-parameter case, as was done by Kull *et al.* [82]. Let the weight matrix be $W = \begin{pmatrix} w & \\ & 2-w \end{pmatrix}$ and let $\mathcal{C}(w)$ be the bound as a function of the weighting. I use this weight matrix choice so that the standard unweighted case ($W = \mathbb{1}$) is present for some value of w. Other parameterisations could be used and would result in the same eventual relation; the ratio of the diagonal elements is important, but their scale only sets the scale for $\mathcal{C}(w)$—rescaling the weight matrix rescales the lower bound, but the restriction on the MSE in Eq. (4.10) is unchanged. The defining equation of points (pairs of variances) on each line is

$$wv_1 + (2-w)v_2 = \mathcal{C}(w). \tag{4.11}$$

Points on the boundary of the combined region maximise v_2 with respect to w while keeping v_1 fixed (or vice versa, maximising v_1 with v_2 fixed). This construction will be demonstrated explicitly in Sec. 5.1.

4.2.3 Three (or More) Parameter Uncertainty Relations

Part of the utility of the Cramér–Rao bounds is that, unlike the joint-measurement uncertainty relations (from which the LWUR was derived), they can be applied to more than two parameters[5]. The curve of a two-parameter uncertainty relation can be generalised to a surface or hypersurface for three or more parameter estimation respectively. The uncertainty relation construction is identical: points on the surface (hypersurface) can be determined by eliminating the weight matrix parameters from the defining equation by finding stationary points. The construction of a three-parameter estimation uncertainty relation will be demonstrated and discussed in Chapter 7.

[4]For systems with more than two parameters, the lines are replaced with planes or hyperplanes.

[5]I note that non-trivial preparation uncertainty relations have been developed for an arbirary number of observables in Ref. [83], but this has not been consolidated into the joint measurement formalism.

Chapter 5

General Results for Two-Parameter Estimation

In this chapter, I present several general results for two-parameter estimation. These results are included for several reasons: (i) to demonstrate proof-of-principle implementation of the methods presented in Chapter 4, (ii) to derive results which will be used in later chapters, and (iii) to provide partial, general comparisons of the Lu-Wang estimation bound and the Nagaoka Cramér–Rao bound. First, I derive an attainable uncertainty relation for qubit systems using the method presented in Sec. 4.2. Next, I calculate the LWB, as defined in Sec. 4.1.1, for the cases where the incompatibility coefficient satisfies $\tilde{c} = 0$ and $\tilde{c} = 1$. These edge cases of the incompatibility coefficient have been shown to be relevant for physical estimation problems [16, 81, 84]. Finally, I compare the derived result for the LWB with $\tilde{c} = 1$ to the analytic expression for the NCRB for qubit systems.

5.1 Qubit Nagaoka Uncertainty Relation

For estimating two parameters in any qubit system, the NCRB is attainable and has an analytical expression [68]. Here, I present a new uncertainty relation based on this expression and the method described in Sec. 4.2. This derivation will also be an explicit example of deriving an uncertainty relation from a Cramér–Rao bound. The NCRB for qubits with weight matrix W is [68]

$$\mathsf{Tr}[WV_{\boldsymbol{\theta}}] \geq \mathcal{C}_{\mathrm{N}} = \mathsf{Tr}[W\mathcal{F}^{-1}] + 2\sqrt{\mathsf{Det}[W\mathcal{F}^{-1}]} \tag{5.1}$$

where $\mathcal{F}$ is the SLDFIM. For convenience, I let $\mathcal{F}^{-1} = \left(\begin{smallmatrix} a & c \\ c & b \end{smallmatrix}\right)$ and $W = \left(\begin{smallmatrix} w & 0 \\ 0 & 2-w \end{smallmatrix}\right)$ with $0 < w < 2$. Then, Eq. (5.1) can be expanded to

$$wv_1 + (2-w)v_2 \geq wa + (2-w)b + 2\sqrt{w(2-w)}\sqrt{ab - c^2}, \tag{5.2}$$

which defines the individual weighting regions. The boundary of these regions is defined by the equality in Eq. (5.2), which defines a line for each w:

$$v_2 = \frac{w}{2-w}(a - v_1) + b + 2\sqrt{\frac{w}{2-w}}\sqrt{ab - c^2}. \tag{5.3}$$

It then remains to find the intersection of the regions defined by these lines. Points on the boundary of the combined region are found by maximising v_2 with respect to w while fixing v_1. This maximum is determined simply by finding stationary points:

$$0 = \frac{\partial v_2}{\partial w} = (a - v_1)\left(\frac{1}{2-w} + \frac{w}{(2-w)^2}\right) + \sqrt{\frac{2-w}{w}}\frac{2}{(2-w)^2}\sqrt{ab - c^2}. \tag{5.4}$$

Eqs. (5.3) and (5.4) can then be combined to eliminate w, resulting in the equation of a curve

$$v_2 = \frac{bv_1 - c^2}{v_1 - a}. \tag{5.5}$$

Equivalently and more elegantly,

$$(v_1 - a)(v_2 - b) = ab - c^2 = \mathsf{Det}[\mathcal{F}^{-1}] = 1/\mathsf{Det}[\mathcal{F}], \tag{5.6}$$

which defines a trade-off or uncertainty relation between v_1 and v_2. In particular, for qubit systems, this is a tight uncertainty relation: any pair (v_1, v_2) satisfying Eq. (5.6) can be achieved by a separable measurement because the weighted NCRB for qubits is attainable. Furthermore, I note that Eq. (5.6) reduces to the single-parameter SLDCRB for v_1 (v_2) in the limit that v_2 (v_1) tends to infinity.

5.2 Lu–Wang Estimation Bound for $\tilde{c} = 0$

For the specific case of the incompatibility coefficient $\tilde{c} = 0$, with SLDFIM written in the form $\left(\begin{smallmatrix} \mathcal{F}_{11} & \mathcal{F}_{12} \\ \mathcal{F}_{12} & \mathcal{F}_{22} \end{smallmatrix}\right)$, the LWUR is

$$\frac{1}{v_1\mathcal{F}_{11}} + \frac{1}{v_2\mathcal{F}_{22}} - 2\sqrt{\left(1 - \frac{1}{v_1\mathcal{F}_{11}}\right)\left(1 - \frac{1}{v_2\mathcal{F}_{22}}\right)} \leq 2. \tag{5.7}$$

For convenience, let

$$1 - \frac{1}{v_1 \mathcal{F}_{11}} = \delta_1, \text{ and } 1 - \frac{1}{v_1 \mathcal{F}_{11}} = \delta_2. \tag{5.8}$$

Then, the inequality (5.7) can be rearranged to

$$\left(\sqrt{\delta_1} + \sqrt{\delta_2}\right)^2 \geq 0, \tag{5.9}$$

which implies that $\sqrt{\delta_1} + \sqrt{\delta_2}$ is real-valued. Because δ_1 and δ_2 are real, this is satisfied if and only if $\delta_1 \geq 0$ and $\delta_2 \geq 0$, which are equivalent to the single-parameter SLDCRB conditions.

Thus, when $\tilde{c} = 0$, the condition imposed by the LWUR is only the single-parameter SLDCRB conditions, and the LWB is the sum of the variances saturating these single-parameter bounds. That is,

$$\mathcal{C}_{\text{LW},\,\tilde{c}=0} = \frac{1}{\mathcal{F}_{11}} + \frac{1}{\mathcal{F}_{22}} = \frac{\mathcal{F}_{11} + \mathcal{F}_{22}}{\mathcal{F}_{11}\mathcal{F}_{22}}. \tag{5.10}$$

This is bounded above (i.e., at best, equally as informative) as the SLDCRB, which is given by

$$\mathcal{C}_{\text{SLD}} = \mathsf{Tr}\left[\mathcal{F}^{-1}\right] = \frac{\mathcal{F}_{11} + \mathcal{F}_{22}}{\mathcal{F}_{11}\mathcal{F}_{22} - \mathcal{F}_{12}^2} \tag{5.11}$$

5.3 Lu–Wang Estimation Bound for $\tilde{c} = 1$

Here I use the Lagrange dual optimisation problem to derive the LWB when $\tilde{c} = 1$. In this scenario, the minimisation problem is

$$\mathcal{C}_{\text{LW},\,\tilde{c}=1} := \min\left\{ f_0 = v_1 + v_2 \;\middle|\; f_1 = \frac{1}{v_1 \mathcal{F}_{11}} + \frac{1}{v_2 \mathcal{F}_{22}} - 1 \leq 0 \right\}, \tag{5.12}$$

for a SLDFIM of the form $\left(\begin{smallmatrix} \mathcal{F}_{11} & \mathcal{F}_{12} \\ \mathcal{F}_{12} & \mathcal{F}_{22} \end{smallmatrix}\right)$. Recall from Sec. 4.1.2 that strong duality holds for the LWB minimisation problem, so we may compute the LWB by solving the dual

problem. The Lagrange dual function is the infimum of the Lagrangian:

$$\begin{aligned} g(\lambda) &= \inf_{v_1,v_2} \{f_0(v_1,v_2) + \lambda f_1(v_1,v_2)\} \\ &= \inf_{v_1,v_2} \left\{ v_1 + v_2 + \lambda \left(\frac{1}{v_1 \mathcal{F}_{11}} + \frac{1}{v_2 \mathcal{F}_{22}} - 1 \right) \right\} \\ &= \inf_{v_1} \left\{ v_1 + \frac{\lambda}{v_1 \mathcal{F}_{11}} \right\} + \inf_{v_2} \left\{ v_2 + \frac{\lambda}{v_2 \mathcal{F}_{22}} \right\} - \lambda. \end{aligned} \tag{5.13}$$

By separating the terms, each infimum can be determined by finding the stationary point. For the first:

$$\partial_{v_1} \left(v_1 + \frac{\lambda}{v_1 \mathcal{F}_{11}} \right) = 0 \Longrightarrow v_1 = \pm \frac{\sqrt{\lambda}}{\sqrt{\mathcal{F}_{11}}} \Longrightarrow \inf_{v_1} \left\{ v_1 + \frac{\lambda}{v_1 \mathcal{F}_{11}} \right\} = \frac{2\sqrt{\lambda}}{\sqrt{\mathcal{F}_{11}}}, \tag{5.14}$$

where the positivity of v_1 is used to select the positive solution. Solving the second infimum similarly gives the dual function

$$g(\lambda) = \sqrt{\lambda} \left(\frac{2}{\sqrt{\mathcal{F}_{11}}} + \frac{2}{\sqrt{\mathcal{F}_{22}}} \right) - \lambda. \tag{5.15}$$

Finally, using strong duality, the LWB is equal to the supremum of the dual function (subject to $\lambda \geq 0$):

$$\mathcal{C}_{\mathrm{LW},\, \tilde{c}=1} = \sup_{\lambda} \{ g(\lambda) \mid \lambda \geq 0 \} = \left(\frac{1}{\sqrt{\mathcal{F}_{11}}} + \frac{1}{\sqrt{\mathcal{F}_{22}}} \right)^2, \tag{5.16}$$

where the supremum is found by finding the stationary point of $g(\lambda)$.

5.4 Partial Comparison of the LWB and NCRB

Here I present a comparison between the LWB and the NCRB in a general qubit estimation problem with the restriction on the incompatibility coefficient: $\tilde{c} = 1$. I use the same general form for the SLDFIM as in Sec. 5.3. The NCRB can then be calculated using

Eq. (5.1)

$$\mathcal{C}_{\mathrm{N}} = \mathsf{Tr}[\mathcal{F}^{-1}] + 2\sqrt{\mathsf{Det}[\mathcal{F}^{-1}]} \tag{5.17}$$

$$= \frac{\mathcal{F}_{11} + \mathcal{F}_{22}}{\mathcal{F}_{11}\mathcal{F}_{22} - \mathcal{F}_{12}^2} + \frac{2}{\sqrt{\mathcal{F}_{11}\mathcal{F}_{22} - \mathcal{F}_{12}^2}}. \tag{5.18}$$

From the previous section, the LWB when $\tilde{c} = 1$ is

$$\mathcal{C}_{\mathrm{LW}} = \left(\frac{1}{\sqrt{\mathcal{F}_{11}}} + \frac{1}{\sqrt{\mathcal{F}_{22}}}\right)^2. \tag{5.19}$$

Now, notice that when $\mathcal{F}_{12} = 0$, $\mathcal{C}_{\mathrm{N}} = \mathcal{C}_{\mathrm{LW}}$, i.e., when $\tilde{c} = 1$ and the SLDFIM is diagonal, the NCRB and LWB are equal. This equality is actually the result of an underlying equivalence: for $\tilde{c} = 1$ and diagaonal SLDFIM, the Nagaoka uncertainty relation (Eq. (5.6)) is identical to the LWUR, which is hence attainable.

For $\mathcal{F}_{12} \neq 0$, the LWB is unchanged, but the NCRB increases because its denominators decrease. For $\tilde{c} = 1$, thus, $\mathcal{C}_{\mathrm{N}} \geq \mathcal{C}_{\mathrm{LW}}$. Additionally, as $\mathcal{F}_{12}$ tends to $\sqrt{\mathcal{F}_{11}\mathcal{F}_{22}}$, the NCRB tends to infinity as $\mathcal{F}$ becomes singular. This behaviour is sensible because the SLDCRB also tends to infinity in that limit. In this regard, the LWUR fails to capture the effect of the off-diagonal SLDFIM elements on the precision limits.

Chapter 6

Estimating Qubit State Rotations

In the previous chapter we saw that there is a class of estimation problems (two-parameter estimation for qubits, $\tilde{c} = 1$, and $\mathcal{F}$ diagonal) for which the NCRB and LWB are equally informative—they place the same limit on the MSE. It remains, however, to investigate whether there can be a gap between the NCRB and LWB, and how large such a gap can be. In this chapter, through a simple estimation problem, I demonstrate that the gap between the NCRB and LWB can be arbitrarily large. I also explicitly construct measurements which saturate the NCRB.

6.1 The Estimation Problem

The estimation problem I consider in this chapter is the problem of estimating the magnitude of two orthogonal rotations of a Bloch vector in the Bloch sphere. I will refer to this as the *qubit rotations* problem, and I use the problem as a testbed to compare the different estimation bounds. These parameters were chosen for their high incompatibility—the operators which encode them onto the probe state are non-commuting. There are two parts to the description of the estimation problem: the choice of the probe state, and the encoding of the parameters. As will become clear in Chapter 7, this rotations problem can be generalised to higher-dimensional quantum systems.

Consider the probe state described by the density matrix

$$\rho_0(r, \vartheta, \varphi) = \frac{1}{2}\left(\mathbb{1} + s_x\sigma_x + s_y\sigma_y + s_z\sigma_z\right) = \frac{1}{2}\begin{pmatrix} 1 + r\cos\vartheta & re^{-i\varphi}\sin\vartheta \\ re^{i\varphi}\sin\vartheta & 1 - r\cos\vartheta \end{pmatrix}. \tag{6.1}$$

Here, $(s_x, s_y, s_z)^\top$ is the Bloch vector defined by the spherical polar coordinates (r, ϑ, φ) that describes the probe state, as discussed in Chapter 2. This choice of description

allows for both the purity[1] of the state and its orientation relative to the rotation axes to be changed.

The parameters of interest are the rotation angles of the probe state about the x- and y-axes of the Bloch sphere: $\boldsymbol{\theta} = (\theta_x, \theta_y)^\top$. These may be encoded onto the state via the unitary channel:

$$\rho_{\boldsymbol{\theta}} = U(\theta_x, \theta_y)\rho_0 U^\dagger(\theta_x, \theta_y), \quad U(\theta_x, \theta_y) = R_y(\theta_y)R_x(\theta_x), \tag{6.2}$$

where $R_j(\theta) = e^{-i\theta\sigma_j/2}$ is called the rotation operator because it acts to rotate a Bloch vector by angle θ around the j-axis of the Bloch sphere [17]. This rotation is depicted in Figure 6.1. Note that the choice of rotation about the x- and y-axes is arbitrary—the system with rotations about any pair of orthogonal axes would be equivalent up to redefinition of the probe state, due to the commutation relations between σ_x, σ_y, and σ_z.

Intuitively, we should expect that the MSE will be smallest when the probe state is pure ($r = 1$), maximising the length of the Bloch vector, and when the probe state is aligned with the z-axis ($\theta = 0$ or π), because in this case the distinction between x and y rotations is greatest. On the other hand, we should expect that the MSE can be infinite. For example, if the probe state is along the x-axis it is insensitive to rotations about the x-axis, and θ_x cannot be estimated.

As described in Sec. 3.1.1, in this thesis I consider local parameter estimation. As such, the derivatives of the state ($\rho_{\boldsymbol{\theta}}$) with respect to the parameters may be approximated as independent of the true value of the parameter by evaluating at $\boldsymbol{\theta} = \mathbf{0}$, which simplifies calculations. It is straightforward to verify, using the chain rule on Eq. (6.2), that these state derivatives are:

$$\partial_j \rho_{\boldsymbol{\theta}} := \left.\frac{\partial \rho_{\boldsymbol{\theta}}}{\partial \theta_j}\right|_{\boldsymbol{\theta}=\mathbf{0}} = -\frac{i}{2}\sigma_j\rho_0 + \frac{i}{2}\rho_0\sigma_j = \frac{i}{2}[\rho_0, \sigma_j]. \tag{6.3}$$

[1] The purity of ρ is $\mathrm{Tr}[\rho^2]$. A pure state has a purity of 1.

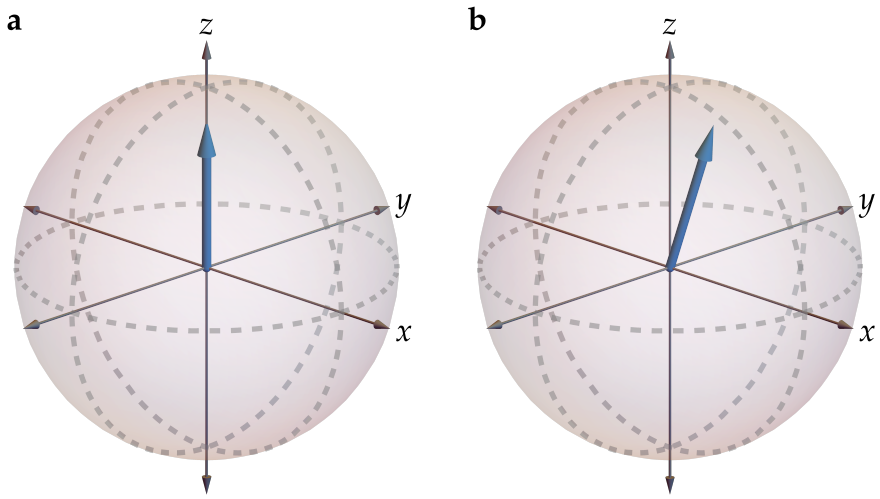

Figure 6.1: Depiction of rotation of a qubit state. **a**, probe state (blue) represented in the Bloch sphere. **b**, the probe state is rotated about the x- and y- axes of the Bloch sphere.

6.2 Qubit Rotations Estimation Bounds

In this section I calculate the estimation lower bounds for the aforementioned qubit rotations problem, so that they can be compared. Where possible, I calculate the bounds for an arbitrary probe state identified by (r, ϑ, φ). Owing to the system being two-dimensional, the NCRB and HCRB for this problem have exact solutions. Nevertheless, I demonstrate how these expressions can be obtained from numerical calculations, as these techniques will be used in Chapter 7, where general analytic expressions do not exist. A major focus of this study is the comparison of the NCRB and LWB, to elucidate when they provide different estimation limits.

6.2.1 Cramér–Rao Bounds

For this two-parameter estimation problem in a two-dimensional system, the Cramér–Rao bounds are easily computed, which I outline here. For simplicity, and for the purpose of comparing to the LWB, I consider the Cramér–Rao bounds with the identity weight matrix only.

The SLD operators can be computed using Eq. (3.25), which depends on the state

(Eq. (6.1)) and its derivatives (Eq. (6.3)). This yields the two operators

$$L_x = \begin{pmatrix} r\sin\varphi\sin\vartheta & ir\cos\vartheta \\ -ir\cos\vartheta & -r\sin\varphi\sin\vartheta \end{pmatrix}, \quad L_y = \begin{pmatrix} -r\cos\varphi\sin\vartheta & r\cos\vartheta \\ r\cos\vartheta & r\cos\varphi\sin\vartheta \end{pmatrix}. \tag{6.4}$$

The SLDFIM is then (using Eq. (3.22))

$$\mathcal{F} = r^2 \begin{pmatrix} 1-\cos^2\varphi\sin^2\vartheta & -\sin\varphi\cos\varphi\sin^2\vartheta \\ -\sin\varphi\cos\varphi\sin^2\vartheta & 1-\sin^2\varphi\sin^2\vartheta \end{pmatrix}, \tag{6.5}$$

and the SLDCRB is

$$\mathcal{C}_{\mathbf{SLD}} = \mathsf{Tr}\left[\mathcal{F}^{-1}\right] = \frac{1+\sec^2\theta}{r^2}. \tag{6.6}$$

For qubit systems the NCRB is given by Eq. (3.31), which using Eq. (6.5) is

$$\mathcal{C}_{\mathbf{N}} = \frac{(1+|\sec\vartheta|)^2}{r^2}. \tag{6.7}$$

Similarly, the HCRB has an analytic solution for qubit systems (see Appendix A), which can be computed as

$$\mathcal{C}_{\mathbf{H}} = \frac{1+2r|\sec\vartheta|+\sec^2\vartheta}{r^2}, \tag{6.8}$$

where it is evident that, as expected, for a pure state ($r = 1$) the NCRB and HCRB coincide. The above bounds do not depend on φ, however, for a general weight matrix one can verify that the SLDCRB, NCRB, and HCRB for the qubit rotations problem do depend on φ—the φ-symmetry is broken.

Determining bounds from numerical calculations

Here, I outline how the above expressions for the NCRB and HCRB can be determined based on the numerical results of the semidefinite program. As described in Sec. 3.2.3, the minimisation problem for the NCRB and HCRB is over Hermitian operators, hereafter referred to as "X matrices". Determining the solution of the minimisation problem is equivalent to determining the minimiser X matrices.

In the case of two-parameter estimation in qubits, there are two X matrices, one for each parameter, which are 2×2 matrices (due to the dimension of the Hilbert space). The X matrices are Hermitian so they each have four parameters, which are partially constrained by the three unbiased conditions in Eq. (3.29), which constrain the matrices to depend on only one free variable each.

For the qubit rotations problem, numerical results from the semidefinite program suggest that the minimiser X matrices for the NCRB and HCRB are identical. The optimal matrices were observed to be traceless, which provides a constraint. Note that the traceless condition may not be necessary for a solution to the minimisation problem—there may be other solutions, but nevertheless, enforcing it gives a solution. Solving these conditions gives the X matrices

$$X_1 = \begin{pmatrix} \frac{\sin\vartheta\sin\varphi}{r} & \frac{\cos\vartheta(\sin\varphi+i\cos\varphi)\left(\cos\varphi\sec^2\vartheta+i\sin\varphi\right)}{r} \\ \frac{\cos\vartheta(\sin\varphi-i\cos\varphi)\left(\cos\varphi\sec^2\vartheta-i\sin\varphi\right)}{r} & -\frac{\sin\vartheta\sin\varphi}{r} \end{pmatrix},$$
$$X_2 = \begin{pmatrix} -\frac{\cos\varphi\sin\vartheta}{r} & \frac{\sec\vartheta(\cos\varphi-i\sin\varphi)\left(\cos\varphi\cos^2\vartheta+i\sin\varphi\right)}{r} \\ \frac{\sec\vartheta(\cos\varphi+i\sin\varphi)\left(\cos\varphi\cos^2\vartheta-i\sin\varphi\right)}{r} & \frac{\cos\varphi\sin\vartheta}{r} \end{pmatrix}, \tag{6.9}$$

from which the NCRB and HCRB can be computed using Eqs. (3.30) and (3.27), respectively, and agree with the results from the exact solutions Eqs. (6.7) and (6.8). If it was not the case that exact solutions were available, the analytic expressions obtained from the X matrices could be verified by comparing to the values of the bounds from the semidefinite program.

6.2.2 Lu–Wang Uncertainty Relation, Bound, and Comparisons

The incompatibility coefficient of the LWUR is calculated from Eq. (3.35) and is

$$\tilde{c} = \frac{|\cos\vartheta|}{\sqrt{\left(\cos^2\vartheta+\cos^2\varphi\sin^2\vartheta\right)\left(\cos^2\vartheta+\sin^2\varphi\sin^2\vartheta\right)}}. \tag{6.10}$$

The fact that $\tilde{c}$ is not constant means that the analytic computation of the LWB for arbitrary (r, ϑ, φ) is non-trivial. Instead, in the following I present a series of comparisons between the LWB and the other bounds, for particular choices of the probe state. Of

particular interest is the existence of a gap between the LWB and NCRB.

Cases with zero gap

In Sec. 5.4 I proved that the NCRB and LWB coincide for a qubit estimation problem with $\tilde{c} = 1$ and a diagonal SLDFIM. In the case of the qubit rotations problem, it is straightforward to verify that these conditions are satisfied for the following probe states:

$$\vartheta = k\pi, \quad k \text{ integer, and arbitrary } r, \varphi, \tag{6.11}$$

$$\varphi = k\frac{\pi}{2}, \quad k \text{ integer, and arbitrary } r, \vartheta. \tag{6.12}$$

In Figure 6.2**a**, I present the regions defined by the LWUR and Cramér–Rao bounds for a single probe state with $(r, \vartheta, \varphi) = (\frac{1}{2}, 0, 0)$. In Figure 6.2**b**, I present the LWB and Cramér–Rao bounds for probe states with the same orientation ($\vartheta = \varphi = 0$), but varying r.

Case with a finite gap

For a probe state with $\vartheta = \varphi = \pi/4$ and arbitrary r, the incompatibility coefficient is $\tilde{c} = \sqrt{8/9}$, which is neither 0 nor 1, so we cannot use the results derived in Chapter 5. In Appendix B, I present the calculation of the LWB for this probe state using the KKT conditions described in Sec. 4.1.2. The result is $\mathcal{C}_{\mathrm{LW}} = 4/r^2$. It is interesting to compare this lower bound to the Cramér–Rao bounds.

In Figure 6.2**c**, I present the regions defined by the LWUR and Cramér–Rao bounds in this example with $r = 1/2$, and in Figure 6.2**d** I present the r-dependence of the estimation bounds. Most notably, the LWB is less than the NCRB for all values of r. This gap means that the NCRB is tighter than the LWB, and hence the LWB is not in general a tight estimation limit. Furthermore, for $r > 1/\sqrt{8}$, the LWB is even less informative than the HCRB.

The gap between the LWB and NCRB is present at $r = 1$ when the probe state is pure, despite the fact that the Lu–Wang IRTR, which the LWUR is based on, is tight

for pure states. Saturation of the IRTR for this pure state estimation problem will be revisited in Sec. 8.2.

A case with an infinite gap

For the probe state $\vartheta = \pi/2$ and arbitrary r, φ, the incompatibility coefficient $\tilde{c}$ is zero and the SLDFIM is

$$\mathcal{F} = r^2 \begin{pmatrix} \sin^2\varphi & -\sin\varphi\cos\varphi \\ -\sin\varphi\cos\varphi & \cos^2\varphi \end{pmatrix}. \tag{6.13}$$

Then, invoking the result from Sec. 5.2, the LWB is

$$\mathcal{C}_{\mathrm{LW}} = \frac{1}{r^2\sin^2\varphi} + \frac{1}{r^2\cos^2\varphi}, \tag{6.14}$$

which is finite for $r > 0$ and values of φ such that $\sin\varphi \neq 0$ and $\cos\varphi \neq 0$. On the other hand, the SLDFIM is singular (its determinant is zero), so the Cramér–Rao bounds are all infinite. The divergence of the Cramér–Rao bounds as $\vartheta \to \pi/2$ can be understood by considering the rotation of the probe state when $\vartheta = \pi/2$, which is when the probe state is in the x-y plane. In this case, one cannot distinguish between small rotations about the x- and y-axes, so there is no way to produce a precise estimate of both angles. This information is manifested in the off-diagonal elements of the SLDFIM, which the LWUR does not explicitly account for. Consequently, in this case the LWUR does not provide useful information about the precision limit.

Further comparison and trends with the incompatibility coefficient

It is interesting to note that the NCRB does not depend on the probe state angle φ while the incompatibility coefficient does. This further suggests that the LWB depends on φ too. To illustrate this, I calculated the NCRB, LWB and incompatibility coefficient for probe states with varying φ. I keep r fixed (at $r = 1$) because $\tilde{c}$ does not depend on r, and perform calculations for multiple values of ϑ. The results are presented in Figure 6.3, in which I use colour to indicate the value of φ used for each point—φ is varied only from 0 to $\pi/4$ due to the φ-symmetry of the problem.

The results in Figure 6.3 demonstrate that the ultimate precision limit is not neces-

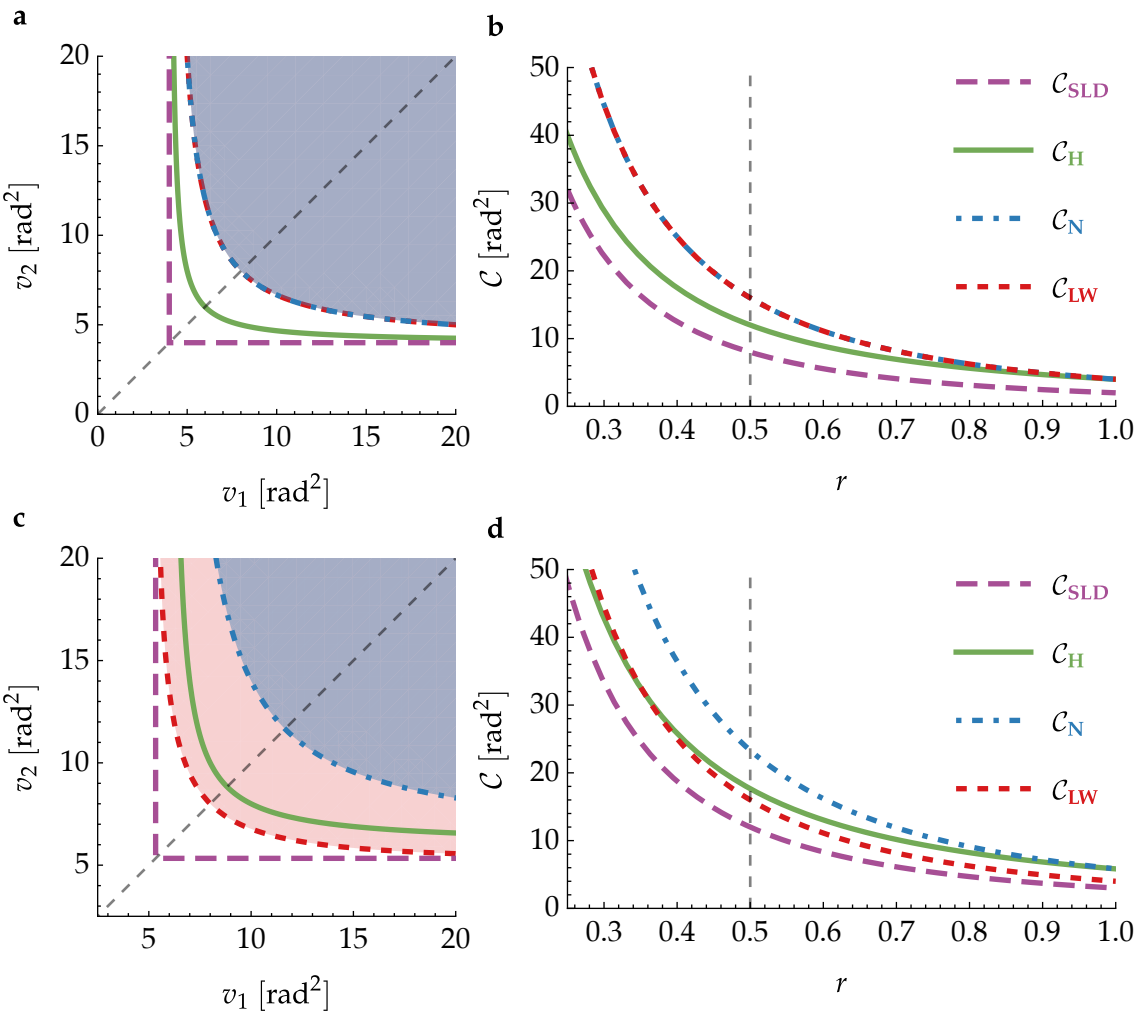

Figure 6.2: Visualisations of qubit rotations estimation limits. **a**, estimation limit regions for a $\tilde{c} = 1$ case, $(r, \vartheta, \varphi) = (\frac{1}{2}, 0, 0)$, where the regions for the NCRB-region and LWUR are identical. **b**, estimation bounds ($\mathcal{C}$) for a $\tilde{c} = 1$ case ($\vartheta = \varphi = 0$), as a function of r. **c**, estimation limit regions for a $\tilde{c} = \sqrt{8/9}$ case, $(r, \vartheta, \varphi) = (\frac{1}{2}, \frac{\pi}{4}, \frac{\pi}{4})$, where the regions for the NCRB and LWUR are different. **d**, estimation bounds as a function of r for the probe state with $\vartheta = \varphi = \pi/4$. The LWB is less than the NCRB for all r, and is less than the HCRB for $r > 1/\sqrt{8}$. The grey dashed line in the region plots (left) mark equal variances and its intersection with the region boundaries corresponds to the LWB and Cramér–Rao bounds. The grey dashed line in the $\mathcal{C}$-vs-r plots (right) denotes the r-value in the region plots.

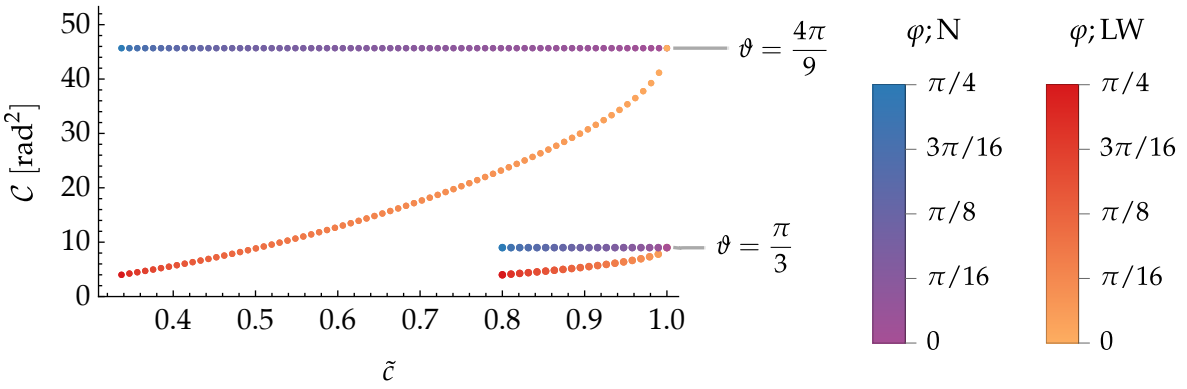

Figure 6.3: Qubit rotations NCRB (blue-purple) and LWB (red-yellow) as a function of $\tilde{c}$ for values $\vartheta = \frac{4\pi}{9}$ and $\frac{\pi}{3}$. The value of $\tilde{c}$ is determined by sampling values of φ, which are indicated by colour. For these calculations the probe state has $r = 1$, and note that the range of $\tilde{c}$ is different for the different values of ϑ; the full range $0 \leq \varphi \leq \pi/4$ is used for both series. This demonstrates that the ultimate precision limit (here, the NCRB) need not be related to the incompatibility coefficient ($\tilde{c}$).

sarily influenced by the incompatibility coefficient, because here the NCRB is a constant function of $\tilde{c}$. Furthermore, it is interesting that the NCRB and LWB coincide only when $\tilde{c} = 1$—this behaviour will be explored further in Sec. 8.1.

6.3 Optimal Measurements Saturating the NCRB

Owing to Nagaoka, we know that the NCRB in this two-parameter qubit estimation problem is tight. In this section, I present a POVM with a high degree of symmetry saturating the bound.

I numerically find optimal 4-element POVMs with elements that are projection operators of the form $\Pi_i = a_i \, |\psi_i\rangle\langle\psi_i|$. After experimentation, I determined that the states $|\psi_i\rangle$ can be constrained to have certain symmetry while not affecting the performance of the measurement. The result is a POVM which has two free parameters (which I

denote α and β) that need to be optimised over, using the ansatz

$$\begin{aligned}
\Pi_j &= \frac{1}{2}\left|\psi_j\right\rangle\left\langle\psi_j\right|,\ j=1,\ldots,4,\ \text{with}\\
\left|\psi_1\right\rangle &= \cos\left(\frac{\alpha}{2}\right)\left|0\right\rangle + \sin\left(\frac{\alpha}{2}\right)e^{i\left(\beta+\varphi-\frac{\pi}{4}\right)}\left|1\right\rangle,\\
\left|\psi_2\right\rangle &= -\sin\left(\frac{\alpha}{2}\right)\left|0\right\rangle + \cos\left(\frac{\alpha}{2}\right)e^{i\left(\beta+\varphi-\frac{\pi}{4}\right)}\left|1\right\rangle,\\
\left|\psi_3\right\rangle &= \cos\left(\frac{\alpha}{2}\right)\left|0\right\rangle + \sin\left(\frac{\alpha}{2}\right)e^{i\left(-\beta+\varphi+\frac{\pi}{4}\right)}\left|1\right\rangle,\\
\left|\psi_4\right\rangle &= -\sin\left(\frac{\alpha}{2}\right)\left|0\right\rangle + \cos\left(\frac{\alpha}{2}\right)e^{i\left(-\beta+\varphi+\frac{\pi}{4}\right)}\left|1\right\rangle.
\end{aligned} \tag{6.15}$$

A priori, the variables α and β may depend on r, ϑ, and φ, but I found that they only depend on ϑ—the φ-dependence is accounted for in the phase factors. The Bloch vectors of the POVM elements form an 'X' shape in the Bloch sphere, as depicted in the inset of Figure 6.4**a**. The parameter α is the angle between the Bloch vectors and the z-axis, β controls the separation between the two arms of the 'X', and φ rotates all of the vectors around the z-axis.

I determined the optimal values of α and β by computing the CFIM ($F(\Pi)$) of the measurement (using Eq. (3.13) and the Born rule Eq. (2.5)) and minimising $\mathrm{Tr}\left[F^{-1}\right]$ numerically in `Mathematica`. Finally, to verify that the measurement is optimal, I compare $\mathrm{Tr}\left[F^{-1}\right]$ to the exact calculation of the NCRB. In Figure 6.4**a**, I present the numerically determined optimal values of α and β as a function of ϑ. The values define the optimal POVM for the qubit rotations problem with any probe state; interestingly, the Bloch vectors of the optimal POVM elements lie in the plane in the Bloch sphere orthogonal to the probe state. In Figure 6.4**b**, I present the normalised difference between the measurement MSE from the CFIM, and the NCRB. The difference is normalised because both quantities diverge as ϑ approaches $\pi/2$. This demonstrates saturation of the NCRB.

This POVM (Eq. (6.15)) is similar to an optimal POVM determined by Conlon *et al.* in Ref. [13], who investigated a qubit rotations problem for a probe state with $\theta = 0$. From their result, I expect that the values $\alpha = \pi/2$ and $\beta = 0$ should be an optimal solution for $\vartheta = 0$, which is indeed what I find numerically.

To describe the optimal estimation process using the above optimal POVM it re-

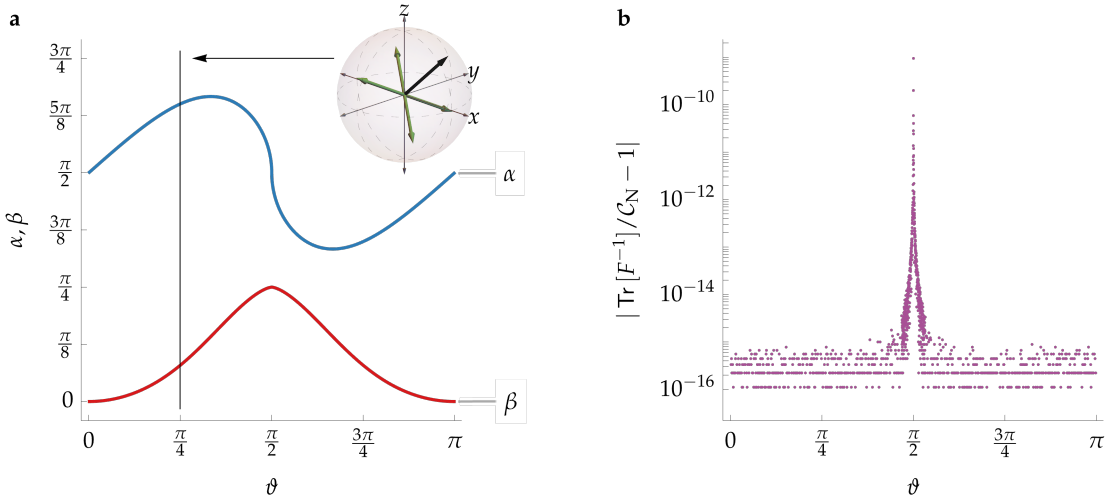

Figure 6.4: **a**, Numerically determined values for α and β which optimise the POVM in Eq. (6.15), as a function of ϑ. The inset depicts the Bloch vectors of the POVM (green) and probe state (black) for the values of ϑ, α, β indicated by the line and $\varphi = \pi/8$. **b**, Normalised difference between the performance of the POVM (quantified by $\mathrm{Tr}\left[F^{-1}\right]$) and the NCRB. The normalisation is included because both quantities diverge at $\vartheta = \pi/2$.

mains to determine the estimator, which here is described by the estimator coefficients $\{\hat{\theta}_{ij} \mid i = 1, 2,\ j = 1, \ldots, 4\}$. These coefficients are related to the optimal POVM and the observables X_1, X_2 (Eq. (6.9)) by

$$X_i = \sum_j \hat{\theta}_{ij} \Pi_j, \tag{6.16}$$

where the θ_i term in Eq. (3.28) is zero. To solve Eq. (6.16) for the estimator coefficients, it is convenient to rewrite the equation by converting the matrices into vectors. Let $u_i = \left(\hat{\theta}_{i1}, \hat{\theta}_{i2}, \hat{\theta}_{i3}, \hat{\theta}_{i4}\right)^\top$ (to be determined), $v_i = \left((X_i)_{11}, (X_i)_{12}, (X_i)_{21}, (X_i)_{22}\right)^\top$, and

$$A = \begin{pmatrix} (\Pi_1)_{11} & (\Pi_2)_{11} & (\Pi_3)_{11} & (\Pi_4)_{11} \\ (\Pi_1)_{12} & (\Pi_2)_{12} & (\Pi_3)_{12} & (\Pi_4)_{12} \\ (\Pi_1)_{21} & (\Pi_2)_{21} & (\Pi_3)_{21} & (\Pi_4)_{21} \\ (\Pi_1)_{22} & (\Pi_2)_{22} & (\Pi_3)_{22} & (\Pi_4)_{22} \end{pmatrix}. \tag{6.17}$$

Then, Eq. (6.16) is equivalent to the linear equation $v_i = A u_i$. The matrix A can be shown to be singular, so the solution cannot be found by inverting A. Nevertheless,

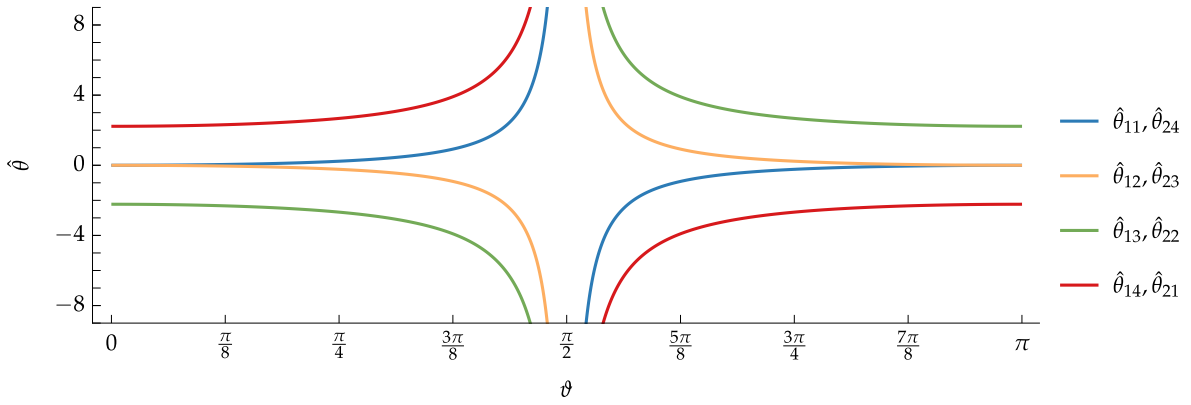

Figure 6.5: Estimator coefficients for the POVM in Eq. (6.15) and the probe state with $\varphi = \pi/4$ and $r = 0.9$, as a function of ϑ. Due to the symmetry of this probe state, pairs of the coefficients are equal.

a solution can be determined using the Moore–Penrose pseudoinverse, the matrix A^+ which satisfies $AA^+A = A$, which gives the solution $u_i = A^+v_i$ [85]. A depends on the values of α and β, and the solution for the estimator coefficients were determined numerically using `PseudoInverse` in `Mathematica`. The dependence of the estimator coefficients on the probe state angle ϑ, for the probe state with $\varphi = \pi/4$ and $r = 0.9$, is presented in Figure 6.5. In this case, owing to the x-y symmetry when $\varphi = \pi/4$, there are pairs of coefficients which are identical, but this is not the case in general. For any probe state, the estimator coefficients can be calculated following the same procedure.

The fact that the matrix A is singular is a consequence of the POVM elements not being linearly independent: $\Pi_1 + \Pi_2 = \Pi_3 + \Pi_4$. This suggests that there exists a 3-outcome POVM that is as optimal as the POVM I found. Such a 3-element POVM may be simpler because it has one fewer outcome, but the elements may not be rank-1, and likely do not have as much symmetry as the elements in Eq. (6.15).

In this chapter, I have presented the estimation limits for a model two-parameter estimation problem. I explicitly demonstrated that the NCRB is tight by finding an optimal POVM saturating the bound. For configurations of this model problem, I found examples where the LWUR is as informative as the NCRB, and also examples where there is a finite or infinite gap between the NCRB and LWB. Somewhat surprisingly, I also demonstrated that the incompatibility coefficient need not be related to the ultimate precision limit.

Chapter 7

Higher-Dimensional Rotations

In the previous chapter, I demonstrated that the qubit rotations problem unmasked a variety of differences between the NCRB and the LWUR approach to estimation bounds. In this chapter, I present an extension of the qubit rotations problem to three-parameter estimation and construct, via the NHCRB, an attainable three-parameter estimation uncertainty relation. This construction cannot be performed using the LWUR because the trade-off relation only simultaneously considers two parameters. In order to extend the qubit rotations problem, it is necessary to generalise to a higher-dimensional system. This necessity can be understood by noticing that any rotation of a qubit state can be decomposed as two rotations about orthogonal axes—uniquely estimating a third angle is not possible. I describe the generalisation and then specialise to the three-dimensional case in the following.

7.1 Generalising Rotations

The qubit rotations analysed in Chapter 6 were encoded via unitary operators in the *special unitary Lie group* SU(2) of 2×2 unitary matrices with determinant 1. In particular, the rotation matrices are generated by the Pauli matrices (that is, obtained via exponentiating the Pauli matrices), which are the *generators* of the *special unitary Lie algebra* $\mathfrak{su}(2)$. Here, I will describe only the necessary details of the Lie groups and Lie algebras; for further details, I refer the reader to Ref. [86]. Importantly, there exist higher-dimensional analogues of $\mathrm{SU}(2)$ and $\mathfrak{su}(2)$, denoted $\mathrm{SU}(d)$ and $\mathfrak{su}(d)$ in dimension d.

The analogue of the Pauli matrices in Hilbert space dimension $d = 3$ (i.e., for qutrit systems), called the Gell-Mann matrices, are

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \tag{7.1}$$
$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

In higher dimensions, the $\mathfrak{su}(d)$ generators are called generalised Gell-Mann matrices, of which $d(d-1)/2$ are real symmetric matrices, $d(d-1)/2$ are imaginary antisymmetric matrices, and $d-1$ are diagonal matrices [87].

Unitary operators can be constructed analogously to the qubit case, $R_j(\theta) = e^{-i\theta\lambda_j/2}$, and the commutativity of these operators is tied to the commutativity of the Gell-Mann or generalised Gell-Mann matrices. I will refer to these operators as *rotation* operators and call the problem of estimating the rotation angles (θ) the *qudit rotations* problem, despite the lack of correspondence between the representations of the Hilbert space and physical rotations. The quantum state derivatives follow the same relation as in the qubit case, with σ_j replaced by λ_j: $\partial_j \rho_{\boldsymbol{\theta}} = \frac{i}{2}[\rho_0, \lambda_j]$.

The second part of generalising the rotation problem to higher dimensions is choosing a probe state—the analogue of choosing (r, ϑ, φ) for the qubit case. The topic of Sec. 7.2 is qutrit rotations, where I will discuss possible choices of probe states and "rotation axes". For higher dimensions, I anticipate that it will not be particularly instructive to generalise these choices explicitly—the qutrit example will be sufficient to demonstrate an $(n > 2)$-parameter uncertainty relation—so I omit further discussion on more general qudit probe states and rotations from this thesis.

7.2 Qutrit Rotations Estimation Bounds

A qutrit system is the simplest quantum system where we can encode three parameters via rotation operators. I use this system and generalised rotations to describe a

three-parameter estimation problem. A general qutrit state can be written as

$$\rho = \frac{1}{3}\mathbb{1}_3 + \mathbf{n} \cdot \boldsymbol{\lambda}, \tag{7.2}$$

where $\mathbf{n}$ is an 8-dimensional real vector satisfying a normalisation constraint to ensure ρ's positivity and $\boldsymbol{\lambda}$ is the vector of Gell-Mann matrices. A full parameterisation of a qutrit thus requires eight parameters, analogous to the three Bloch sphere coordinates for qubits. In this thesis, I focus on the relationship between the estimation bounds and the probe state. A subset of qutrit states will be sufficient for this investigation provided that the subset is broad enough, for example, that it contains both pure and mixed states. For simplicity, I will use the 1-parameter family described in Ref. [88] and defined by the diagonal density matrix

$$\rho_0(z) = \frac{1}{3}\mathbb{1}_3 + \frac{z}{\sqrt{3}}\lambda_8 = \text{Diag}\left(\frac{1+z}{3}, \frac{1+z}{3}, \frac{1-2z}{3}\right), \tag{7.3}$$

where $-1 \leq z \leq 1/2$. At $z = -1$, the state is pure, and the purity decreases as z goes to zero from either side. The bounds on z ensure that ρ_0 is positive semidefinite.

As discussed in Sec. 7.1, for qutrit rotations, the derivatives of the probe state are proportional to the commutator with Gell-Mann matrices. Thus, for non-trivial encoding of the parameters (i.e., non-zero derivative), rotations generated by Gell-Mann matrices that do not commute with λ_8 can be chosen. Due to the block-diagonal form of the matrices (see Eq. (7.1)), it is evident that those which do not commute with λ_8 are λ_4, λ_5, λ_6, and λ_7. I choose to estimate rotations generated by $\lambda_4, \lambda_5, \lambda_6$ and I denote by $\theta_4, \theta_5, \theta_6$ the (small) angles of rotation.

In the following, I present the calculation of the NHCRB for this qutrit rotations problem with an arbitrary diagonal weight matrix. This will enable the construction of a three-parameter estimation uncertainty relation.

7.2.1 Weighted NHCRB Calculation

In this estimation problem, the NHCRB can be computed numerically. However, we need an analytic expression for the weighted bound to construct an analytic estima-

tion uncertainty relation based on the NHCRB. The way I approach determining this analytic expression is to use the results of the semidefinite program (SDP) to determine the optimiser matrix $\mathbb{L}$ from Eq. (3.32), which a priori is a 9×9 complex Hermitian matrix and may depend on both the probe state and the weight matrix (W). The bound can then be calculated as $\mathcal{C}_{\text{NH}} = \mathbb{T}\text{r}[(W \otimes \rho)\,\mathbb{L}]$. As it turns out, it is easier to determine the form of $\mathbb{L}$ than determining $\mathcal{C}_{\text{NH}}$ directly. All calculations are performed using a diagonal weight matrix of the form $W = \text{Diag}\,(a, b, c)$.

Using the numerical results of the SDP, which outputs both the optimal value of the objective ($\mathbb{T}\text{r}[(W \otimes \rho)\,\mathbb{L}]$) and the optimiser ($\mathbb{L}$), I determined that there was degeneracy in the optimiser. That is, additional constraints could be placed on $\mathbb{L}$ without changing the minimum value of the objective. The additional constraints that make the matrix diagonal and set some other values to zero are

$$\mathbb{L}_{11} = \sqrt{\frac{b}{a}}\,\mathbb{L}_{44}, \quad \mathbb{L}_{22} = \mathbb{L}_{55} = \mathbb{L}_{77} = 0, \quad \mathbb{L}_{ij} = 0 \text{ if } i \neq j, \tag{7.4}$$

which drastically reduces the number of unknown $\mathbb{L}$-elements to six real values. The non-zero entries may each depend on the weights a, b, c and the probe state variable z, but the additional constraints are independent of the probe state. The fact that the off-diagonal terms can be set to zero without changing the solution is not too surprising: with a diagonal weight matrix and diagonal density matrix, any off-diagonal elements of $\mathbb{L}$ do not contribute to the trace and are thus irrelevant to the objective. To determine the dependence of the diagonal $\mathbb{L}$ elements on a, b, c, z, I varied one at a time. I fitted the numerical SDP values of the diagonal elements with the aid of `Mathematica`. The resulting optimal matrix is

$$\mathbb{L} = \text{Diag}\left(\frac{\sqrt{a}+\sqrt{b}+\sqrt{c}}{z^2\sqrt{a}}, 0, \frac{\sqrt{a}+\sqrt{b}}{z^2\sqrt{a}}, \frac{\sqrt{a}+\sqrt{b}+\sqrt{c}}{z^2\sqrt{b}}, 0, \frac{\sqrt{a}+\sqrt{b}}{z^2\sqrt{b}}, 0, \frac{\sqrt{a}+\sqrt{b}+\sqrt{c}}{z^2\sqrt{c}}, \frac{1}{z^2}\right). \tag{7.5}$$

This result is symmetric with respect to swapping a and b, which is expected due to the equality of first and second diagonal elements of the probe state density matrix.

Using the optimal $\mathbb{L}$, the weighted NHCRB is

$$\mathcal{C}_{\mathrm{NH}}(z) = \frac{2\left(\left(\sqrt{a}+\sqrt{b}\right)^2 + \left(\sqrt{a}+\sqrt{b}\right)\sqrt{c}+c\right) - z\left(\sqrt{a}+\sqrt{b}-\sqrt{c}\right)^2}{3z^2}. \tag{7.6}$$

I verified the correctness of this analytic solution by comparing it against numerical SDP results for random states and random weight matrices. In total, 32 random values of $-1 \leq z \leq 1/2$ were tested with 32 of each $0 < a, b, c < 3$, for a total of $32^4 = 1\,048\,576$ random estimation problems. Due to the divergence of the bound at $z = 0$, to assess the validity of the analytic solution, the absolute difference between the analytical and numerical values is normalised by the SDP solution. Out of all the random tests, the maximum normalised difference was 7.4486×10^{-8}. In the limit of setting two of the weight parameters to zero, the bound in Eq. (7.6) reduces to the SLDCRB, as expected.

7.2.2 Three-Parameter Estimation Uncertainty Relation

In this subsection, I demonstrate the construction of a three-parameter uncertainty relation based on the NHCRB. I use the method described in Sec. 4.2 and the weighted bound calculated in the previous section. I further demonstrate, numerically, that there exist measurements saturating the NHCRB and this uncertainty relation.

The LWUR is manifestly suboptimal for this estimation problem (and indeed, any problem with more than two parameters) because it is only a trade-off between pairs of MSE elements. On the contrary, the uncertainty relation I present simultaneously defines a non-trivial trade-off between all three MSE elements.

The weighted NHCRB in Eq. (7.6) defines a series of planes in the real space of MSE elements $\{(v_1, v_2, v_2) \in \mathbb{R}^3\}$ via the equation

$$av_1 + bv_2 + cv_3 = \mathcal{C}_{\mathrm{NH}}(z; a, b, c). \tag{7.7}$$

Points on the uncertainty relation surface can be found by fixing two coordinates (say v_1, v_2) and maximising the third (v_3) with respect to the weights. The plane's orienta-

tion is dictated by the ratio of the weights a, b, c, i.e., their overall scale is unimportant, and hence one can be eliminated. Due to the dependence of the bound on the weights, it is convenient to parameterise the weight matrix as

$$W = \text{Diag}\,(a, b, c) = \text{Diag}\left(s^2, t^2, (1-t-s)^2\right). \tag{7.8}$$

Eq. (7.7) then becomes:

$$s^2 v_1 + t^2 v_2 + (1-s-t)^2 v_3 = \mathcal{C}_{\text{NH}}(z; s^2, t^2, (1-s-t)^2), \tag{7.9}$$

and conditions for maximising v_3 can be obtained by differentiation with respect to s and t:

$$2sv_1 - 2(1-s-t)v_3 = \frac{\partial}{\partial s}\mathcal{C}_{\text{NH}}(z; s^2, t^2, (1-s-t)^2), \tag{7.10}$$

$$2tv_2 - 2(1-s-t)v_3 = \frac{\partial}{\partial t}\mathcal{C}_{\text{NH}}(z; s^2, t^2, (1-s-t)^2). \tag{7.11}$$

Then, Eqs. (7.9), (7.10) and (7.11) can be combined to eliminate s and t. The elimination was performed in `Mathematica`, and the result is a relation between the three variances:

$$v_1\left(3v_2v_3z^4 + (v_2+v_3)z^3 - 2(v_2+v_3)z^2 - 2z + 1\right) + v_2\left(v_3z^3 - 2v_3z^2 - 2z + 1\right) = 0. \tag{7.12}$$

Replacing the equality in Eq. (7.12) with an inequality ($\geq$) defines an estimation uncertainty relation for the MSE elements of the above qutrit estimation problem. The boundary of this uncertainty relation is presented as the surface in Figure 7.1**a** for the probe state with $z = -1/2$. In Figure 7.1**b**, I present contours of the uncertainty relation surface, to aid in its visualisation.

I numerically determined the MSE elements of optimal measurements in `MATLAB` by optimising the weighted trace of the MSE of a measurement described by a POVM, using publicly available code [65]. The result is points (v_1, v_2, v_3) which lie on the surface of the uncertainty relation, as was verified by comparing the numerical MSE to the NHCRB in Eq. (7.6). These points are overlayed on the uncertainty relation in Figure 7.1**a**. The radial pattern of the points is due to the choice of weight parameters.

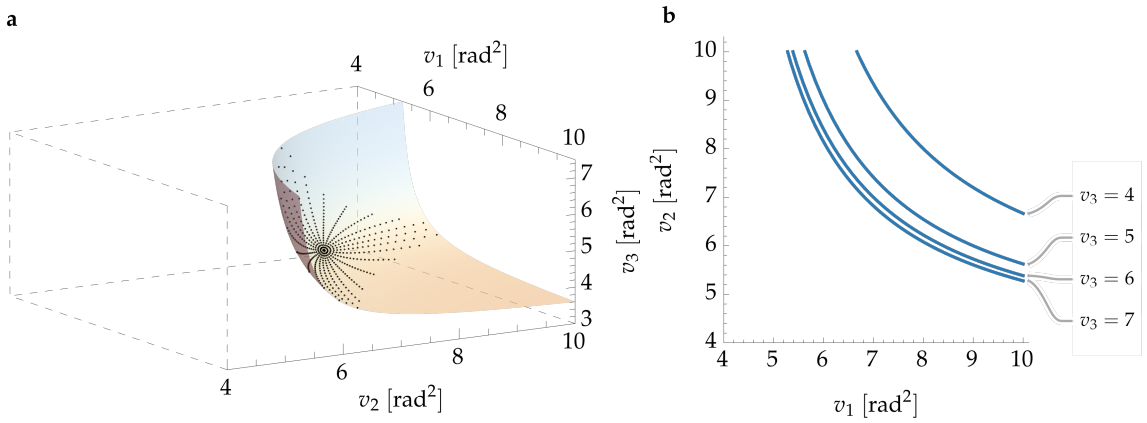

Figure 7.1: a, Three-parameter uncertainty relation for the qutrit rotations problem with $z = -1/2$. The volume on and above the surface is not forbidden by the NHCRB. The points (black) denote the MSE elements of measurements saturating the weighted NHCRB and hence saturating the uncertainty relation. **b**, Contours of the three-parameter uncertainty relation in **a**, for evenly spaced values of the variance v_3 (in units of rad^2).

I performed calculations starting with the unweighted case (centre of the pattern), and subsequently calculated the weighted bound for weight parameters incrementally away from $a = b = c = 1$. In this way, the optimal POVMs could be iteratively used as initial points for subsequent calculations, decreasing computational time.

The result in Eq. (7.12) is important because the relationship between the MSE elements represents the ultimate precision for estimating the three rotations, independent of the weight that is assigned to each parameter. While we have studied a simple, highly symmetric estimation problem, the method demonstrated in this section can in principle be applied to any multiparameter estimation problem. This extends beyond three-parameter estimation, where a hypersurface may not be visualised, but the ultimate relationship between MSE elements can nevertheless be determined.

7.2.3 LWUR for Higher-Dimensional Rotations

In Sec. 5.4, I demonstrated that for a qubit estimation problem with incompatibility coefficient $\tilde{c} = 1$ and diagonal SLDFIM, the LWB, which has a simple expression derived in Sec. 5.3, is equal to the NCRB. This offers a pathway for the simple calculation of the estimation bound, but the argument was restricted to two-dimensional systems. To

investigate this behaviour in a higher-dimensional system, I restrict the qutrit rotations problem to two parameters. In particular, I choose rotations generated by the Gell-Mann matrices λ_4 and λ_5. It is then straightforward to verify, using Eqs. (3.25) and (3.22), that the SLDFIM is diagonal:

$$\mathcal{F} = \begin{pmatrix} \frac{3z^2}{2-z} & \\ & \frac{3z^2}{2-z} \end{pmatrix}, \tag{7.13}$$

and $\tilde{c} = 1$. The LWB, given by Eq. (5.16), is then

$$\mathcal{C}_{\text{LW}} = \left(\frac{1}{\sqrt{\mathcal{F}_{11}}} + \frac{1}{\sqrt{\mathcal{F}_{22}}} \right)^2 = \frac{8 - 4z}{3z^2}. \tag{7.14}$$

Interestingly, this coincides with the NCRB calculated by setting $(a, b, c) = (1, 1, 0)$ in Eq. (7.6). This means that we have an example of a higher-dimensional estimation problem in which the NCRB and LWB are equal when $\tilde{c} = 1$ and the SLDFIM is diagonal, the same conditions as the qubit case. This leads me to make the following conjecture:

Conjecture 1 For any two-parameter estimation problem, if the incompatibility coefficient satisfies $\tilde{c} = 1$, then the NCRB and LWB are equal.

Higher-Dimensional Rotations

To provide evidence supporting Conjecture 1, I present a simple generalisation of the rotations to dimension d. In analogy to the above qutrit rotations, consider the d-dimensional probe state and derivatives

$$\rho = \text{Diag}\left(\frac{1}{d - 1/2}, \ldots, \frac{1}{d - 1/2}, 1 - \frac{d-1}{d - 1/2} \right), \quad \partial_j \rho = \frac{i}{2}[\rho, M_j], \tag{7.15}$$

where

$$M_1 = \begin{pmatrix} & & 1 \\ & & \\ 1 & & \end{pmatrix} \text{ and } M_2 = \begin{pmatrix} & & -i \\ & & \\ i & & \end{pmatrix}, \tag{7.16}$$

where all entries in $M_{1,2}$, except the top-right and bottom-left corners, are zero. I find that in this problem $\tilde{c} = 1$ and the NCRB and LWB are equal; see Appendix C for details. I present further evidence supporting the conjecture in the next chapter.

Chapter 8

Studies Involving Randomness

In Chapters 6 and 7, I demonstrated a number of properties of the Cramér–Rao bounds and the LWUR as well as their differences via explicit examples of estimation problems. While this was able to uncover several features of the estimation limits, it does not necessarily give a complete picture. In order to explore a more complete problem space, in this chapter I investigate features of the estimation limits using randomly generated estimation problems and randomly generated measurements.

8.1 Numerical Comparison of the NCRB and LWB

In this section, I present the results of using randomly generated estimation problems to investigate the differences between the NCRB and LWB. The purpose of this investigation was to understand whether the ordering $\mathcal{C}_{\mathrm{N}} \geq \mathcal{C}_{\mathrm{LW}}$ holds in dimensions other than $d = 2$, and if so, whether this provides evidence for Conjecture 1, that equality holds if $\tilde{c} = 1$. I begin by outlining how to generate a random estimation problem.

8.1.1 Generating Random Estimation Problems

As far as the estimation limits—in this case, the NCRB and LWB—are concerned, any local estimation problem consists of two parts: the probe state density matrix onto which the parameters are encoded (ρ), and the derivatives of the probe state post-encoding with respect to the parameters: $\{\partial_1\rho, \partial_2\rho\}$. This means that generating a random estimation problem is equivalent to generating a random density matrix and random derivatives. Here, I consider only two-parameter estimation because we are investigating the limits imposed by the LWUR.

A density operator is a postive semidefinite Hermitian operator with unit trace.

Thus, a random[1] d-dimensional density matrix, ρ, can be constructed by first generating a random complex $d \times d$ matrix, M, and then setting $\rho = M^\dagger M / \operatorname{Tr}[M^\dagger M]$, which ensures both positivity and $\operatorname{Tr}[\rho] = 1$.

The derivatives of the probe state must be traceless Hermitian operators. The traceless condition arises due to the unit trace property of the density operator: $0 = \partial_i \operatorname{Tr}[\rho] = \operatorname{Tr}[\partial_i \rho]$. A random Hermitian operator can be constructed by generating a random complex $d \times d$ matrix M and taking the construction $M + M^\dagger$. Then, the trace can be removed by subtracting $\operatorname{Tr}[M + M^\dagger]/d$ from each diagonal element.

8.1.2 The NCRB, LWB, and Incompatibility Coefficient

I generated random estimation problems in different Hilbert space dimensions. For each problem, I computed the NCRB using its semidefinite program and the LWB using the numerical code described in Sec. 4.1.4. For each problem, I also calculate the incompatibility coefficient $\tilde{c}$. Here, I present the results of these simulations for a million random problems in each of the dimensions $2, 3, 4$, and 5.

The absolute scales of the estimation problems are arbitrary due to their random construction, so a plot of the absolute estimation limits is not instructive. Instead, In Figure 8.1, I present the difference between the NCRB and LWB, normalised by the NCRB, as a function of the incompatibility coefficient. The scatter plots for dimensions 3-5 include only 1% of the collected data; the rest of the data is presented in the shaded regions which are the smallest convex region containing all points (the convex hull). I also include the point $(\tilde{c}, (\mathcal{C}_{\mathrm{N}} - \mathcal{C}_{\mathrm{LW}})/\mathcal{C}_{\mathrm{N}}) = (1, 0)$ in these regions, because this point is possible as demonstrated in Sec. 7.2.3. For the qubit problems, I present all points as the convex hull is not representative of the data—it is the region bounded by the black dashed line and the line (not shown) connecting the points $(0, 1)$ and $(1, 0)$, a large portion of which no points were found in. The coloured regions are representative of the space that is accessible to any estimation problem in that dimension, and it is evident that the region for qubit problems is qualitatively different to the regions for the higher dimensions. Nevertheless, the observation that the vertical coordinate of

[1] Here, I use density operators that are distributed uniformly in the Hilbert–Schmidt measure. See Ref. [89] for more details.

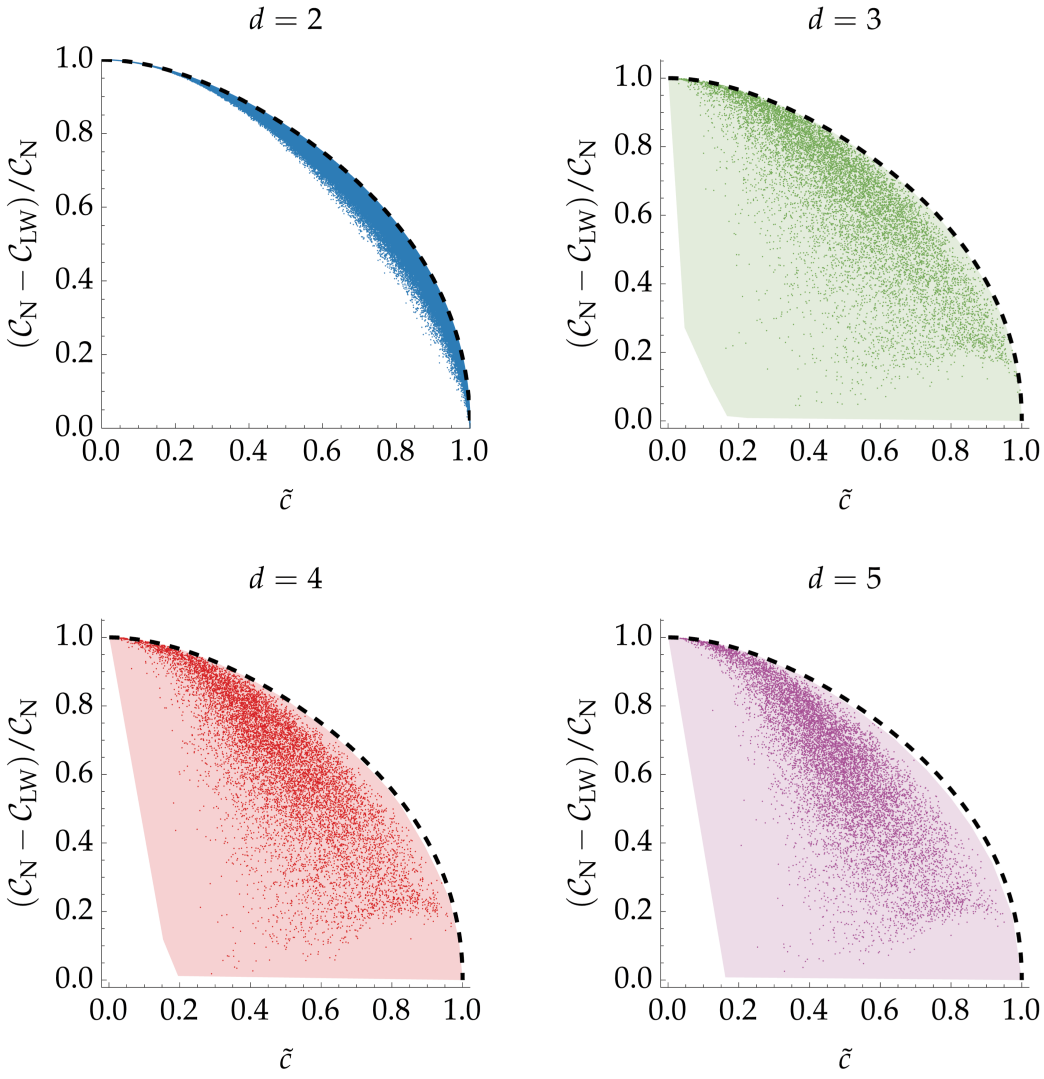

Figure 8.1: Normalised difference between the NCRB and LWB as a function of the incompatibility coefficient ($\tilde{c}$), for different Hilbert space dimensions (d). The difference is normalised by the NCRB. The presented points for dimensions 3–5 are a sample (1%) of the data collected from random estimation problems and the entirety of the data is represented by the shaded regions. The black dashed line is the upper boundary of the region of the $d = 2$ points. All points have a vertical coordinate ≤ 1, which suggests that the ordering $\mathcal{C}_{LW} \leq \mathcal{C}_N$ does not only hold in qubit systems, but higher dimensional systems too.

all data points is less than or equal to 1 suggests that the ordering $\mathcal{C}_{\text{LW}} \leq \mathcal{C}_{\text{N}}$ does hold in higher dimensions and is not a feature restricted to qubit systems.

In the subplots for dimensions 3–5 in Figure 8.1, I overlay the upper boundary of the region for qubit problems (black dashed line). Interestingly, the coloured regions each lie below this curve, suggesting that the normalised difference between the NCRB and LWB is bounded by a monotonic function of $\tilde{c}$, regardless of the Hilbert space dimension. In particular, this curve goes to 0 as $\tilde{c}$ goes to 1, which supports Conjecture 1. That there is a small gap between the coloured regions and the curve is likely due to the random sampling—due to the large number of degrees of freedom in the higher-dimensional systems, randomly generating extremal cases is unlikely.

8.2 Measurements Saturating the Lu–Wang IRTR

8.2.1 Discrepancy Between the IRTR and LWUR

Recall that the qubit rotations problem with the pure probe state described by $(r, \vartheta, \varphi) = (r, \pi/4, \pi/4)$ has estimation bounds (see Sec. 6.2.2 for details)

$$\mathcal{C}_{\text{N}} = \frac{(1+\sqrt{2})^2}{r^2} \quad \text{and} \quad \mathcal{C}_{\text{LW}} = \frac{4}{r^2}, \tag{8.1}$$

which are unequal even in the pure state case ($r = 1$), the NCRB being tighter. This gap means that the LWB is not saturated by measurements minimising the trace of the MSE. As it turns out, measurements minimising a weighted trace of the MSE do not saturate the LWUR either—this is demonstrated in Figure 8.2**a** in which, for the pure state case, I present the curve defined by the LWUR, the curve defined by the weighted NCRB (forming an uncertainty relation as described in Sec. 5.1), and points denoting the MSE of measurements saturating the weighted NCRB. It is evident that the LWUR is unattainable in this estimation problem, even though the IRTR, from which the LWUR was derived, is tight for pure states.

Furthermore, while the IRTR is tight, it is not saturated by the measurements which saturate the weighted NCRB (the measurements from Figure 8.2**a**). This is demon-

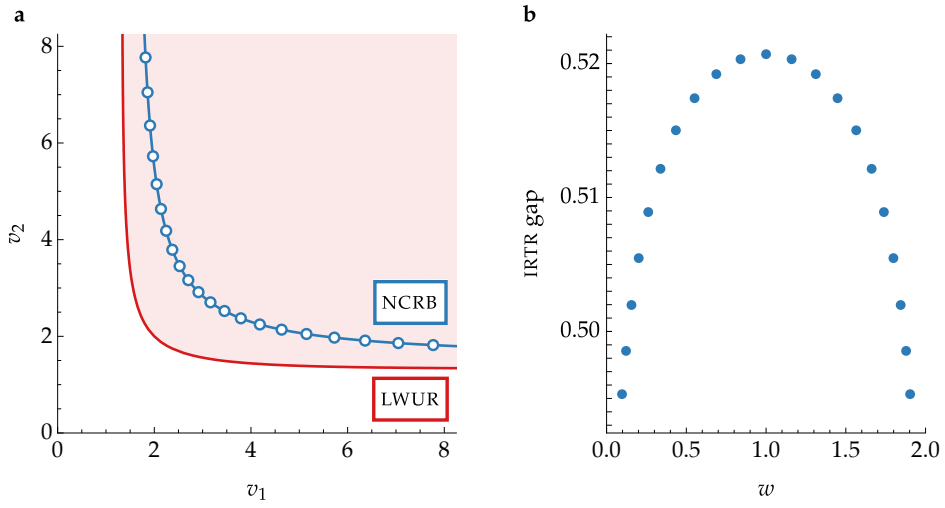

Figure 8.2: Difference between the LWUR and IRTR and the NCRB for the qubit rotations problem with $(r, \vartheta, \varphi) = (1, \pi/4, \pi/4)$. **a**, measurements (with MSE elements (v_1, v_2) marked by open circles) saturating the weighted NCRB (blue curve, as derived in Sec. 5.1). The red shaded region is the non-forbidden region of the LWUR, the boundary of which is strictly below the optimal measurements. The measurements are determined for weight matrices of the form $\mathrm{Diag}(w, 2-w)$ and the points from left to right are in order of decreasing w. **b**, gap of the IRTR inequality for the NCRB-optimal measurements in **a**. The gap is plotted as a function of the weight parameter w used for the weight matrices, and is non-zero.

strated in Figure 8.2**b**, where I plot the IRTR inequality (Eq. (3.34)) gap (the difference between the left- and right-hand sides of the inequality) of the NCRB-optimal measurements, as a function of the weighting (w). Here, w is the element of the weight matrix of the form $W = \mathrm{Diag}(w, 2-w)$. This begs the question: what measurement saturates the IRTR, and what is its performance in a MSE-based metric? I investigated the answer to this question using randomly generated measurements.

8.2.2 MSE and IRTR of Random Measurements

Recall that the IRTR is

$$\underbrace{\Delta_1^2 + \Delta_2^2 + 2\sqrt{1-\tilde{c}^2}\Delta_1\Delta_2 - \tilde{c}^2}_{G} \geq 0, \tag{8.2}$$

where $\Delta_i = \sqrt{(\mathcal{F} - F(\Pi))_{ii}/\mathcal{F}_{ii}}$ for SLDFIM $\mathcal{F}$ and CFIM $F(\Pi)$ of the measurement $\{\Pi\}$. One route to investigating the discrepancy between the IRTR and LWUR could be

to determine measurements saturating the IRTR by numerically minimising the left-hand side of Eq. (8.2) (which I will call "G"). The MSE of this measurement could then be determined using the CFIM, and compared to the LWUR inequality. This could aid in understanding the discrepancy, but it would not necessarily give a complete picture of the situation. For example, there may be multiple POVMs that saturate the IRTR, but may give different MSEs, in particular because the IRTR only takes into account the diagonal elements of the CFIM.

An alternative method, which I present here, is to use random measurements to investigate the discrepancy between the IRTR and LWUR. Rather than optimising the measurement, I generate random POVMs and look at the trends between the value of G and the trace of the MSE. As I further elaborate below, this turns out to be surprisingly insightful.

I generate random measurements by generating rank-1 operators of the form $\Pi_i = |\psi_i\rangle\langle\psi_i|$ with random vectors $|\psi_i\rangle$. Then, I normalise these operators so that they sum to the identity operator (see Refs. [89, 90] for more details). It transpired that for the purpose of this investigation it is sufficient to use 3-element POVMs.

For each randomly generated POVM, I calculate the value of G and calculate the MSE as the inverse of the CFIM, because the CCRB can always be saturated by an appropriate estimator (Sec. 3.1.5). The performance of the random measurements is presented in Figure 8.3 where each point represents a random POVM. The results demonstrate that, in this estimation problem, as a measurement becomes more optimal for the IRTR (decreasing G), its estimation variance worsens. There is a range of values of the CFIM for the same G, but as G becomes close to zero all values of trace of the inverse of the CFIM become increasingly large.

8.2.3 Understanding the Diverging MSE

To better understand the behaviour seen in Figure 8.3—where the sum of the MSE elements diverges—I further analysed the results of the random measurements. I repeated the calculation of the trace of the MSE and the value of G for 500 000 random POVMs. I then arrange the data into bins by the G values—taking vertical slices in

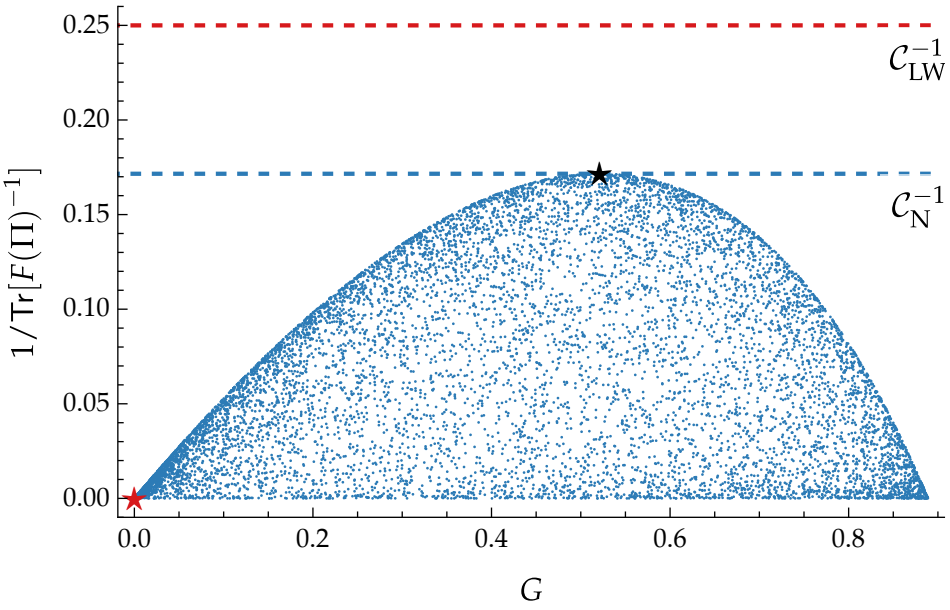

Figure 8.3: Results of random measurements for the $(r, \vartheta, \varphi) = (1, \pi/4, \pi/4)$ qubit rotations problem. The points denote the IRTR "G" value and the inverse trace of the inverse CFIM (giving the inverse of the sum of estimation variances) of random POVMs. The point with maximal vertical coordinate denotes the most informative measurement (denoted by the black star). A value of $G = 0$ corresponds to saturation of the IRTR (denoted by the red star). The blue dashed line denotes the inverse NCRB, which is saturated, and the red dashed line denotes the inverse LWB. The minimum value of G from 10 000 random trials is less than 2×10^{-7}, indicating that the IRTR is tight.

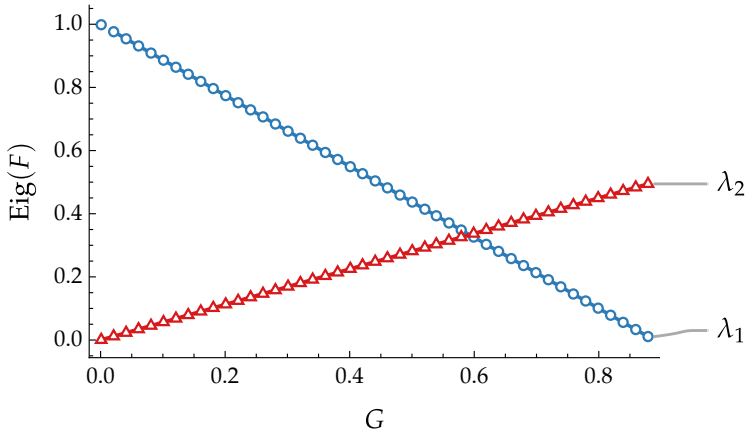

Figure 8.4: Eigenvalues (λ_i) of CFIM for "optimal" ($\mathrm{Tr}[F^{-1}]$-minimising) measurements for different values of G. As a measurement approaches saturating the IRTR ($G \to 0$), the CFIM becomes singular as λ_2 tends to zero. The estimation variances hence diverge.

Figure 8.3. Within each slice, I then find the measurement with the minimum value of $\mathrm{Tr}[F(\Pi)^{-1}]$. By finding these values, I find the measurements which form the upper boundary of the region in Figure 8.3, the "optimal" measurements for each value of G.

For the "optimal" measurement at each value of G, I calculate the eigenvalues of the CFIM, which I present in Figure 8.4. Interestingly, the eigenvalues vary linearly with G, which was confirmed via regression. Important to the trend observed in Figure 8.3 is that one of the eigenvalues (labelled λ_2) is zero when $G = 0$. This means that the CFIM is singular, and the estimation variances—the diagonal elements of the inverse of the CFIM—diverge.

To summarise, using random measurements to investigate the qubit rotations problem with a finite gap between the NCRB and LWB, I find that the IRTR is tight, but measurements saturating the IRTR have infinite MSE. This observation from Figure 8.3 is supported by the trend of the eigenvalues of the CFIM, which becomes singular as equality in the IRTR is approached. Fundamentally, this means that information regret—the difference between the classical and quantum Fisher information—at least how it is implemented in the IRTR, is not a good metric for quantifying multiparameter estimation performance. A measurement can saturate the IRTR while gaining no information about the parameters of the system, whilst a measurement gaining the most possible information (here, saturating the NCRB) does not saturate the IRTR.

Chapter 9

Discussion and Conclusion

In this thesis, I have investigated the limits that the Cramér–Rao bounds and the LWUR place on the MSE of quantum multiparameter estimation, with attainability measuring how useful an estimation limit is. The comparison between these fundamentally different limits was enabled by developing methods to convert between an uncertainty relation and a MSE-sum lower bound. As a result of applying these methods to different estimation problems, I can draw some conclusions about the general behaviour of the LWUR. I remind the reader that the LWUR is particularly interesting because, unlike the NCRB, NHCRB, and HCRB, it places a limit on the MSE without requiring the solving of an optimisation problem. As such, if it defines an attainable limit, it is a powerful tool for optimal measurement design.

In the rest of this chapter, I summarise the major results of this thesis and discuss how they relate to the questions posed in Chapter 1. I then discuss the future directions of this work.

9.1 n-Parameter Uncertainty Relations

The LWUR is restricted to two-parameter estimation because it only determines the trade-off between two MSE elements—if a third parameter is to be considered, the LWUR determines no information about its relation to the others. This motivated the determination of a framework for determining uncertainty relations that can apply to any number of parameters, based on the NCRB and NHCRB, and answers question (iii) in Chapter 1 in the affirmative.

In Sec. 5.1, I derived a tight uncertainty relation for any two-parameter estimation problem in a qubit system, which depends simply on the SLDFIM. In Chapter 7, I considered a specific three-parameter estimation problem with an arbitrary diagonal

weight matrix and obtained an analytic expression for its NHCRB. Using this weighted lower bound, I determined the surface that describes the trade-off between the three MSE elements: a three-parameter uncertainty relation. I further demonstrated that this uncertainty relation is tight by numerically determining optimal POVMs that saturate the relation.

The importance of the result of Chapter 7 is that the method for determining an uncertainty relation based on the NHCRB can be applied to estimation problems with any number of parameters. In principle, this can be done even if an analytic expression for the NHCRB cannot determined; the process for obtaining the uncertainty relation involves maximisation with respect to the weight matrix parameters, which could be performed numerically. For example, the weighted bound could be calculated for a discrete grid of weight parameters, allowing approximate points on the uncertainty relation (hyper)surface to be determined. Such a hypersurface could be visualised by plotting lower-dimensional slices of it.

I emphasise that the presented method is applicable to any Cramér–Rao bound. For example, the HCRB could be used to define an uncertainty relation that applies to collective measurements on any number of probe state copies, or the NHCRB could be replaced by a tighter n-parameter lower bound, if one is developed in the future.

The estimation uncertainty relations have practical use for experimental design. For example, the information about the MSE trade-off may aid optimal measurement design in situations where the importance of minimising each parameter's MSE is variable.

9.2 Conditional Attainability of the LWUR

The general result in Sec. 5.4 and the explicit examples in Chapter 6 answer questions (i) and (ii) posed in Chapter 1 as: "sometimes". That is, the LWUR in some cases defines the same estimation limit as the NCRB (for example, when $\tilde{c} = 1$ and the SLD-FIM is diagonal in qubit problems), and it is sometimes attainable. This suggests a more refined question: under what conditions does the LWUR and NCRB define the same estimation limit, and why? In the following, I summarise the evidence I found

for answers to this question. In particular, investigating the "*why?*" suggests avenues for future work.

The LWUR depends on two quantities: the incompatibility coefficient ($\tilde{c}$) and the SLDFIM. It is sensible that conditions for the attainability of the LWUR should be related to these. Indeed, the condition determined in Sec. 5.4 concerns only these quantities. One route to determining the relationship between LWUR attainability and these quantities is to compare the LWB and NCRB and observe when they are equal. Based on the observations in Chapters 5, 6, and 7, I proposed that if $\tilde{c} = 1$, then the LWB and NCRB are equal, for any two-parameter estimation problem (Conjecture 1).

It should be noticed that I omit the diagonal SLDFIM condition from Conjecture (as opposed to the condition determined in Sec. 5.4). This choice is supported by the results in Figure 8.1, which suggests that there is a relationship between the value of $\tilde{c}$ and whether the SLDFIM is diagonal. In Sec. 5.4, I found that for qubits, when $\tilde{c} = 1$, the NCRB and LWB are equal if and only if the SLDFIM is diagonal. However, the numerical results found that whenever $\tilde{c} = 1$, the NCRB and LWB are equal, which suggests that $\tilde{c} = 1$ implies that the SLDFIM is diagonal. This condition should be explored further in future work and may yield insights about the physical meaning of the incompatibility coefficient.

Proving Conjecture 1 is beyond the scope of this thesis, but there is extensive evidence supporting it. In Chapter 8, I presented the results of random estimation problems. These results suggest that the normalised difference between the NCRB and LWB is bounded by a monotonic function of $\tilde{c}$, which in particular goes to zero at $\tilde{c} = 1$. The consequence of this finding is that if an estimation problem satisfies $\tilde{c} = 1$, then the precision limit has a simple expression (Sec. 5.3), irrespective of the Hilbert space dimension. For the qubit case, the results of the random problems also suggest that the LWUR is tight if and only if $\tilde{c} = 1$, but this may not be the case for higher dimensions.

A common theme in the explicit examples in Chapter 6 where the LWUR is not attainable is that the SLDFIM is not diagonal. This is not entirely surprising because the LWUR does not explicitly depend on the off-diagonal SLDFIM terms. In the extreme case, the SLDFIM can be singular, meaning that the SLDCRB, NCRB, and HCRB are infinite, while the LWB is finite because the diagonal SLDFIM elements are finite.

The ignorance of the LWUR to the off-diagonal elements of the SLDFIM appears to be detrimental to its general attainability. However, this problem may originate from more fundamental problems with the IRTR. In Sec. 8.2, I found that while the IRTR is tight for pure states, this need not be useful for optimal parameter estimation. In a pure state problem where the LWUR is not attainable, the optimal measurement (according to the physically relevant MSE-trace measure) does not saturate the IRTR. Instead, the measurements that do saturate the IRTR have singular CFIM, and hence gain zero information about the parameters. Again, this is likely due to the IRTR only depending on the diagonal elements of the information regret matrix. In the next section, I suggest modifications to the IRTR and LWUR which may resolve this problem.

9.3 Future Directions

9.3.1 Improving the IRTR

The optimality of the LWUR in some cases and the sub-optimality in others suggests that there may be room for its improvement. Indeed, the path taken by Lu and Wang to apply Ozawa's uncertainty relation to multiparameter estimation need not be the only possible route, and another may lead to different results.

The above-discussed importance of the off-diagonal elements of the SLDFIM motivates me to propose the following alternative: instead of using the diagonal elements of the information-regret matrix and the SLDFIM, the IRTR could be reformulated to use the eigenvalues of both matrices. In the case that the matrices are diagonal, there is no change (because the eigenvalues of a diagonal matrix are the diagonal elements), but when they are non-diagonal, the off-diagonal elements are accounted for in the eigenvalues. This seems to be a sensible choice because it is the positivity of the eigenvalues of the information regret ($R = \mathcal{F} - F$) that relates to the bounding property $F \preceq \mathcal{F}$, while the positivity of the diagonal elements is a weaker condition.

This proposed change should retain the optimality of Lu and Wang's IRTR in the cases where it is optimal, and may improve the situation in the other cases. In particular, the closer relation to the CFIM upper bound may provide a better correspondence

between an information-regret trade-off relation and MSE-optimal measurements.

9.3.2 Infinite-Dimensional Systems

A potential shortcoming of this study is the sole consideration of low-dimensional quantum systems. While these are the simplest to model and investigate, and they provide a wealth of interesting phenomena, they do not necessarily expose the entire picture of quantum multiparameter estimation limits. Indeed, important quantities such as position, momentum, and electromagnetic field amplitudes and phases are continuous and thus described by infinite-dimensional Hilbert spaces.

It is not clear whether the results of this thesis would be different if considering higher-dimensional and infinite-dimensional quantum systems. For example, the results of random estimation problems in dimensions 3, 4 and 5 were qualitatively very similar (see Fig. 8.1), so exploring higher dimensions may not provide further insight. Nevertheless, higher-dimensional systems should be considered in future for a complete investigation.

9.4 Conclusions

In this thesis, I derived analytic expressions for the LWB in limiting cases of the incompatibility coefficient which are known to be physically relevant. I further identified underlying flaws in the IRTR and LWUR which prevent it from always defining an attainable trade-off relation for quantum multiparameter estimation. I proposed a fix for these issues and also demonstrated that the issues can nevertheless be circumvented by defining an uncertainty relation based on Cramér–Rao bounds. I further demonstrated that this method can be generalised for determining estimation uncertainty relations for any number of parameters. Such uncertainty relations explicitly define the trade-off of the MSE due to operator incompatibility—a fundamental part of quantum mechanics—and have practical use in optimal experimental design.

Appendix A

Right Logarithmic Derivative and HCRB

The right logarithmic derivative (RLD) is a non-commutative analog of the logarithmic derivative—an alternative to the symmetric logarithmic derivative. The RLD operator corresponding to the parameter θ_i is defined implicitly by $\frac{\partial \rho_\theta}{\partial \theta_i} = \rho_{\boldsymbol{\theta}} \tilde{L}_i$, and is not necessarily self-adjoint. The right logarithmic derivative Fisher information matrix has elements $\tilde{\mathcal{F}}_{ij} = \frac{1}{2} \operatorname{Tr}[\rho_{\boldsymbol{\theta}} \{\tilde{L}_i, \tilde{L}_j\}]$, from which the right logarithmic derivative Cramér–Rao bound is defined

$$\mathcal{C}_{\mathrm{RLD}} = \mathsf{Tr}[W \operatorname{Re}[\tilde{\mathcal{F}}^{-1}]] + \mathsf{TrAbs}[W \operatorname{Im}[\tilde{\mathcal{F}}^{-1}]]. \tag{A.1}$$

Suzuki [70] proved that for any two-parameter qubit estimation problem with unitarily encoded parameters, the HCRB has the explicit expression

$$\mathcal{C}_{\mathrm{H}} = \begin{cases} \mathcal{C}_{\mathrm{RLD}}, & \text{if } \mathcal{C}_{\mathrm{RLD}} \geq \frac{1}{2} \left(\mathcal{C}_{\mathrm{Z}} + \mathcal{C}_{\mathrm{SLD}} \right), \\ \mathcal{C}_{\mathrm{RLD}} + S, & \text{otherwise,} \end{cases} \tag{A.2}$$

where the quantity $\mathcal{C}_{\mathrm{Z}}$ is defined as

$$\mathcal{C}_{\mathrm{Z}} = \mathsf{Tr}[W \operatorname{Re}[Z]] + \mathsf{TrAbs}[W \operatorname{Im}[Z]], \text{ where } Z_{ij} = \operatorname{Tr}[\rho_{\boldsymbol{\theta}} L^i L^j] \text{ and } L^i = \sum_k [\mathcal{F}^{-1}]_{ik} L_k, \tag{A.3}$$

and

$$S = \frac{\left(\frac{1}{2} \left(\mathcal{C}_{\mathrm{Z}} + \mathcal{C}_{\mathrm{SLD}} \right) - \mathcal{C}_{\mathrm{RLD}} \right)^2}{\mathcal{C}_{\mathrm{Z}} - \mathcal{C}_{\mathrm{RLD}}}. \tag{A.4}$$

Here, L^i is a linear combination of the SLD operators L_k, weighted by the elements of the inverse SLDFIM $\mathcal{F}$.

Appendix B

Calculation of LWB for Qubit Rotations

The values relevant to the LWUR for this probe state ($\vartheta = \varphi = \pi/4$, arbirary r) are

$$\mathcal{F} = r^2 \begin{pmatrix} \frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} \end{pmatrix}, \quad \tilde{c} = \sqrt{\frac{8}{9}}. \tag{B.1}$$

For convenience in the following, I define $\kappa = 1/\mathcal{F}_{xx} = 1/\mathcal{F}_{yy} = 4/(3r^2)$.

I use the Karush–Kuhn–Tucker (KKT) conditions, which were discussed in Sec. 4.1.2 to determine the LWB. In this problem, the KKT conditions are simpler than the Lagrange dual function approach due to the non-trivial dependence of the constraints on v_x, v_y. The minimisation problem can be stated as

$$\mathcal{C}_{\text{LW}} := \min\Bigg\{ v_x + v_y \;\Bigg|\; f_1(v_x, v_y) = \frac{\kappa}{v_x} + \frac{\kappa}{v_y} - \frac{2}{3}\sqrt{\left(1 - \frac{\kappa}{v_x}\right)\left(1 - \frac{\kappa}{v_y}\right)} - \frac{10}{9} \leq 0, \\ f_2(v_x, v_y) = \kappa - v_x \leq 0,\ f_3(v_x, v_y) = \kappa - v_y \leq 0 \Bigg\} \tag{B.2}$$

The KKT conditions (following numbering in Eq. (4.5)) are then

$$\begin{aligned} &\text{(I; III)} && f_1(v_x^\star, v_y^\star) \leq 0,\ \kappa - v_x^\star \leq 0,\ \kappa - v_y^\star \leq 0; \quad \lambda_1^\star \geq 0,\ \lambda_2^\star \geq 0,\ \lambda_3^\star \geq 0, \\ &\text{(IV)} && \lambda_1^\star f_1(v_x^\star, v_y^\star) = 0,\ \lambda_2^\star(\kappa - v_x^\star) = 0,\ \lambda_3^\star(\kappa - v_y^\star) = 0, \\ &\text{(V)} && \boldsymbol{\nabla} f_0(v_x^\star, v_y^\star) + \lambda_1^\star \boldsymbol{\nabla} f_1(v_x^\star, v_y^\star) + \lambda_2^\star \boldsymbol{\nabla}(\kappa - v_x^\star) + \lambda_3^\star \boldsymbol{\nabla}\left(\kappa - v_y^\star\right) = 0 \end{aligned} \tag{B.3}$$

The minimisation problem is convex and strong duality holds, so the KKT conditions are necessary and sufficient for the solution. That is, it is sufficient to find *a* solution to the above equations. First I consider the complementary slackness conditions (IV).

Suppose $\lambda_2^\star \neq 0$ and $\lambda_3^\star \neq 0$. Then $v_x^\star = v_y^\star = \kappa$, but $f_1(\kappa, \kappa) = 2 - \frac{10}{9} > 0$ in contradiction with (I). Instead, suppose that one of $\lambda_2^\star, \lambda_3^\star$ is zero while the other is non-zero, say $\lambda_2^\star \neq 0$. Then $v_x^\star = \kappa$ and $f_1(\kappa, v_y^\star) = \frac{\kappa}{v_y^\star} - \frac{1}{9}$, and enforcing $f_1 \leq 0$ is equivalent to $v_y^\star \geq 9\kappa$. This implies that $v_x^\star + v_y^\star \geq 10\kappa$. However, as I will show in the following, enforcing $\lambda_2^\star = \lambda_3^\star = 0$ leads to a smaller (correct minimum) value of the sum.

With $\lambda_2^\star = \lambda_3^\star = 0$, condition (V) can be expanded and simplified as (dropping the '$\star$' for brevity)

$$\begin{pmatrix} 1 - \frac{\lambda_1 \kappa}{v_x^2}\left(1 + \frac{1}{3}\sqrt{\frac{1-\frac{\kappa}{v_y}}{1-\frac{\kappa}{v_x}}}\right) \\ 1 - \frac{\lambda_1 \kappa}{v_y^2}\left(1 + \frac{1}{3}\sqrt{\frac{1-\frac{\kappa}{v_x}}{1-\frac{\kappa}{v_y}}}\right) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \frac{v_x^2}{1 + \frac{1}{3}\sqrt{\frac{1-\frac{\kappa}{v_y}}{1-\frac{\kappa}{v_x}}}} = \frac{v_y^2}{1 + \frac{1}{3}\sqrt{\frac{1-\frac{\kappa}{v_x}}{1-\frac{\kappa}{v_y}}}}. \tag{B.4}$$

This has a solution $v_x^\star = v_y^\star \equiv v$. Finally, this solution must also satisfy $f_1(v, v) = 0$ because $\lambda_1^\star \neq 0$, as seen in Eq. (B.4). Solving:

$$0 = f_1(v, v) = \frac{2\kappa}{v} - \frac{2}{3}\left(1 - \frac{\kappa}{v} - \frac{10}{9}\right) \tag{B.5}$$

$$\implies v = \frac{3\kappa}{2} = \frac{2}{r^2}, \tag{B.6}$$

which gives the LWB as $\mathcal{C}_{\text{LW}} = \frac{4}{r^2}$.

Higher-Dimensional Rotations

Consider the d-dimensional probe state and derivatives

$$\rho = \text{Diag}\left(\frac{1}{d-1/2}, \ldots, \frac{1}{d-1/2}, 1 - \frac{d-1}{d-1/2}\right), \quad \partial_j \rho = \frac{i}{2}[\rho, M_j], \tag{C.1}$$

where

$$M_1 = \begin{pmatrix} & & 1 \\ & & \\ 1 & & \end{pmatrix} \text{ and } M_2 = \begin{pmatrix} & & -i \\ & & \\ i & & \end{pmatrix}. \tag{C.2}$$

Here, and in the following, blank entries in the $d \times d$ matrices are zero. It is straightforward to verify (by computing the commutator) that the derivatives are

$$\partial_1 \rho = \frac{-1}{4d-2} M_2 \text{ and } \partial_2 \rho = \frac{1}{4d-2} M_1, \tag{C.3}$$

and that the SLD operators are

$$L_1 = -\frac{1}{3} M_2 \text{ and } L_2 = \frac{1}{3} M_1. \tag{C.4}$$

Then, the SLDFIM is

$$\mathcal{F} = \begin{pmatrix} \frac{1}{6d-3} & 0 \\ 0 & \frac{1}{6d-3} \end{pmatrix}. \tag{C.5}$$

It is then straightforward to verify that $\tilde{c} = 1$, so the LWB is given by

$$\mathcal{C}_{\text{LW}} = \left(\frac{1}{\sqrt{\mathcal{F}_{11}}} + \frac{1}{\sqrt{\mathcal{F}_{22}}}\right)^2 = 12(2d-1). \tag{C.6}$$

Inspired by the symmetry of the above operators, an ansatz for X matrices is chosen as

$$X_i = \begin{pmatrix} & a_i \\ b_i & \end{pmatrix}, \tag{C.7}$$

for which the locally unbiased conditions enforce that

$$a_1 = i(2d-1), \quad b_1 = -i(2d-1), \quad a_2 = 2d-1, \quad b_2 = 2d-1. \tag{C.8}$$

Substituting these X matrices into the NCRB function (argument of min in Eq. (3.30)) gives

$$\mathcal{C}_{\mathrm{N}} \leq \mathsf{Tr}[Z(X)] + \mathrm{TrAbs}[\rho[X_1, X_2]] = 12(2d-1), \tag{C.9}$$

which is equal to the LWB. I verified, using the numerical solution to the semidefinite program for the NCRB, that these X matrices are optimisers, and hence there is equality in Eq. (C.9). It follows that for this estimation problem, $\tilde{c} = 1$ and the NCRB and LWB are equal.

Bibliography

[1] T. Quinn and J. Kovalevsky, "The development of modern metrology and its role today", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **363**, 2307–2327 (2005) (cit. on p. 1).

[2] M. Zhan and X. Xie, "Precision measurement physics: physics that precision matters", National Science Review **7**, 1795–1795 (2020) (cit. on p. 1).

[3] S. G. Karshenboim, "Precision physics of simple atoms: QED tests, nuclear structure and fundamental constants", Physics Reports **422**, 1–63 (2005) (cit. on p. 1).

[4] V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology", Nature Photonics **5**, 222–229 (2011) (cit. on p. 1).

[5] J. Aasi, J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, C. Adams, T. Adams, et al., "Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light", Nature Photonics **7**, 613–619 (2013) (cit. on p. 1).

[6] M. Tse, H. Yu, N. Kijbunchoo, A. Fernandez-Galiana, P. Dupej, L. Barsotti, C. D. Blair, D. D. Brown, et al., "Quantum-Enhanced Advanced LIGO Detectors in the Era of Gravitational-Wave Astronomy", Physical Review Letters **123**, 231107 (2019) (cit. on p. 1).

[7] M. Tsang, R. Nair, and X.-M. Lu, "Quantum Theory of Superresolution for Two Incoherent Optical Point Sources", Physical Review X **6**, 031033 (2016) (cit. on pp. 1, 26).

[8] C. A. Casacio, L. S. Madsen, A. Terrasson, M. Waleed, K. Barnscheidt, B. Hage, M. A. Taylor, and W. P. Bowen, "Quantum-enhanced nonlinear microscopy", Nature **594**, 201–206 (2021) (cit. on p. 1).

[9] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, "Quantum Fisher information matrix and multiparameter estimation", Journal of Physics A: Mathematical and Theoretical **53**, 023001 (2019) (cit. on pp. 2, 23).

[10] F. Albarelli, M. Barbieri, M. G. Genoni, and I. Gianani, "A perspective on multiparameter quantum metrology: From theoretical tools to applications in quantum imaging", Physics Letters A **384**, 126311 (2020) (cit. on pp. 2, 23).

[11] R. Demkowicz-Dobrzański, W. Górecki, and M. Guţă, "Multi-parameter estimation beyond quantum Fisher information", Journal of Physics A: Mathematical and Theoretical **53**, 363001 (2020) (cit. on pp. 2, 23).

[12] J. S. Sidhu and P. Kok, "Geometric perspective on quantum parameter estimation", AVS Quantum Science **2**, 014701 (2020) (cit. on pp. 2, 23).

[13] L. O. Conlon, T. Vogl, C. D. Marciniak, I. Pogorelov, S. K. Yung, F. Eilenberger, D. W. Berry, F. S. Santana, et al., "Approaching optimal entangling collective measurements on quantum computing platforms", Nature Physics **19**, 351–357 (2023) (cit. on pp. 2, 10, 41, 57).

[14] A. S. Holevo, "Noncommutative analogues of the Cramér-Rao inequality in the quantum measurement theory", in Proceedings of the Third Japan — USSR Symposium on Probability Theory, edited by G. Maruyama and J. V. Prokhorov, Lecture Notes in Mathematics (1976), pp. 194–222 (cit. on pp. 2, 28).

[15] M. Hayashi, *Asymptotic Theory of Quantum Statistical Inference: Selected Papers* (World Scientific, 2005) (cit. on p. 2).

[16] X.-M. Lu and X. Wang, "Incorporating Heisenberg's Uncertainty Principle into Quantum Multiparameter Estimation", Physical Review Letters **126**, 120503 (2021) (cit. on pp. 3, 32, 43).

[17] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, 10th Anniversary Ed (Cambridge University Press, Cambridge, 2010), 676 pp. (cit. on pp. 4–6, 8, 9, 49).

[18] A. Babazadeh, M. Erhard, F. Wang, M. Malik, R. Nouroozi, M. Krenn, and A. Zeilinger, "High-Dimensional Single-Photon Quantum Gates: Concepts and Experiments", Physical Review Letters **119**, 180510 (2017) (cit. on p. 5).

[19] Y. Chi, J. Huang, Z. Zhang, J. Mao, Z. Zhou, X. Chen, C. Zhai, J. Bao, et al., "A programmable qudit-based quantum processor", Nature Communications **13**, 1166 (2022) (cit. on p. 5).

[20] M. Kacprowicz, R. Demkowicz-Dobrzański, W. Wasilewski, K. Banaszek, and I. A. Walmsley, "Experimental quantum-enhanced estimation of a lossy phase shift", Nature Photonics **4**, 357–360 (2010) (cit. on p. 5).

[21] W. Wang, Y. Wu, Y. Ma, W. Cai, L. Hu, X. Mu, Y. Xu, Z.-J. Chen, et al., "Heisenberg-limited single-mode quantum metrology in a superconducting circuit", Nature Communications **10**, 1–6 (2019) (cit. on p. 5).

[22] K. C. McCormick, J. Keller, S. C. Burd, D. J. Wineland, A. C. Wilson, and D. Leibfried, "Quantum-enhanced sensing of a single-ion mechanical oscillator", Nature **572**, 86–90 (2019) (cit. on p. 5).

[23] D. J. Griffiths, *Introduction to quantum mechanics*, Second edition [2017 edition] (Cambridge University Press, Cambridge, 2017), 468 pp. (cit. on p. 8).

[24] A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, 1st ed., Vol. 1 (Edizioni della Normale Pisa, 2011) (cit. on pp. 8, 9, 28, 30).

[25] M. Neumark, "Spectral functions of a symmetric operator", Izvestiya Akademii Nauk Seriya Matematicheskaya **4**, 277–318 (1940) (cit. on p. 9).

[26] P. Kurzyński, "Quantum Walk as a Generalized Measuring Device", Physical Review Letters **110**, 10.1103/PhysRevLett.110.200404 (2013) (cit. on pp. 10, 12, 13).

[27] Z. Bian, "Realization of Single-Qubit Positive-Operator-Valued Measurement via a One-Dimensional Photonic Quantum Walk", Physical Review Letters **114**, 10.1103/PhysRevLett.114.203602 (2015) (cit. on pp. 10, 12).

[28] Y.-y. Zhao, N.-k. Yu, P. Kurzyński, G.-y. Xiang, C.-F. Li, and G.-C. Guo, "Experimental realization of generalized qubit measurements based on quantum walks", Physical Review A **91**, 042101 (2015) (cit. on pp. 10, 12).

[29] Z. Li, H. Zhang, and H. Zhu, "Implementation of generalized measurements on a qudit via quantum walks", Physical Review A **99**, 062342 (2019) (cit. on pp. 10, 12).

[30] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, "Simulating Positive-Operator-Valued Measures with Projective Measurements", Physical Review Letters **119**, 190501 (2017) (cit. on p. 10).

[31] M. Oszmaniec, F. B. Maciejewski, and Z. Puchała, "Simulating all quantum measurements using only projective measurements and postselection", Physical Review A **100**, 012351 (2019) (cit. on p. 10).

[32] T. Singal, F. B. Maciejewski, and M. Oszmaniec, "Implementation of quantum measurements using classical resources and only a single ancillary qubit", npj Quantum Information **8**, 1–7 (2022) (cit. on p. 10).

[33] H. Chen, Y. Chen, and H. Yuan, "Information Geometry under Hierarchical Quantum Measurement", Physical Review Letters **128**, 250502 (2022) (cit. on p. 10).

[34] H. Chen, Y. Chen, and H. Yuan, "Incompatibility measures in multiparameter quantum estimation under hierarchical quantum measurements", Physical Review A **105**, 062442 (2022) (cit. on p. 10).

[35] A. Franzen, *ComponentLibrary: a free vector graphics library for optics*, (2009) http://www.gwoptics.org/ComponentLibrary/ (cit. on p. 12).

[36] J. Kempe, "Quantum random walks: An introductory overview", Contemporary Physics **44**, 307–327 (2003) (cit. on p. 12).

[37] S. E. Venegas-Andraca, "Quantum walks: a comprehensive review", Quantum Information Processing **11**, 1015–1106 (2012) (cit. on p. 12).

[38] Z. Hou, J.-F. Tang, J. Shang, H. Zhu, J. Li, Y. Yuan, K.-D. Wu, G.-Y. Xiang, et al., "Deterministic realization of collective measurements via photonic quantum walks", Nature Communications **9**, 1414 (2018) (cit. on pp. 12, 13).

[39] W. Heisenberg, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik", Zeitschrift für Physik **43**, 172–198 (1927) (cit. on p. 13).

[40] E. H. Kennard, "Zur Quantenmechanik einfacher Bewegungstypen", Zeitschrift für Physik **44**, 326–352 (1927) (cit. on p. 13).

[41] H. Weyl, "Quantenmechanik und Gruppentheorie", Zeitschrift für Physik **46**, 1–46 (1927) (cit. on p. 13).

[42] L. E. Ballentine, "The Statistical Interpretation of Quantum Mechanics", Reviews of Modern Physics **42**, 358–381 (1970) (cit. on p. 14).

[43] H. P. Robertson, "The Uncertainty Principle", Physical Review **34**, 163–164 (1929) (cit. on p. 14).

[44] E. Schrödinger, "Zum heisenbergschen unschärfeprinzip", Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse **14**, 296–303 (1930) (cit. on p. 14).

[45] M. Ozawa, "Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement", Physical Review A **67**, 042105 (2003) (cit. on pp. 14, 15).

[46] M. Ozawa, "Uncertainty relations for joint measurements of noncommuting observables", Physics Letters A **320**, 367–374 (2004) (cit. on p. 15).

[47] C. Branciard, "Error-tradeoff and error-disturbance relations for incompatible quantum measurements", Proceedings of the National Academy of Sciences **110**, 6742–6747 (2013) (cit. on p. 15).

[48] M. Ozawa, *Error-disturbance relations in mixed states*, arXiv:1404.3388 (arXiv, 2014), arXiv:1404.3388[quant-ph] (cit. on pp. 15, 32).

[49] R. A. Fisher, "Theory of Statistical Estimation", Mathematical Proceedings of the Cambridge Philosophical Society **22**, 700–725 (1925) (cit. on pp. 16, 20, 22).

[50] J. A. Rice, *Mathematical statistics and data analysis*, 2nd ed (Duxbury Press, Belmont, CA, 1995), 602 pp. (cit. on p. 19).

[51] R. A. Fisher, "On the mathematical foundations of theoretical statistics", Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character **222**, 309–368 (1922) (cit. on p. 21).

[52] Harald Cramér, *Mathematical Methods Of Statistics* (Princeton University Press, 1946) (cit. on p. 21).

[53] C. R. Rao, "Information and the Accuracy Attainable in the Estimation of Statistical Parameters", in *Breakthroughs in Statistics: Foundations and Basic Theory*, edited by S. Kotz and N. L. Johnson, Springer Series in Statistics (Springer, New York, NY, 1992), pp. 235–247 (cit. on p. 21).

[54] C. R. Rao, "Minimum variance and the estimation of several parameters", Mathematical Proceedings of the Cambridge Philosophical Society **43**, 280–283 (1947) (cit. on p. 21).

[55] A. Fujiwara and H. Nagaoka, "An estimation theoretical characterization of coherent states", Journal of Mathematical Physics **40**, 4227–4239 (1999) (cit. on pp. 22, 41).

[56] M. G. Genoni, P. Giorda, and M. G. A. Paris, "Optimal estimation of entanglement", Physical Review A **78**, 032303 (2008) (cit. on p. 23).

[57] C. W. Helstrom, "Minimum mean-squared error of estimates in quantum statistics", Physics Letters A **25**, 101–102 (1967) (cit. on p. 25).

[58] C. Helstrom, "The minimum variance of estimates in quantum signal detection", IEEE Transactions on Information Theory **14**, 234–242 (1968) (cit. on p. 25).

[59] S. L. Braunstein and C. M. Caves, "Statistical distance and the geometry of quantum states", Physical Review Letters **72**, 3439–3443 (1994) (cit. on pp. 25, 26).

[60] C. W. Helstrom, *Quantum detection and estimation theory*, Mathematics in science and engineering v. 123 (Academic Press, New York, 1976), 309 pp. (cit. on p. 26).

[61] M. G. A. Paris, "Quantum estimation for quantum technology", International Journal of Quantum Information **07**, 125–137 (2009) (cit. on p. 26).

[62] T. Baumgratz and A. Datta, "Quantum Enhanced Estimation of a Multidimensional Field", Physical Review Letters **116**, 030801 (2016) (cit. on p. 26).

[63] P. J. D. Crowley, A. Datta, M. Barbieri, and I. A. Walmsley, "Tradeoff in simultaneous quantum-limited phase and loss estimation in interferometry", Physical Review A **89**, 023845 (2014) (cit. on p. 26).

[64] S. P. Boyd and L. Vandenberghe, *Convex optimization* (Cambridge University Press, New York, United States of America, 2004), 716 pp. (cit. on pp. 27, 35, 36, 38, 39).

[65] M. Zhang, H.-M. Yu, H. Yuan, X. Wang, R. Demkowicz-Dobrzański, and J. Liu, "QuanEstimation: An open-source toolkit for quantum parameter estimation", Physical Review Research **4**, 043057 (2022) (cit. on pp. 27, 65).

[66] J. Lofberg, "YALMIP : a toolbox for modeling and optimization in MATLAB", in 2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508) (2004), pp. 284–289 (cit. on p. 27).

[67] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 9.3.* (2021) (cit. on p. 27).

[68] H. Nagaoka, "A New Approach to Cramér-Rao Bounds for Quantum State Estimation", in *Asymptotic Theory of Quantum Statistical Inference* (World Scientific, 2005), pp. 100–112 (cit. on pp. 27–29, 31, 43).

[69] F. Albarelli, J. F. Friel, and A. Datta, "Evaluating the Holevo Cramér-Rao Bound for Multiparameter Quantum Metrology", Physical Review Letters **123**, 200503 (2019) (cit. on p. 29).

[70] J. Suzuki, "Explicit formula for the Holevo bound for two-parameter qubit-state estimation problem", Journal of Mathematical Physics **57**, 042201 (2016) (cit. on pp. 29, 81).

[71] H. Nagaoka, "A Generalization of the Simultaneous Diagonalization of Hermitian Matrices and its Relation to Quantum Estimation Theory", in *Asymptotic Theory of Quantum Statistical Inference* (World Scientific, 2005), pp. 133–149 (cit. on pp. 29, 31).

[72] L. O. Conlon, J. Suzuki, P. K. Lam, and S. M. Assad, "Efficient computation of the Nagaoka–Hayashi bound for multiparameter estimation with separable measurements", npj Quantum Information **7**, 1–8 (2021) (cit. on p. 30).

[73] M. Hayashi and Y. Ouyang, *Tight Cramér-Rao type bounds for multiparameter quantum metrology through conic programming*, 2022, arXiv:2209.05218[quant-ph] (cit. on p. 31).

[74] K. Matsumoto, "A new approach to the Cramér-Rao-type bound of the pure-state model", Journal of Physics A: Mathematical and General **35**, 3111 (2002) (cit. on p. 31).

[75] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, "Compatibility in multiparameter quantum metrology", Physical Review A **94**, 052108 (2016) (cit. on p. 31).

[76] J. Kahn and M. Guţă, "Local Asymptotic Normality for Finite Dimensional Quantum Systems", Communications in Mathematical Physics **289**, 597–652 (2009) (cit. on p. 31).

[77] K. Yamagata, A. Fujiwara, and R. D. Gill, "Quantum local asymptotic normality based on a new quantum likelihood ratio", The Annals of Statistics **41**, 2197–2217 (2013) (cit. on p. 31).

[78] Y. Yang, G. Chiribella, and M. Hayashi, "Attaining the Ultimate Precision Limit in Quantum State Estimation", Communications in Mathematical Physics **368**, 223–293 (2019) (cit. on p. 31).

[79] L. O. Conlon, J. Suzuki, P. K. Lam, and S. M. Assad, *The gap persistence theorem for quantum multiparameter estimation*, 2022, arXiv:2208.07386[math-ph, physics:quant-ph] (cit. on p. 31).

[80] J. Shao and X.-M. Lu, "Performance-tradeoff relation for locating two incoherent optical point sources", Physical Review A **105**, 062416 (2022) (cit. on p. 32).

[81] Y. Shi and X.-M. Lu, "Joint optimal measurement for locating two incoherent optical point sources near the Rayleigh distance", Communications in Theoretical Physics **75**, 045102 (2023) (cit. on pp. 32, 43).

[82] I. Kull, P. A. Guérin, and F. Verstraete, "Uncertainty and trade-offs in quantum multiparameter estimation", Journal of Physics A: Mathematical and Theoretical **53**, 244001 (2020) (cit. on pp. 41, 42).

[83] H.-H. Qin, S.-M. Fei, and X. Li-Jost, "Multi-observable Uncertainty Relations in Product Form of Variances", Scientific Reports **6**, 31192 (2016) (cit. on p. 42).

[84] M. D. Vidrighin, G. Donati, M. G. Genoni, X.-M. Jin, W. S. Kolthammer, M. S. Kim, A. Datta, M. Barbieri, et al., "Joint estimation of phase and phase diffusion for quantum metrology", Nature Communications **5**, 3532 (2014) (cit. on p. 43).

[85] S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: machine learning, dynamical systems, and control* (Cambridge University Press, Cambridge, 2019) (cit. on p. 59).

[86] W. Pfeifer, *The Lie algebras su(N), an introduction* (Birkhäuser, Basel, 2003) (cit. on p. 60).

[87] R. A. Bertlmann and P. Krammer, "Bloch vectors for qudits", Journal of Physics A: Mathematical and Theoretical **41**, 235303 (2008) (cit. on p. 61).

[88] I. P. Mendaš, "The classification of three-parameter density matrices for a qutrit", Journal of Physics A: Mathematical and General **39**, 11313 (2006) (cit. on p. 62).

[89] N. Johnston, A. Cosentino, and V. Russo, *{QETLAB}: A {MATLAB} toolbox for quantum entanglement, version 0.9.* Version v0.9, 2016 (cit. on pp. 69, 73).

[90] T. Heinosaari, M. A. Jivulescu, and I. Nechita, "Random positive operator valued measures", Journal of Mathematical Physics **61**, 042202 (2020) (cit. on p. 73).