Journal of Physics A: Mathematical and Theoretical

PAPER • **OPEN ACCESS**

Uncertainty and trade-offs in quantum multiparameter estimation

To cite this article: Ilya Kull *et al* 2020 *J. Phys. A: Math. Theor.* **53** 244001

OPEN ACCESS
IOP Publishing
Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **53** (2020) 244001 (27pp)
https://doi.org/10.1088/1751-8121/ab7f67

Uncertainty and trade-offs in quantum multiparameter estimation

Ilya Kull[1,2,4], **Philippe Allard Guérin**[2] and **Frank Verstraete**[3]

[1] Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria
[2] Institute for Quantum Optics and Quantum Information (IQOQI), Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria
[3] Department of Physics and Astronomy, Ghent University, Krijgslaan 281, 9000 Gent, Belgium

E-mail: Ilya.kull@oeaw.ac.at

Received 23 December 2019, revised 10 March 2020
Accepted for publication 12 March 2020
Published 26 May 2020

Abstract

Uncertainty relations in quantum mechanics express bounds on our ability to simultaneously obtain knowledge about expectation values of non-commuting observables of a quantum system. They quantify trade-offs in accuracy between complementary pieces of information about the system. In quantum multiparameter estimation, such trade-offs occur for the precision achievable for different parameters characterizing a density matrix: an uncertainty relation emerges between the achievable variances of the different estimators. This is in contrast to classical multiparameter estimation, where simultaneous optimal precision is attainable in the asymptotic limit. We study trade-off relations that follow from known tight bounds in quantum multiparameter estimation. We compute trade-off curves and surfaces from Cramér–Rao type bounds which provide a compelling graphical representation of the information encoded in such bounds, and argue that bounds on simultaneously achievable precision in quantum multiparameter estimation should be regarded as measurement uncertainty relations. From the state-dependent bounds on the expected cost in parameter estimation, we derive a *state-independent* uncertainty relation between the parameters of a qubit system.

Keywords: quantum Fisher information, measurement uncertainty relations, trade-off, quantum multiparameter estimation

(Some figures may appear in colour only in the online journal)

[4] Author to whom any correspondence should be addressed.

Ever since its first formulation, the uncertainty principle has seen many refinements and clarifications. As quantum theory developed, its state-of-the-art concepts and mathematical tools were used to formulate in precise terms the ideas which were put forward in Heisenberg's 1927 paper [1]. As a result, our current understanding of the uncertainties inherent in quantum mechanics is spelled out in a collection of theorems pertaining to well defined operational tasks.

Soon after Heisenberg's paper, rigorous proofs of his uncertainty relations were formulated [2–4]. Those are usually referred to as *preparation uncertainty relations*. Most well known is the relation due to Weyl and Robertson

$$\sigma_A \sigma_B \geqslant \frac{1}{2} |\langle [A, B] \rangle|, \tag{1}$$

where $\sigma_A = \sqrt{\langle (A - \langle A \rangle)^2 \rangle}$ is the standard deviation of an observable A in a given state ψ ($\langle \cdot \rangle \equiv \langle \psi | \cdot | \psi \rangle$). For canonically conjugate observables such as position and momentum the right-hand side of equation (1) equals $\hbar/2$. This relation implies that it is impossible *to prepare* a particle in a state with arbitrarily sharp statistics for both position and momentum. Note that such uncertainty relations do not tell anything about statistics of joint measurements of both observables. Rather, the standard deviations on the left-hand side of equation (1) correspond to measurements of A and B on two independent ensembles of identical copies of the state $|\psi\rangle$. The preparation uncertainty relation between position and momentum is tight, as equality is achieved for specific states [5]. The relation equation (1) hence quantifies an attainable trade off between the sharpness of the position and momentum measurement statistics. Subsequent works formulated preparation uncertainty relations which involve other measures for the spread of a distribution [6–9].

The development of quantum measurement theory [10–12] allowed to formulate *accuracy–disturbance* uncertainty relations which quantify the disturbance caused by a positive operator valued measure (POVM) measurement to the statistics of a subsequent measurement of another POVM [13–19]. *Joint measurement uncertainty relations* have been discussed by many authors [20–24] and most recently in reference [25]. They describe the deviation of the statistics in a joint approximate measurement of two quantities from their statistics when measured separately. Many more authors have considered these two notions of uncertainty, for a more complete list see references in references [18, 19]. There is still debate between the proponents of the most recent approaches regarding which of them most properly captures Heisenberg's qualitative considerations [26–30].

Quantum parameter estimation theory provides yet another way to quantify quantum uncertainty. In this framework one considers a family of quantum states parameterized by real numbers, and the task is to estimate the parameters corresponding to a given state by performing measurements on identical copies of the state. In the one parameter case, the quantum Fisher information (QFI) Cramér–Rao bound provides a lower bound on the asymptotic scaling of the variance of an unbiased estimator [31, 32]. The bound is achievable in the asymptotic limit of many copies of the state with a separable measurement [33, 34]. Of particular importance is the case when the parameter to be estimated is elapsed time t for a state $|\psi(t)\rangle = \exp(-\mathrm{i}tH)|\psi(0)\rangle$ evolved with a given Hamiltonian. In that case, the quantum Cramér–Rao bound is proportional to the expectation value of the Hamiltonian [35], and hence yields the well known time-energy uncertainty relation [36]. Results of this type can be seen as *hybrid preparation–measurement uncertainty relations*, as they describe a trade off between the accuracy of a *measured* quantity, namely, the estimator for the desired parameter; and the

variance of the operator generating translations in that parameter—a quantity pertaining to the preparation. Quantum parameter estimation has been also used to formulate joint measurement uncertainty relations [37, 38] and error-disturbance relations [39].

Classically, going from single parameter estimation to a multiparameter setting involves replacing the scalar Cramér–Rao bound by a matrix inequality. This multiparameter bound is still asymptotically achievable [40], which means that the optimal precision can be achieved for all parameters *simultaneously*. In quantum multiparameter estimation however, the quantum Cramér–Rao bound is in general no longer attainable as the measurements required to attain the single parameter bound for the individual parameters might not be compatible [41]. In this setting one expects there to be *trade-offs* between the precision achievable for the estimators of different parameters. This is clearly a pure quantum phenomenon, and such trade-offs should hence be viewed as yet another manifestation of quantum uncertainty. Such bounds on quantum multiparameter estimation belong to the measurement type of uncertainty relations. The 'no go' part of such uncertainty relations is the unattainability of the multiparameter QFI Cramér–Rao bound. It implies that, in contrast to the classical case, optimal precision for all parameters simultaneously is impossible to achieve—acquiring better statistics for one parameter automatically leads to worse statistics for the complementary ones. The positive content is the characterization of the achievable trade-off and the measurement schemes attaining it. Various bounds on quantum multiparameter estimation that appear in the literature already encode such trade-off relations. The aim of this paper is to focus attention on this particular feature of the known tight bounds.

A point of distinction from other kinds of uncertainty relations is that in quantum multiparameter estimation, the quantities that one tries to estimate do have simultaneously well defined values. The task is to uncover classical information encoded in the state, e.g. the settings that were chosen on the device that prepared the state. Furthermore, given arbitrarily many copies of the state, all of the parameters can be estimated with arbitrary precision. Trade-offs appear when one considers the precision increase for each unknown parameter *per copy* of the state. A sharp distinction has to be also made between bounds for separable measurements—the realistic situation in experiments—versus collective measurements—which involve highly entangled measurements between the different copies.

Quantum multiparameter estimation has been an active field of research for nearly five decades. It has seen significant recent development which was stimulated in part by the increasing relevance of multiparameter estimation to quantum metrology tasks. We do not attempt to provide a comprehensive review of the field. Rather, we present the minimal background needed for the presentation of our results in a self contained manner. For a proper and up-to-date introduction to the field we refer to several very recent reviews which cover the state-of-the-art theoretical results as well as applications to concrete tasks [42–45].

Attainable bounds for multiparameter estimation are known for several quantum statistical models. For estimation of shift parameters of Gaussian states, Holevo proved an achievable bound [32]; this bound is referred to in the literature as the Holevo Cramér–Rao bound [41, 46]. The theory of local asymptotic normality [47, 48] implies that this bound is achievable for finite dimensional quantum systems if one allows *collective* measurements to be performed on many identical copies of the state. Attainable bounds for a qubit system have been proven by Nagaoka, Hayashi, and Gill and Massar [49–51]. Attainability of the quantum Fisher information Cramér–Rao bound with collective measurements has been shown to be equivalent to what is called the commutation condition [41], which involves

the commutators of the operators whose measurement provides the optimal one parameter precision.

In this paper, we translate the various known bounds on quantum multiparameter estimation into trade-off curves (or hyper-surfaces, in the case of more that two parameters). Such curves provide a visually clear representation of the information encoded in bounds on estimation. They highlight the trade-off, which is not evident when the bounds are written down as inequalities. To demonstrate this, we use trade-off curves to compare the bounds for estimation in a qubit system such as the Gill–Massar (GM) bound (which is attainable with separable measurements) and the Holevo Cramér–Rao bound (which is only attainable with collective measurements). We show how to sample points from the trade-off surface corresponding to the GM bound for different parameterizations of a qubit state and discuss the family of measurements which attain the bound in different parameterizations.

Our main result is the derivation of a *state-independent* trade-off relation between the three parameters of a qubit system when estimated using separable measurements. This result follows from the GM bound which is state-dependent—like many other Cramér–Rao type bounds in quantum parameter estimation. Our state-independent result is obtained by superimposing the trade-off surfaces corresponding to different states in the same plot to obtain a region in the 3-dimensional space of the variances of the estimators which is unattainable for all states and all separable measurements. This result implies a *state-independent measurement uncertainty relation* between the three Pauli operators $\sigma_x, \sigma_y, \sigma_z$. We prove the corresponding uncertainty relation for two parameters which turns out to have a simple additive form $\mathrm{Var}(\hat{\theta}_i) + \mathrm{Var}(\hat{\theta}_j) \geqslant 1/(4N)$, for $i \neq j \in \{x, y, z\}$, where $\mathrm{Var}(\hat{\theta}_i)$ is the variance of the estimator for the parameter $\theta_i = \langle \sigma_i/2 \rangle$ and N is the number of copies of the state. We further show that the bound $\mathrm{Var}(\hat{\theta}_x) + \mathrm{Var}(\hat{\theta}_y) + \mathrm{Var}(\hat{\theta}_z) \geqslant 1/N$ holds and forms part of the trade-off surface.

Finally, we compute the Holevo Cramér–Rao bound for a three level system model and describe the structure of its trade-off surface which is generic to any d-dimensional quantum system.

The paper is organized as follows. In section 1 we briefly review the required background and set up our notation; in section 2 we show how to obtain trade-off curves from Cramér–Rao type bounds for two parameters; in section 3 we present our results for the qubit model which include a state-independent trade-off surface; section 4 describes the structure of the trade-off surface of the Holevo Cramér–Rao bound for a qutrit; we conclude with a discussion in section 5.

1. Preliminaries

We start by reviewing estimation theory. In classical estimation theory [40] we are given a family of probability distributions with probability density $p(\boldsymbol{\theta})$ parametrized by a vector of parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$. The task is to estimate the unknown values $\boldsymbol{\theta}_0$ by sampling from $p(\boldsymbol{\theta}_0)$. In order to do so we shall pick an estimator, a function that produces an estimated value $\hat{\boldsymbol{\theta}}(x_1, x_2, \ldots, x_N)$ given the N samples drawn $\{x_i\}$. The estimation statistics are then described by the random variable $\hat{\boldsymbol{\theta}}(X_1, X_2, \ldots, X_N)$, where the random variables X_i are distributed according to $p(x_i|\boldsymbol{\theta}_0) := p(X = x_i|\boldsymbol{\theta}_0)$. An estimator is called locally unbiased if $\mathbb{E}\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0$, where $\mathbb{E}$ is the expectation value is with respect to $p(\boldsymbol{\theta}_0)$.

1.1. The Cramér–Rao bound

Let θ be a single parameter. The Cramér–Rao bound is a lower bound on the variance of the estimator $\mathrm{Var}(\hat{\theta}) = \mathbb{E}(\hat{\theta} - \theta_0)^2$. When the estimator is unbiased the bound is given by the inverse of the Fisher information $\mathfrak{f}(\theta_0) := \mathbb{E}\left(\frac{\mathrm{d}\ \log\ p(x|\theta)}{\mathrm{d}\theta}\right)^2\Big|_{\theta=\theta_0}$

$$\mathrm{Var}(\hat{\theta}) \geqslant \frac{1}{N\mathfrak{f}(\theta_0)}, \tag{2}$$

with N the number of samples.

In the multiparameter case we define the covariance matrix of the estimators (we shall suppress the θ_0 dependence in the notation)

$$V(\hat{\boldsymbol{\theta}})_{ij} = \mathbb{E}(\hat{\theta}_i - \theta_i)(\hat{\theta}_j - \theta_j),$$

and the Fisher information matrix

$$\mathfrak{F}_{ij} = \mathbb{E}\frac{\mathrm{d}\ \log\ p}{\mathrm{d}\theta_i}\frac{\mathrm{d}\ \log\ p}{\mathrm{d}\theta_j}. \tag{3}$$

The Cramér–Rao bound then takes the form of an inequality between positive semi definite matrices

$$V(\hat{\boldsymbol{\theta}}) \geqslant \mathfrak{F}(\boldsymbol{\theta}_0)^{-1}/N. \tag{4}$$

This bound is achievable asymptotically by the maximum likelihood method. More precisely, it is shown that there is a locally unbiased estimator for which the rescaled covariance matrix $NV \approx \mathfrak{F}^{-1}$ in the limit of large N [40]. To compensate for the overall $1/N$ improvement in precision due to the use of many copies of the source $p(\boldsymbol{\theta})$, we pick the rescaled covariance matrix NV as the figure of merit for the precision of the estimators in the asymptotic regime. We keep the N explicit in the notation as a reminder.

1.2. Quantum parameter estimation

In quantum parameter estimation, instead of a probability distribution we are given a quantum state $\rho(\theta)$ (satisfying $\rho \geqslant 0, \mathrm{Tr}\rho = 1$) which depends on θ. For a given measurement $\boldsymbol{M}$ with POVM elements $\{M_i\}$ (satisfying $M_i \geqslant 0$, $\sum M_i = \mathbb{I}$) we obtain a probability distribution for the outcomes $p^{\boldsymbol{M}}(i|\theta) = \mathrm{Tr}\, M_i\rho(\theta)$ which depends on θ through the state $\rho(\theta)$. Classical estimation theory can now be applied to the estimation of θ from $p^{\boldsymbol{M}}$. The problem of quantum parameter estimation is hence equivalent to the one of finding the measurement which maximizes this classical Fisher information. The Fisher information associated with the measurement $\boldsymbol{M}$ is

$$\mathfrak{f}^{\boldsymbol{M}}(\theta_0) := \mathbb{E}\left(\frac{\mathrm{d}\ \log\ p^{\boldsymbol{M}}}{\mathrm{d}\theta}\right)^2 = \sum_i \frac{\left(\mathrm{Tr}\ M_i\frac{\mathrm{d}\rho}{\mathrm{d}\theta}\right)^2}{\mathrm{Tr}\ M_i\rho(\theta_0)}, \tag{5}$$

where $\frac{\mathrm{d}\rho}{\mathrm{d}\theta}$ is evaluated at θ_0. The symmetric logarithmic derivative quantum Fisher information (SLD-QFI) is defined as

$$\mathfrak{h}(\theta_0) = \mathrm{Tr}(\rho(\theta_0)L(\theta_0)^2),$$

where $L(\theta_0)$ is the symmetric logarithmic derivative (SLD) defined implicitly by

$$\left.\frac{\mathrm{d}\rho}{\mathrm{d}\theta}\right|_{\theta_0} =: \frac{1}{2}(L(\theta_0)\rho(\theta_0) + \rho(\theta_0)L(\theta_0)). \tag{6}$$

When ρ is of full rank, solutions to equation (6) are unique as the only matrix that anti-commutes with ρ is the zero matrix. We will always assume that this is the case. For treatment of the case of degenerate states see references [33, 52, 53]. The SLD-QFI bound [33] states that for any measurement $\boldsymbol{M}$

$$\mathfrak{f}^{\boldsymbol{M}}(\theta_0) \leqslant \mathfrak{h}(\theta_0), \tag{7}$$

(we shall suppress the θ_0 dependence from now on). The proof is obtained by the use of the Cauchy–Schwarz inequality:

$$\begin{aligned}\mathfrak{f}^{\boldsymbol{M}} &= \sum_i \frac{\left(\mathrm{Tr}\ M_i\frac{\mathrm{d}\rho}{\mathrm{d}\theta}\right)^2}{\mathrm{Tr}\ M_i\rho} = \sum_i \frac{\mathrm{Re}(\mathrm{Tr}\ M_i\rho L)^2}{\mathrm{Tr}\ M_i\rho} \leqslant \sum_i \frac{|\mathrm{Tr}\ M_i\rho L|^2}{\mathrm{Tr}\ M_i\rho} \\ &= \sum_i \frac{\left|\mathrm{Tr}\sqrt{M_i}\sqrt{M_i}\sqrt{\rho}\sqrt{\rho}L\right|^2}{\mathrm{Tr}\ M_i\rho} \leqslant \sum_i \frac{\mathrm{Tr}(M_i\rho)\mathrm{Tr}(\rho L M_i L)}{\mathrm{Tr}\ M_i\rho} = \mathrm{Tr}(\rho L^2) = \mathfrak{h},\end{aligned}$$

where we used the definition of the SLD in the second equality and $\sum M_i = \mathbb{I}$ in going to the last line. Braunstein and Caves [33] proved that equality in equation (7) is attained when $\boldsymbol{M}$ is a projective measurement in the basis which diagonalizes L, hence identifying the optimal measurement strategy.

One can also define the right logarithmic derivative (RLD) and corresponding to it is the right logarithmic derivative quantum Fisher information (RLD-QFI) bound. This bound will be discussed later.

In the case of multiple parameters, the Fisher information matrix of the measurement $\boldsymbol{M}$ is defined according to equation (3) as

$$\mathfrak{F}^{\boldsymbol{M}}_{ij} := \mathbb{E}\frac{\mathrm{d}\ \log\ p^{\boldsymbol{M}}}{\mathrm{d}\theta_i}\frac{\mathrm{d}\ \log\ p^{\boldsymbol{M}}}{\mathrm{d}\theta_j} = \sum_k \frac{\mathrm{Tr}\ M_k\frac{\mathrm{d}\rho}{\mathrm{d}\theta_i}\mathrm{Tr}\ M_k\frac{\mathrm{d}\rho}{\mathrm{d}\theta_j}}{\mathrm{Tr}\ M_k\rho}.$$

The quantum Fisher information matrix is defined as

$$\mathfrak{H}_{ij} = \frac{1}{2}\mathrm{Tr}\ \rho(L_iL_j + L_jL_i),$$

where L_i is the symmetric logarithmic derivative with respect to θ_i. The multiparameter SLD-QFI bound is an inequality in the sense of positive semidefinite matrices:

$$\mathfrak{F}^{\boldsymbol{M}} \leqslant \mathfrak{H}. \tag{8}$$

This bound is a consequence of the one parameter bound equation (7). To see this, let $\boldsymbol{v}$ be a vector in the space of parameters $\mathbb{R}^K$,[5] let $\theta_{\boldsymbol{v}} := \sum_i v_i\theta_i$. From linearity of the

[5] A note on notation: to reduce confusion between state vectors in Hilbert space and vectors in parameter space we will stick to Dirac notation $\langle\psi|O|\phi\rangle$ for the former and vector notation $\boldsymbol{v}^\intercal M\boldsymbol{v}$ for the latter.

definition of the SLD equation (6), it follows that the corresponding symmetric logarithmic derivative is

$$L_{\boldsymbol{v}} = \sum_i v_i L_i. \tag{9}$$

We then have

$$\boldsymbol{v}^\intercal \mathfrak{F}^{\boldsymbol{M}} \boldsymbol{v} = \sum_{ijk} v_i \frac{\mathrm{Tr}\ M_k \frac{\mathrm{d}\rho}{\mathrm{d}\theta_i} \mathrm{Tr}\ M_k \frac{\mathrm{d}\rho}{\mathrm{d}\theta_j}}{\mathrm{Tr}\ M_k \rho} v_j = \sum_k \frac{\mathrm{Tr}\ M_k \frac{\mathrm{d}\rho}{\mathrm{d}\theta_{\boldsymbol{v}}} \mathrm{Tr}\ M_k \frac{\mathrm{d}\rho}{\mathrm{d}\theta_{\boldsymbol{v}}}}{\mathrm{Tr}\ M_k \rho} = \mathfrak{f}_{\boldsymbol{v}} \leqslant$$

$$\mathfrak{h}_v = \mathrm{Tr}\ \rho L_{\boldsymbol{v}} L_{\boldsymbol{v}} = \sum_{ij} v_i \frac{1}{2} \mathrm{Tr}\ \rho (L_i L_j + L_i L_j) v_j = \boldsymbol{v}^\intercal \mathfrak{H} \boldsymbol{v},$$

where $\mathfrak{f}_v^{\boldsymbol{M}}$, and $\mathfrak{h}_v$ denote the one parameter Fisher information of the measurement $\boldsymbol{M}$ and the quantum Fisher information for the estimation of $\theta_{\boldsymbol{v}}$ respectively.

In other words, in the multi parameter setting, the SLD-QFI bound equation (8) can be stated as the following: for any linear combination of the parameters $\theta_{\boldsymbol{v}} = \sum v_i \theta_i$, a one parameter SLD-QFI bound $\mathfrak{f}_v \leqslant \mathfrak{h}_v$ applies. In addition, for any $\boldsymbol{v}$ the bound is attainable with a projective measurement in the basis diagonalizing $L_{\boldsymbol{v}}$.[6]

Further notice that because of their quadratic forms, the covariance matrix V, the Fisher information matrix $\mathfrak{F}$, and the quantum Fisher information matrix $\mathfrak{H}$ all transform in the same way under linear coordinate transformations. When $\theta_i \mapsto \tilde{\theta}_i := \sum_j R_{ij} \theta_j$, all three matrices transform as $(\cdot) \mapsto R(\cdot)R^\intercal$. This implies that matrix inequalities between them are invariant under rotations of coordinates[7].

2. Trade-off

If two linear combinations of the parameters $\{\theta_i\}$ defined by the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ result in commuting SLDs $[L_{\boldsymbol{u}}, L_{\boldsymbol{v}}] = 0$, then optimal estimation of the two parameters $\theta_{\boldsymbol{u}}$ and $\theta_{\boldsymbol{v}}$ can be achieved simultaneously by performing a measurement in the basis which diagonalizes both of them.

However, $[L_{\boldsymbol{u}}, L_{\boldsymbol{v}}] = 0$ will typically not be satisfied. In general, we expect there to be a trade-off between the achievable precision in the two parameters in the following sense. Let $\boldsymbol{M}(\lambda)$, $\lambda \in [0, 1]$ be a family of measurements with POVM elements $\{\boldsymbol{M}_i(\lambda)\}$ such that $\boldsymbol{M}(0)$ is the optimal measurement for $\theta_{\boldsymbol{v}}$ and $\boldsymbol{M}(1)$ is the optimal measurement for $\theta_{\boldsymbol{u}}$. For intermediate values of λ the precisions of the estimators for $\theta_{\boldsymbol{x}}$ ($\{\boldsymbol{x}\}=\{\boldsymbol{u}\}$ or $\{\boldsymbol{v}\}$) which we quantify by $\mathrm{Var}(\theta_{\boldsymbol{x}})$ will take values larger than optimal.

[6] Note that for the measurement $\tilde{M}$ which is optimal for the estimation of the parameter θ_1 in a one parameter setting (i.e. when all other parameters are kept fixed) we have asymptotically $NV \approx \mathfrak{F}^{\tilde{M}^{-1}}$ which implies

$$NV_1 = (\mathfrak{F}^{\tilde{M}^{-1}})_{11} \geqslant (\mathfrak{H}^{-1})_{11} \geqslant 1/\mathfrak{H}_{11} = 1/\mathfrak{h}_1 = 1/\mathfrak{f}_1^{\tilde{M}},$$

where the second inequality is a general property of positive matrices, and the last equality is due to the optimality assumption about $\tilde{M}$. That is, the optimal measurement for one parameter when estimated alone might perform worse for the estimation of the same parameter when additional parameters are unknown [41].

[7] Because we are dealing with local estimation, only linear coordinate transformations are of interest (see reference [51]). An arbitrary (smooth) coordinate transformation will be approximated to first order by a linear one $\theta_i \mapsto \tilde{\theta}_i(\boldsymbol{\theta}) = \tilde{\theta}_i(\boldsymbol{\theta}_0) + \sum_j \partial_{\theta_j} \tilde{\theta}_i(\boldsymbol{\theta}_0)(\theta_j - \theta_{0j}) + o(|\boldsymbol{\theta} - \boldsymbol{\theta}_0|)$.

Trade-off curves are commonly used in detection theory. In particular receiver operating characteristic curves (ROC curves) are a convenient way to represent how the probability for false positive detection increases as one increases the sensitivity [54]. In the context of uncertainty relations, a similar representation was used in [55] for preparation uncertainties of angular momentum components. As we will now show, trade-off curves (or surfaces) are a convenient representation of the data which is typically encoded in uncertainty relations.

The known bounds on precision in parameter estimation are most often stated as lower bounds on the expected cost, resulting from a given positive definite $K \times K$ cost matrix G [32, 46, 51].[8] In general these are bounds of the form

$$\mathrm{Tr}\ VG \geqslant f(G), \quad \forall G \geqslant 0,$$

where V is the covariance matrix of the estimator $\hat{\boldsymbol{\theta}}$ and f is a real scalar function on semi definite matrices. This family of inequalities defines a region in $\mathbb{R}^K$ of allowed values for the vector of variances $(\mathrm{Var}(\theta_1), \mathrm{Var}(\theta_2), \ldots, \mathrm{Var}(\theta_K))$. The boundary of this region is the trade-off surface. We now show how this is obtained by considering specific examples.

2.1. Classical trade-off curves: the quantum Fisher information Cramér–Rao bound

By *classical* we refer to the situation when the optimal precision values for the different parameters are independent of each other. This is automatically the case in classical parameter estimation where the maximum likelihood method asymptotically achieves the optimal values for all the variances $\mathrm{Var}(\theta_i)$ simultaneously [40].

Let us begin by plotting the trade-off curve resulting from the SLD-QFI bound equation (8). As discussed above, this bound can be interpreted as the assertion that for every direction in parameter space, the single parameter bound applies. Therefore we do not expect to be able to extract nontrivial trade-off relations from it.

The matrix inequalities equations (4) and (8) imply

$$\mathrm{Tr}\ NVG \geqslant \mathrm{Tr}\ \mathfrak{H}^{-1}G, \quad \forall G \geqslant 0. \tag{10}$$

Consider the case of two parameters and let $G = \begin{pmatrix} t & \\ & 1-t \end{pmatrix}$ for $t \in (0,1)$. This form of cost matrix corresponds to a fixed total cost of 1 which is divided between θ_1 and θ_2 with proportion $t/(1-t)$. Let $\mathfrak{H}^{-1} = \begin{pmatrix} u_1 & b \\ b & u_2 \end{pmatrix}$. equation (10) becomes

$$N(tV_1 + (1-t)V_2) \geqslant tu_1 + (1-t)u_2,$$

where V_i is the variance of θ_i. This implies that for every value of $t \in (0,1)$ the points in the (NV_1, NV_2) plane which are not excluded by equation (10) lie above the line $NV_2 = u_2 + (u_1 - NV_1)\frac{t}{1-t}$. All of these lines pass through the point (u_1, u_2) and as t varies between 0 and 1 the

[8] Once again, because we are dealing with local estimation representing the cost function by a positive matrix is general enough. Expanding an arbitrary cost function $f(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ around the minimum $\boldsymbol{\theta}_0$ and taking the expectation value we obtain $\mathbb{E}f(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = f(\boldsymbol{\theta}_0) + 1/2 \sum (\partial_{\theta_i}\partial_{\theta_j} f)\mathbb{E}(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_{0i})(\hat{\boldsymbol{\theta}}_j - \boldsymbol{\theta}_{0j}) + o(|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0|^2) = f(\boldsymbol{\theta}_0) + 1/2\mathrm{Tr}\ GV + o(|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0|^2)$, where $G_{ij} \coloneqq \partial_{\theta_i}\partial_{\theta_j} f$ is the Hessian.

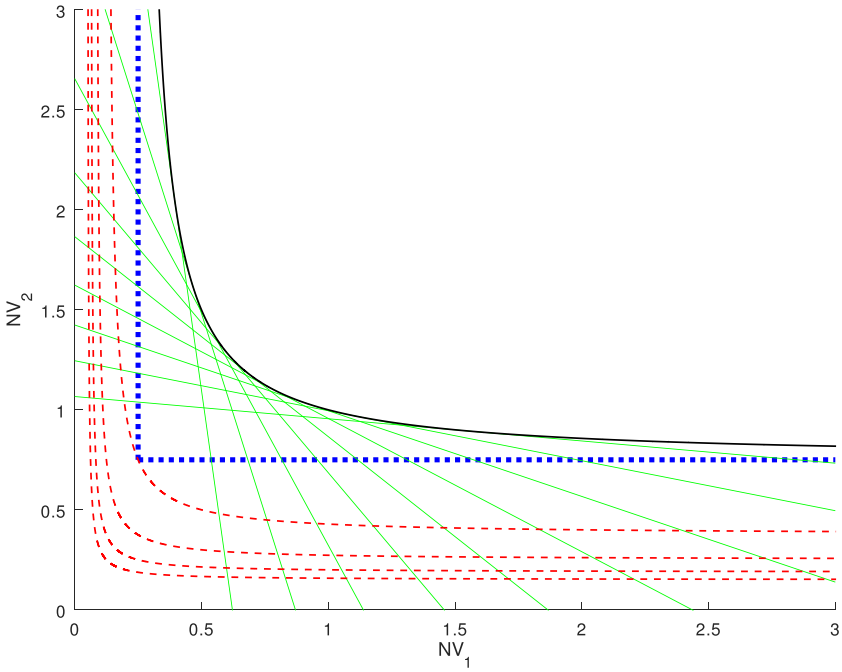


Figure 1. Trade-off curves for the rescaled variances of the estimators of a 2-parameter density matrix with quantum Fisher information matrix $\mathfrak{H} = \begin{pmatrix} u_1 & b \\ b & u_2 \end{pmatrix}^{-1}$. The SLD-QFI bound equation (8) implies the dotted blue classical trade-off curve. The curves resulting from the GM bound equation (12) are plotted for fixed values of $u_1 = 0.25$ and $u_2 = 0.75$ for systems of different Hilbert space dimensions: $d = 2$ (black solid curve), and $d = 3, 4, 5, 6$ (red dashed curves). The GM curves are obtained as the upper envelopes of the lines given by equation (14), several of which are plotted (green) for the $d = 2$ case. For $d = 2$ the GM trade-off curve is asymptotic to the SLD-QFI curve, whereas for $d > 2$ the GM curves are below the SLD-QFI curve.

slope of the line varies between 0 and $-\infty$. The allowed region (not excluded by any value of t) is $\{NV_1 \geqslant u_1\} \cap \{NV_2 \geqslant u_2\}$. In particular, the bound equation (10) does not exclude the point $(NV_1 = u_1, NV_2 = u_2)$, which corresponds to optimal precision for both θ_1 and θ_2 simultaneously. This classical—or trivial—trade-off bound is plotted in figure 1 as the blue dotted curve.

2.2. *Non-trivial trade-off curves: the Gill–Massar bound*

To demonstrate nontrivial trade-off we shall introduce the bound proved by Gill and Massar in reference [51]. They showed that for separable measurements on N identical copies of finite, d-dimensional quantum systems the following holds:

$$\mathrm{Tr}\ \mathfrak{F}^{\boldsymbol{M}}\mathfrak{H}^{-1} \leqslant N(d-1). \tag{11}$$

This bound implies [51] that for any $G \geqslant 0$

$$\mathrm{Tr}\, NVG \approx \mathrm{Tr}(\mathfrak{F}^{M})^{-1}G \geqslant \frac{1}{d-1}\left(\mathrm{Tr}\sqrt{\sqrt{G}\mathfrak{H}^{-1}\sqrt{G}}\right)^{2}. \tag{12}$$

We will refer to equation (12) as the Gill–Massar (GM) bound.

The non-linear dependence of the right-hand side of equation (12) on G results in a non-trivial trade-off curve. Let G and $\mathfrak{H}$ be parameterized as before. Using the following expression for the fidelity of 2×2 matrices [56] which appears in the right-hand side of equation (12)

$$\left(\mathrm{Tr}\sqrt{\sqrt{A}B\sqrt{A}}\right)^{2} = \mathrm{Tr}\, AB + 2\sqrt{\det(AB)}, \tag{13}$$

we obtain the following family of lines in the (NV_1, NV_2) plane:

$$tNV_1 + (1-t)NV_2 = \frac{1}{d-1}\left(tu_1 + (1-t)u_2 + 2\sqrt{t(1-t)}\sqrt{\det\mathfrak{H}^{-1}}\right). \tag{14}$$

To obtain a formula for the trade-off curve fix V_1 and maximize V_2 with respect to t. This results in the following parameterization of the curve in terms of $t\in(0,1)$:

$$NV_1(t) = \frac{1}{d-1}\left(u_1 + \sqrt{\frac{1-t}{t}}\sqrt{\det\mathfrak{H}^{-1}}\right)$$

$$NV_2(t) = \frac{1}{d-1}\left(u_2 + \sqrt{\frac{t}{1-t}}\sqrt{\det\mathfrak{H}^{-1}}\right).$$

Figure 1 shows the trade-off curves obtained for fixed values of u_1, u_2 and for $d = 2, 3, \ldots, 6$. In addition the trivial trade-off curve resulting from the SLD-QFI bound is shown. The figure clearly shows that for $d > 2$ the GM bound is unattainable as it allows a higher precision for each of the parameters than that allowed by the SLD-QFI bound. Furthermore, figure 1 shows that for $d > 2$ the GM bound does not exclude any region above the trivial trade-off curve. This is in agreement with reference [51] where it was concluded that when the number of parameters K satisfies $K \leqslant d-1$, the SLD-QFI bound is stronger then the GM bound.

2.3. The right logarithmic derivative quantum Fisher information bound and the Holevo Cramér–Rao bound

We shall now introduce the right logarithmic derivative quantum Fisher information (RLD-QFI) bound. This bound exhibits nontrivial trade-off, with the 'strength' of the trade-off between the variances of θ_i and θ_j depending directly on the expectation value of the commutator of the corresponding SLDs $\mathrm{Tr}\rho[L_i, L_j]$.

The right logarithmic derivative (RLD) is defined implicitly by

$$\frac{\partial\rho}{\partial\theta_i} = \rho Ł_i.$$

The RLD-QFI matrix is then defined by

$$\mathfrak{R}_{ij} = \mathrm{Tr}\ \rho Ł_j Ł_i^{\dagger}.$$

Just as the SLD-QFI matrix, the RLD-QFI matrix bounds the covariance matrix of any locally unbiased estimator [46]:

$$V(\hat{\boldsymbol{\theta}}) \geqslant \mathfrak{R}^{-1}/N.$$

This bound implies, as before, a lower bound on the expected cost associated with any positive cost matrix $G > 0$, which, due to the fact that $\mathfrak{R}$ is a Hermitian matrix (whereas $\mathfrak{H}$ is real and symmetric) takes the form [32, lemma 6.6.1]

$$\operatorname{Tr} NVG \geqslant \operatorname{Tr} G \operatorname{Re}\left(\mathfrak{R}^{-1}\right) + \operatorname{Tr}\left|\sqrt{G}\operatorname{Im}\left(\mathfrak{R}^{-1}\right)\sqrt{G}\right|, \tag{15}$$

where $|\cdot|$ is the absolute value function defined for Hermitian matrices via their spectral decomposition; and Re and Im refer to the real and imaginary parts of a matrix taken entry-wise. The imaginary part results in a non-trivial trade-off curve. To see this, consider the case of two parameters. Because $\mathfrak{R}^{-1}$ is Hermitian, its imaginary part is anti-symmetric. Let $\mathfrak{R}^{-1} = \begin{pmatrix} r_1 & b + \mathrm{i}a \\ b - \mathrm{i}a & r_2 \end{pmatrix}$, and $G = \begin{pmatrix} t & \\ & 1-t \end{pmatrix}$. Equation (15) becomes

$$tNV_{11} + (1-t)NV_{22} \geqslant tr_1 + (1-t)r_2 + 2a\sqrt{t(1-t)}. \tag{16}$$

The right-hand side has the same functional dependence on t as in equation (14) with $d = 2$. From this we conclude that this bound results in a non-trivial trade-off curve which is asymptotic to the lines $NV_1 = r_1$ and $NV_2 = r_2$.

In certain cases, it is possible to express the RLD-QFI matrix in terms of the SLDs. In the case of what is called a $\mathcal{D}$-invariant model[9] [32, 46] the following holds:

$$\mathfrak{R}^{-1} = \mathfrak{H}^{-1} + \frac{\mathrm{i}}{2}\mathfrak{H}^{-1}D\mathfrak{H}^{-1}, \tag{17}$$

where D is a matrix whose entries are proportional to the expectation values of the commutators of the SLDs:

$$D_{ij} = \mathrm{i}\operatorname{Tr}\rho\left[L_i, L_j\right]. \tag{18}$$

As $\mathfrak{H}$ and D are real, the imaginary part of $\mathfrak{R}^{-1}$ is $\mathfrak{H}^{-1}D\mathfrak{H}^{-1}/2$, which together with equation (15) implies

$$\operatorname{Tr} NVG \geqslant \operatorname{Tr} G\mathfrak{H}^{-1} + \frac{1}{2}\operatorname{Tr}\left|\sqrt{G}\mathfrak{H}^{-1}D\mathfrak{H}^{-1}\sqrt{G}\right|. \tag{19}$$

Comparing to equation (16) we see that in this case $\operatorname{Tr}\rho\left[L_i, L_j\right]$ determines how much area the trade-off curve excludes above the trivial curve resulting from the SLD-QFI bound equation (10) (which has only the $\operatorname{Tr} G\mathfrak{H}^{-1}$ term).

We mention the Holevo Cramér–Rao bound, which is in general stronger than both the SLD-QFI and the RLD-QFI bounds [46]. In the $\mathcal{D}$-invariant case the Holevo bound coincides

[9] The $\mathcal{D}$ operator is defined implicitly by $\mathcal{D}(X)\rho_0 + \rho_0\mathcal{D}(X) = 2i[X, \rho_0]$. A model is called $\mathcal{D}$–invariant if the space spanned by the SLDs is invariant under the action of $\mathcal{D}$. For a further classification of statistical models see reference [57].

with the RLD-QFI bound [32, 46]. As we will be dealing only with such cases, we shall not present the Holevo Cramér–Rao bound here and only mention results we will need for our discussion[10]. The Holevo Cramér–Rao bound was shown to be equal to the SLD-QFI bound if the expectation values of the commutators between all SLDs vanish [41]. In Gaussian state shift models where one estimates the displacement parameters, it has been shown that the Holevo Cramér–Rao bound is attainable [32]. The theory of local asymptotic normality maps any quantum estimation problem involving many copies of the same state to a Gaussian shift model [47]. This implies asymptotic attainability of the Holevo Cramér–Rao bound with *collective* measurements [41, 48].

3. The qubit model

Let us next move to the estimation of the most general density matrix of a qubit, which is parameterized by three parameters. This problem is also known as quantum state tomography [60]. In order to observe trade-off relations between more than two parameters, it is enough to consider a qubit system. In the qubit case, the GM bound is attainable with a measurement performed on single copies of the state [51, 61].

In this section we compare the GM bound and the Holevo Cramér–Rao bound (which in this case is equal to the RLD bound) through the resulting trade-off surfaces. We also investigate the set of optimal measurements which saturate the inequalities. We characterize this set in two cases: when the parameterization is aligned with ρ_0 (when the z axis is pointing in the direction of the Bloch vector of ρ_0); and when it is not aligned. Finally we use the trade-off surfaces computed for different parameterizations to obtain a state-independent trade-off surface, and derive state-independent uncertainty relations.

We work in the Bloch sphere parameterization, using Pauli matrices as a basis, and with $\rho_0 = [\mathbb{I} + z_0\sigma_z]/2$, the full parameterization is $\rho(\boldsymbol{\theta}) = \rho_0 + \sum \theta_i\sigma_i$ (in this parameterization the true parameter value is $\boldsymbol{\theta}_0 = (0, 0, 0)$). Note that the initial state can always be brought to this form by rotating the Bloch sphere and working in the appropriate basis. We will call this coordinate system the *adjusted* one, and later—in section 3.3—we shall return to describe things in a general coordinate system. We will identify $\theta_1 \equiv x$, $\theta_2 \equiv y$ and $\theta_3 \equiv z$. When the state is full rank ($z_0 < 1$) the solution to the equation defining the SLDs is unique and given by

$$\begin{aligned} L_x &= 2\sigma_x, \quad L_y = 2\sigma_y \\ L_z &= 2\begin{pmatrix} (1+z_0)^{-1} & 0 \\ 0 & -(1-z_0)^{-1} \end{pmatrix}. \end{aligned} \tag{20}$$

The resulting SLD-QFI is diagonal and takes the form

$$\mathfrak{H} = 4\begin{pmatrix} 1 & & \\ & 1 & \\ & & (1-z_0^2)^{-1} \end{pmatrix}. \tag{21}$$

[10] I addition we mention that it has been recently shown that the bound in equation (19) is always greater or equal than the Holevo Cramér–Rao bound, and that the Holevo Cramér–Rao bound is less or equal than two times the SLD bound equation (10) [58, 59].

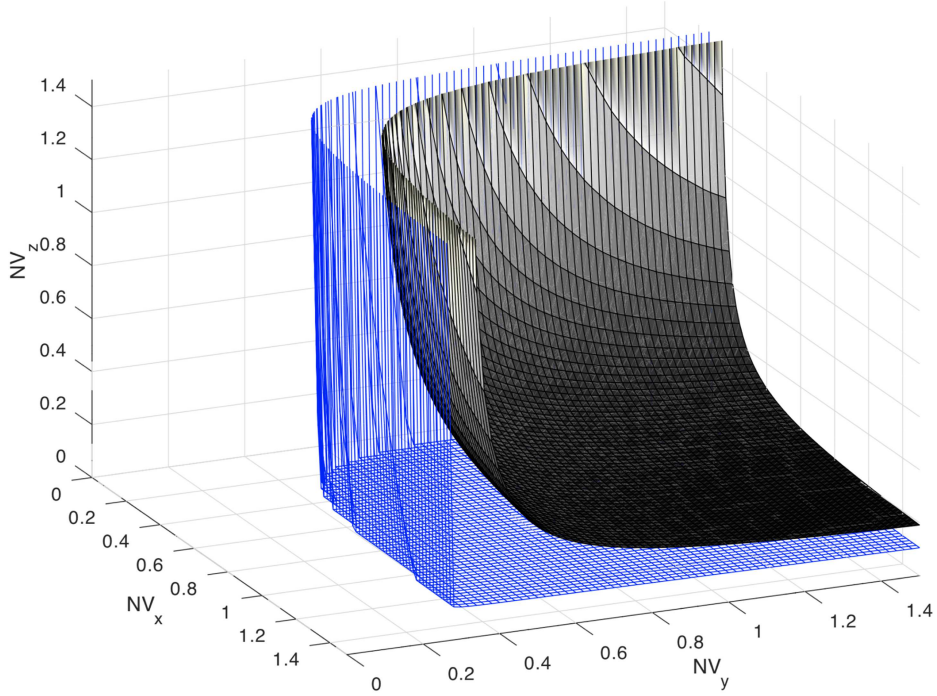


Figure 2. Comparison between the trade-off surfaces obtained from the Gill and Massar (GM) bound (grayscale and filled) and the right logarithmic derivative (RLD) or Holevo Cramér–Rao bound (blue and transparent) for estimation of the three Bloch parameters x, y and z for a qubit state $\rho_0 = (\mathbb{I} + 0.7\sigma_z)$. The axes are the rescaled variances of the parameters, i.e. the variance multiplied by N—the number of copies of ρ_0. The GM trade-off surface lies strictly above the RLD surface. The RLD surface shows nontrivial trade-off only between the x and y parameters (it has a 'flat bottom'), whereas the GM bound exhibits nontrivial trade-off between all three parameters simultaneously. Both surfaces are asymptotic to the trivial trade-off surface implied by the symmetric logarithmic derivative quantum Fisher information bound equation (8).

3.1. Comparison between Gill–Massar and Holevo Cramér–Rao bounds

Let us take a cost matrix parameterized as

$$G = \begin{pmatrix} s & & \\ & t & \\ & & r := 1 - t - s \end{pmatrix}; \quad s \geqslant 0,\ t \geqslant 0,\ s + t \leqslant 1. \tag{22}$$

The GM bound is given by equation (12):

$$\operatorname{Tr} NVG \geqslant \frac{1}{4}\left(s + t + r(1 - z_0^2) + 2\sqrt{ts} + 2\sqrt{r(1 - z_0^2)}(\sqrt{t} + \sqrt{s})\right). \tag{23}$$

The Holevo Cramér–Rao bound is equal to the RLD bound because the model is $\mathcal{D}$-invariant (this is verified by a direct computation). Computing the matrix $D_{ij} = \mathrm{iTr}\ \rho\left[L_i, L_j\right]$ we obtain

$$D = 8z_0 \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{24}$$

According to equation (19) the RLD-QFI bound is then given by

$$\mathrm{Tr}\ NVG \geqslant \frac{1}{4}\left(s + t + r(1 - z_0^2) + 2z_0\sqrt{ts}\right). \tag{25}$$

From this expression one can already guess that the RLD bound exhibits nontrivial trade-off only between the x and y parameters as r appears only in the term coming from $\mathrm{Tr}\ G\mathfrak{H}^{-1}$ on the right-hand side. This is a generic feature of the RLD-QFI bound for finite dimensional quantum systems. We will show that this is the case in a 3-level system in section 4.

Using equations (23) and (25) we find for each bound the smallest allowed value of NV_z for a grid of values of NV_x, NV_y (for fixed NV_x, NV_y we can find NV_z by requiring equality in equations (23) and (25) and maximizing over a grid of values for s and t). The results are plotted in figure 2. For states with $z_0 < 1$, the Holevo Cramér–Rao (=RLD) bound is strictly weaker than the GM bound. Recall that the GM bound is attainable with single copy measurements whereas the Holevo Cramér–Rao bound with collective measurements. This conforms with our expectation that collective entangled measurements should provide an advantage over separable ones. As the state ρ_0 tends towards a pure state, the GM bound tends towards the Holevo Cramér–Rao bound, as can be seen from equations (23) and (25) by setting $z_0 = 1$.

3.2. Measurements attaining the Gill–Massar bound

The bound equation (11) is achievable for qubits. Gill and Massar show that for a qubit system ($d = 2$), every matrix $\mathfrak{F}$ that satisfies equation (11) is obtainable as the Fisher information matrix of a measurement $\boldsymbol{M}^{\mathfrak{F}}$. $\boldsymbol{M}^{\mathfrak{F}}$ is a probabilistic mixture of three projective measurements along the directions which diagonalize $\mathfrak{F}$ (seen as Bloch vectors). By probabilistic mixture we mean combining measurements in the following way: let $\boldsymbol{M}^{(1)}$ and $\boldsymbol{M}^{(2)}$ be measurements with POVM elements $\{M_i^{(1)}\}_{i=1}^{I}$ and $\{M_j^{(2)}\}_{j=1}^{J}$. We say that $\boldsymbol{M}$ is a probabilistic mixture of $\boldsymbol{M}^{(1)}$ and $\boldsymbol{M}^{(2)}$ if $\boldsymbol{M}$ has $I + J$ POVM elements $M_k = \lambda M_k^{(1)}$ for $k = 1, \ldots, I$ and $M_k = (1 - \lambda)M_{k-I}^{(2)}$ for $k = I + 1, \ldots, I + J$ for some $\lambda \in (0, 1)$; and denote $\boldsymbol{M} = \lambda\boldsymbol{M}^1 \cup (1 - \lambda)\boldsymbol{M}^{(2)}$. With the obvious generalization to mixtures of more than two measurements. This corresponds to measuring $\boldsymbol{M}^{(1)}$ in λN copies of ρ out of an ensemble of N copies, and $\boldsymbol{M}^{(2)}$ on the rest. From equation (5) it is easily verified that probabilistic mixtures of measurements result in convex combinations of the Fisher information matrices with the same mixing coefficients, i.e. $\mathfrak{F}^{\boldsymbol{M}} = \lambda\mathfrak{F}^{\boldsymbol{M}^{(1)}} + (1 - \lambda)\mathfrak{F}^{\boldsymbol{M}^{(2)}}$.

In the rest of this section, we will require more detailed notation. We denote the Fisher information matrix corresponding to $\rho_{\boldsymbol{n}} = (\mathbb{I} + \boldsymbol{n} \cdot \sigma)/2$, the estimation of parameters $\boldsymbol{\theta}$, and to a projective measurement $\boldsymbol{M} = \boldsymbol{P}_{\boldsymbol{v}}$ along a Bloch vector $\boldsymbol{v}$ as $\mathfrak{F}(\rho_{\boldsymbol{n}}, \boldsymbol{\theta}, \boldsymbol{P}_{\boldsymbol{v}})$. The following calculation shows that this matrix equals $\frac{4}{1-(\boldsymbol{n}\cdot\boldsymbol{v})^2}\boldsymbol{v}\boldsymbol{v}^{\intercal}$.

$$\mathfrak{F}(\rho_{\boldsymbol{n}}, \boldsymbol{\theta}, \boldsymbol{P}_{\boldsymbol{v}})_{ij} := \frac{\mathrm{Tr}\ P_{\boldsymbol{v}}^{+} \frac{\mathrm{d}\rho}{\mathrm{d}\theta_i} \mathrm{Tr}\ P_{\boldsymbol{v}}^{+} \frac{\mathrm{d}\rho}{\mathrm{d}\theta_j}}{\mathrm{Tr}\ P_{\boldsymbol{v}}^{+}\rho} + \frac{\mathrm{Tr}\ P_{\boldsymbol{v}}^{-} \frac{\mathrm{d}\rho}{\mathrm{d}\theta_i} \mathrm{Tr}\ P_{\boldsymbol{v}}^{-} \frac{\mathrm{d}\rho}{\mathrm{d}\theta_j}}{\mathrm{Tr}\ P_{v}^{-}\rho}$$
$$= 2v_i v_j \left(\frac{1}{1+\boldsymbol{n}\cdot\boldsymbol{v}} + \frac{1}{1-\boldsymbol{n}\cdot\boldsymbol{v}} \right) = 4\frac{v_i v_j}{1-(\boldsymbol{n}\cdot\boldsymbol{v})^2} \tag{26}$$

In reference [51] it is shown, as part of the proof of the bound equation (11), that the optimal rescaled covariance matrix for a given cost matrix G is given by

$$NV_{\mathrm{opt}}(G) = \frac{1}{d-1} \left(\mathrm{Tr}\sqrt{G^{1/2}\mathfrak{H}^{-1}G^{1/2}} \right) G^{-1/2}\sqrt{G^{1/2}\mathfrak{H}^{-1}G^{1/2}}G^{-1/2} \tag{27}$$

Plugging in $d = 2$, $\mathfrak{H}$ from equation (21) and a cost matrix parameterized as in equation (22) we obtain

$$NV_{\mathrm{opt}}(s,t) = \frac{1}{4}\left(\sqrt{s} + \sqrt{t} + \sqrt{(1-s-t)(1-z_0^2)} \right) \begin{pmatrix} s^{-1/2} & 0 & 0 \\ 0 & t^{-1/2} & 0 \\ 0 & 0 & \sqrt{\dfrac{1-z_0^2}{(1-s-t)}} \end{pmatrix}$$

Comparing this with the Fisher information of a probabilistic mixture with proportions $(\alpha, \beta, 1-\alpha-\beta)$ of projective measurement in the $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{z}}$ directions ($\boldsymbol{M}(\alpha,\beta) = \alpha P_{\hat{\boldsymbol{x}}} \cup \beta P_{\hat{\boldsymbol{y}}} \cup (1-\alpha-\beta)P_{\hat{z}}$):

$$\mathfrak{F}(\rho_{z_0\hat{\boldsymbol{z}}}, \boldsymbol{\theta}, \boldsymbol{M}(\alpha,\beta)) = 4 \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \dfrac{1-\alpha-\beta}{1-z_0^2} \end{pmatrix},$$

we can find $\bar{\alpha}(s,t)$ and $\bar{\beta}(s,t)$ such that

$$\mathfrak{F}(\rho_{z_0\hat{\boldsymbol{z}}}, \boldsymbol{\theta}, \boldsymbol{M}(\bar{\alpha},\bar{\beta}))^{-1} = NV_{\mathrm{opt}}(s,t).$$

Those are given by

$$\bar{\alpha}(s,t) = \frac{\sqrt{s}}{\sqrt{s}+\sqrt{t}+\sqrt{(1-s-t)(1-z_0^2)}}; \quad \bar{\beta}(s,t) = \frac{\sqrt{t}}{\sqrt{s}+\sqrt{t}+\sqrt{(1-s-t)(1-z_0^2)}}$$

This gives a simple characterization of the optimal measurements, i.e. the measurements for which the obtained variances lie on the trade-off surface. They are probabilistic mixtures of projective measurements in the $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ directions, with different proportions optimizing for different cost matrices. These projective measurements happen to be the optimal ones in the one-parameter estimation scenario as the SLDs are diagonal in the $\boldsymbol{x}, \boldsymbol{y}$ and $\boldsymbol{z}$ bases respectively (see equation (20)). Note, however, that we have thus far been working in the adjusted coordinate system, where the z axis is aligned with ρ_0. In the next paragraph we analyze the case of a general coordinate system.

3.3. General coordinates

So far we have considered a general state ρ_0 but in order to simplify the analysis, we adjusted our coordinate system such that the z axis was aligned with the Bloch vector of ρ_0. In the last paragraph we saw that in this adjusted coordinate system the optimal trade-off is attained by probabilistic mixtures of the Pauli operators (rotated to the adjusted basis). We are not always free to choose the coordinate system we work in, and it is likely that we would like to optimize our measurement for a cost matrix which is diagonal in a different coordinate system than the one adjusted to ρ_0. We now look at the trade-off surface in a coordinate system which is not aligned with the state, and investigate the measurements which achieve the trade-off surface. This will turn out to be useful for deriving our state-independent result in section 3.4.

Changing the coordinate system rotates the covariance matrix, the SLDs, and the quantum Fisher information matrix as described in section 1.2. In reference [51] it is described how to find the measurement which achieves a desired Fisher information matrix satisfying equation (11) (this is achieved by mixing the projective measurements corresponding to the Bloch vectors which constitute the eigen basis of the desired Fisher information matrix). We used their method to compute the optimal measurements, finding the measurements which result in the inverse of equation (27) for different diagonal costs G. The result is that the optimal measurements no longer belong to an easily characterizable family. As G is varied, the three Bloch vectors describing the projectors of which the measurement is composed travel around the Bloch sphere. The SLD measurements, which when working in the adjusted coordinates could be mixed in different proportions to get variances anywhere on the trade-off surface, no longer play a role. Below we demonstrate that they are far from optimal even in the case of a pure cost matrix (one which assigns all the cost to one parameter), and that in fact, the optimal for such a cost matrix is to measure the corresponding Pauli operator.

In the following we fix an arbitrary coordinate system and test the performance of two families of measurements—one consisting of probabilistic mixtures of the SLD measurements, and the other of the Pauli operators in the chosen coordinate system—and see how they fare compared to the measurements attaining the GM bound. Let $\rho_0 = (\mathbb{I} + z_0\sigma_z)/2$ as before and let $\boldsymbol{\theta}$ be the adjusted coordinate system as before. Any orthogonal coordinate system is related to the adjusted one by a rotation. Let $\tilde{\theta}_i = \sum_j R_{ij}\theta_j$ be the coordinates in which we would like to work, where $R \in O(3)$ is a rotation matrix. The state ρ in these coordinates reads

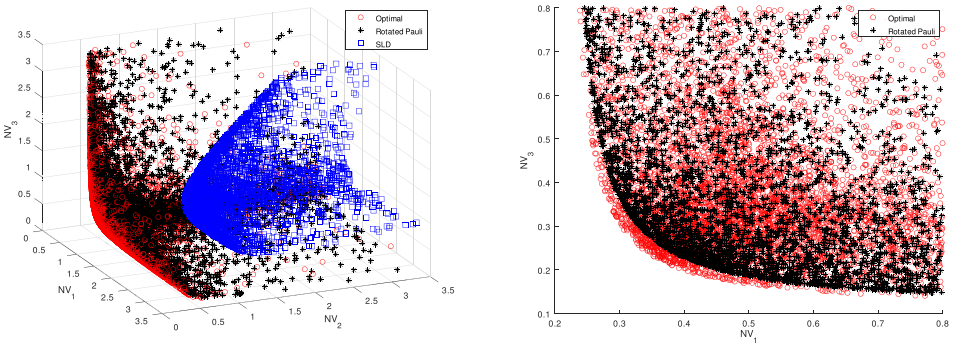
$$\rho(\tilde{\boldsymbol{\theta}}) = \rho_0 + \boldsymbol{\theta}\cdot\boldsymbol{\sigma} = \rho_0 + R^{\mathsf{T}}\tilde{\boldsymbol{\theta}}\cdot\boldsymbol{\sigma} = \rho_0 + \tilde{\boldsymbol{\theta}}\cdot R\boldsymbol{\sigma}.$$

We denote the Pauli matrices in the chosen coordinates by $\tilde{\sigma}_i \coloneqq (R\boldsymbol{\sigma})_i = \sum_j R_{ij}\sigma_j$. As explained in section 1.2, the quantum Fisher information matrix now takes the form

$$\mathfrak{H} = R\mathfrak{H}_{\text{diag}}R^{\mathsf{T}}, \tag{28}$$

where $\mathfrak{H}_{\text{diag}}$ is the QFI matrix in the adjusted coordinates given in equation (21). According to equation (9), the SLDs corresponding to this coordinate system $L_{\tilde{\theta}_i}$ are given by linear combinations of the SLDs in the adjusted coordinates:

$$L_{\tilde{\theta}_i} = \sum_j R_{ij}L_{\theta_j},$$



(a) (b)

Figure 3. The variances of the estimators for three parameters of a qubit in a parameterization rotated with respect to ρ_0 resulting from three families of measurements. The variances are obtained as the diagonal elements of the inverse of the Fisher information matrix $\mathfrak{F}(\mathbf{M})^{-1}$ of a measurement $\mathbf{M}$. Plot (a) shows the variances of the measurements optimizing the expected cost for randomly sampled diagonal cost matrices (red circles), these all lie on the trade-off surface; the variances of random probabilistic mixtures of rotated Pauli measurements (black asterisks); and random probabilistic mixtures of SLD measurements (blue squares). Plot (b) is a projection of the points in plot (a) on to the (NV_1, NV_3) plane. $\rho_0 = (\mathbb{I} + 0.92\sigma_z)/2$ and the rotated coordinates are given by $\tilde{\theta}_i = \sum_j R_{ij}\theta_j$, where θ_i are the coefficients of the Pauli matrices in the coordinate system aligned with ρ_0, and the rotation matrix R is defined by three Euler angles $R = R_x(\alpha)R_y(\beta)R_z(\gamma)$; $\alpha = 25°, \beta = 25°, \gamma = 55°$. It is clearly seen that the SLD measurements perform much worse than the rest, and that the variances of the rotated Pauli measurements lie above, but close to, the trade-off surface. The rotated Pauli measurements approach the trade-off surface far away from the origin as shown in the main text.

Each $L_{\tilde{\theta}_i}$ is diagonal in a basis consisting of two pure states corresponding to two antipodal points on the Bloch sphere. For a general rotation R, the three bases diagonalizing $L_{\tilde{\theta}_i}$, $i = 1, 2, 3$ no longer correspond to three mutually orthogonal lines through the center of the Bloch sphere (because of the non zero $\mathbb{I}$ component in L_z, see equation (20)).

We will now show that the Pauli measurements in the chosen coordinates achieve the optimal expected cost for a pure cost matrix, i.e. $G_i := \boldsymbol{e}_i\boldsymbol{e}_i^\intercal$ (where $\boldsymbol{e_1} = (1, 0, 0)^\intercal$ etc.). The optimal cost according to equation (27) is given by

$$\begin{aligned}\operatorname{Tr} NV_{\text{opt}}(G_i)G_i &= \left[\operatorname{Tr}\sqrt{G_i\mathfrak{H}^{-1}G_i}\right]^2 = (\mathfrak{H}^{-1})_{ii} = \frac{1}{4}\left(R\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1-z_0^2\end{pmatrix}R^\intercal\right)_{ii}\\ &= \frac{1}{4}(R_{i1}^2 + R_{i2}^2 + (1-z_0^2)R_{i3}^2) = \frac{1}{4}(1 - z_0^2R_{i3}^2) \qquad (29)\end{aligned}$$

where the last equality is due to orthogonality of R. The Fisher information for a Pauli measurement is given by

$$\mathfrak{F}\left(\rho_0, \tilde{\boldsymbol{\theta}}, \boldsymbol{P}_{R^\intercal \boldsymbol{e}_i}\right) = \mathfrak{F}\left(\rho_{z_0\hat{z}}, R\boldsymbol{\theta}, \boldsymbol{P}_{R^\intercal \boldsymbol{e}_i}\right) = R\mathfrak{F}\left(\rho_{z_0\hat{z}}, \boldsymbol{\theta}, \boldsymbol{P}_{R^\intercal \boldsymbol{e}_i}\right) R^\intercal$$
$$= \frac{4}{1 - z_0^2 (R^\intercal \boldsymbol{e}_i)_z^2} \boldsymbol{e}_i \boldsymbol{e}_i^\intercal = \frac{4}{1 - z_0^2 R_{i3}^2} \boldsymbol{e}_i \boldsymbol{e}_i^\intercal,$$

where we used equation (26) for the calculation of the Fisher information of a projective measurement in the adjusted coordinates ($\boldsymbol{\theta}$), and the transformation rule for $\mathfrak{F}$ under change of coordinates. Taking probabilistic mixtures of the three Pauli measurements and inverting the resulting Fisher information matrix we obtain the following family of covariance matrices:

$$NV(\alpha, \beta) = \mathfrak{F}(\rho_0, \tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{M}}(\alpha, \beta))^{-1} = \frac{1}{4} \begin{pmatrix} \frac{1 - z_0^2 R_{13}^2}{\alpha} & 0 & 0 \\ 0 & \frac{1 - z_0^2 R_{23}^2}{\beta} & 0 \\ 0 & 0 & \frac{1 - z_0^2 R_{33}^2}{(1 - \alpha - \beta)} \end{pmatrix},$$

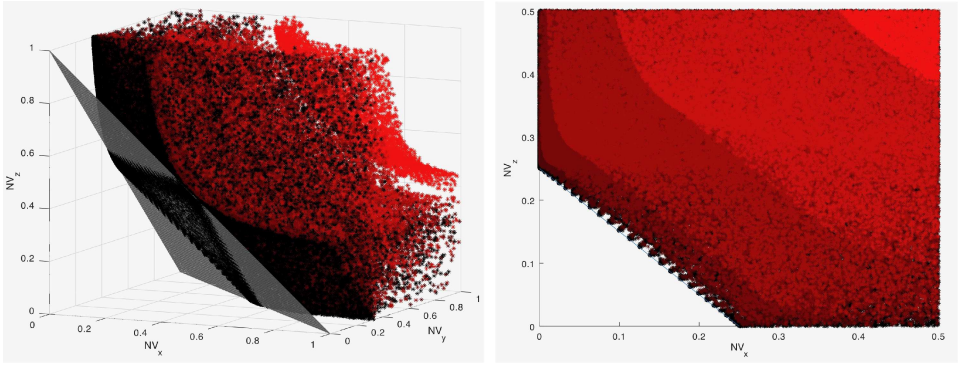
where $\tilde{\boldsymbol{M}}(\alpha, \beta)$ is a probabilistic mixture with proportions $(\alpha, \beta, 1 - \alpha - \beta)$ of the measurements in the Pauli bases corresponding to our chosen coordinates. We see this achieves the optimal cost for pure cost matrices equation (29) for $i = 1, 2, 3$ in the limits $\alpha \to 1$, $\beta \to 1$ and $\alpha, \beta \to 0$ respectively.

For randomly sampled diagonal cost matrices as in equation (22) we computed the optimal covariance matrix using equations (21), (27) and (28). In addition we computed the covariance matrices corresponding to random probabilistic mixtures of the Pauli measurements, and to random probabilistic mixtures of SLD measurements. Figure 3 shows the resulting trade-off surfaces between the variances of the three parameters $\tilde{\theta}_i$. It is clearly seen that the Pauli measurements lie above the optimal surface, and that the SLD measurements preform significantly worse than the other two. This is due to correlations between the parameters in the SLD measurements. In further numerical calculations we performed it was observed that the separation between the Pauli measurements and the optimal measurements is noticeable for states closer to the sphere of pure states ($z_0 > 0.5$), and that it vanished when one of the coordinate axes came close to alignment with ρ_0.

3.4. state-independent trade-off surface

So far we have always considered state-dependent bounds. Indeed, all the bounds we used in order to plot our trade-off surfaces involved explicit dependence on the state ρ_0 (recall that the quantum Fisher information matrix $\mathfrak{H}$ always depends on ρ_0). A state-independent trade-off surface can be obtained as the boundary of the union over all states ρ_0 of the attainable regions—the regions laying above the trade-off surface (equivalently, as the boundary of the intersection of the unattainable regions). To obtain a graphical representation of this state-independent trade-off surface, we would need to plot the trade-off surfaces corresponding to different state ρ_0 all on the same plot, and see what region remains uncovered.

We fix our standard coordinate system to be in terms of the usual Pauli operators $\rho(\boldsymbol{\theta}) = \rho_0 + \sum \theta_i \sigma_i$, and for every state ρ_0 in the Bloch sphere we use equations (21), (27) and (28) to sample points from the trade-off surface corresponding to the GM bound with that state. More precisely, we compute the quantum Fisher information matrix $\mathfrak{H}(z_0)$ in the coordinates aligned



(a) (b)

Figure 4. state-independent trade-off surface. Plot (a) shows points sampled from trade-off surfaces corresponding to different states. The region which is filled with points is the attainable region, and the region which is empty is unattainable for all measurements and all states. The boundary between the regions is the *state-independent trade-off surface*. Plot (a) shows in addition the plane $NV_x + NV_y + NV_z = 1$ which forms part of the trade-off surface, as proven in the main text. Black corresponds to purer states and red to states closer to the center of the Bloch ball. Plot (b) is a projection of plot (a) to the x, z plane (black points were plotted under the red ones) and shows in addition a straight line $NV_z = 0.25 - NV_x$ fitted to the boundary of the region by maximizing the point of intercept. The plots are obtained by random sampling of points from the trade-off surfaces in different coordinate systems, specified by a rotation matrix $R \in O(3)$ which relates the z axes to the Bloch vector of the state ρ_0. As explained in the main text, this is equivalent to varying over all states. The values used for the length of the Bloch vector of ρ_0 are $z_0 \in \{0.5, 0.9, 0.99, 0.999, 1 - 10^{-4}, 1 - 10^{-6}\}$ and the rotations run over a grid of Euler angles: $R = R_x(\alpha)R_y(\beta)R_z(\gamma); \alpha, \beta, \gamma \in \{0, 3^\circ, 6^\circ, \ldots, 360^\circ\}$. The spikes visible on the edge of the covered region in plot (b) are the result of this discrete grid of rotation angles. Pure states ($z_0 = 1$) were avoided for numerical stability.

with ρ_0 (z_0 is the length of the Bloch vector of ρ_0) and then rotate it back to the standard coordinates with the appropriate rotation $R \in O(3)$. We plug the result into equation (27) and plot the diagonal entries of $V_{\text{opt}}(G)$ for randomly sampled diagonal cost matrices G. We do this for a grid of values of $z_0 \in [0, 1]$ and of the angles parameterizing the rotation ($R = R_x(\alpha)R_y(\beta)R_z(\gamma)$, where $R_x(\alpha)$ is a rotation around the x axis by an angle α). This procedure is equivalent to running over a grid of states ρ_0.

The result is shown in figure 4. The figure shows that a non-trivial state-independent trade-off relation holds between the three parameters of a qubit state. This result relies on the GM bound equation (12) and therefore applies whenever the parameters are estimated from the outcomes of separable measurement strategies.

The shape of the trade-off surface in figure 4(a) has features similar to the boundary of the *preparation uncertainty regions* found in [55, figures 6 and 7]. Its projection to the x, z plane shown in figure 4(b) suggests that the following uncertainty relation holds for the rescaled variances[11]:

$$NV(\hat{\theta}_i) + NV(\hat{\theta}_j) \geqslant \frac{1}{4}, \quad i \neq j \in \{x, y, z\}. \tag{30}$$

[11] Recall that in our parameterization θ_i is the deviation of $\langle \sigma_i/2 \rangle$ from its true value, if we were to parametrize the state as $\rho(\boldsymbol{\theta}) = \rho_0 + \boldsymbol{\theta} \cdot \boldsymbol{\sigma}/2$ the lower bounds in equations (30), (34) and (35) would be 4 times bigger.

This bound coincides with the *preparation* uncertainty relation $\Delta(\sigma_x/2)^2 + \Delta(\sigma_z/2)^2 \geqslant 1/4$ proven in [55, 62].

We now prove equation (30). It is enough to prove the case $i = 1, j = 2$, this will become clear from equation (31) below, where we have the freedom to rotate $\mathfrak{H}^{-1}$. We therefore prove $V_1 + V_2 = \mathrm{Tr} VP_2 \geqslant 1/4$, where P_2 is the following matrix:

$$P_2 = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix}.$$

Denote the optimal covariance matrix for a state ρ_0 with Bloch vector of length z_0 and the cost matrix P_2 in a coordinate frame rotated by a rotation $R \in O(3)$ with respect to ρ_0 as $V_{\mathrm{opt}}(P_2, R, z_0)$. As explained above, minimizing the expected cost $\mathrm{Tr}\, VP_2$ over all states ρ_0 is equivalent to minimizing $\mathrm{Tr}\, V_{\mathrm{opt}}(P_2, R, z_0)P_2$ over all choices of coordinate systems (specified by $R \in O(3)$) and all $z_0 \in [0, 1]$. According to equations (27) and (28) we have

$$\mathrm{Tr}\, NV_{\mathrm{opt}}(P_2, R, z_0)P_2 = \left(\mathrm{Tr}\sqrt{\sqrt{P_2}R\mathfrak{H}^{-1}(z_0)R^{\intercal}\sqrt{P_2}}\right)^2. \tag{31}$$

We now proceed to minimize equation (31) over $R \in O(3)$ and $z_0 \in [0, 1]$. First notice that the minimum is always obtained for pure states ($z_0 = 1$) because:

$$\left(\mathrm{Tr}\sqrt{\sqrt{P_2}R\mathfrak{H}^{-1}(z_0)R^{\intercal}\sqrt{P_2}}\right)^2 = \frac{1}{4}\left(\mathrm{Tr}\sqrt{P_2R\left[P_2 + \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1-z_0^2 \end{pmatrix}\right]R^{\intercal}P_2}\right)^2$$

$$\geqslant \frac{1}{4}\left(\mathrm{Tr}\sqrt{P_2RP_2R^{\intercal}P_2}\right)^2 = \mathrm{Tr}NV_{\mathrm{opt}}(P_2, R, z_0 = 1)P_2,$$

where we used the operator monotonicity of the square root function ($A \geqslant B \geqslant 0 \Rightarrow \sqrt{A} \geqslant \sqrt{B}$) going to the second line. We would now like to perform the minimization over $R \in O(3)$. A convenient parameterization of R for this purpose is given by $R(\boldsymbol{u}, \phi)$ where $\boldsymbol{u}$ is a unit vector and ϕ is the angle of rotation. Using the Rodrigues' rotation formula [63], $R(\boldsymbol{u}, \phi)$ is given explicitly by

$$R(\boldsymbol{u}, \phi) = \begin{pmatrix} c + u_x^2(1-c) & u_xu_y(1-c) - u_zs & * \\ u_xu_y(1-c) + u_zs & c + u_y^2(1-c) & * \\ * & * & * \end{pmatrix},$$

where we used the shorthand $c := \cos\phi$ and $s := \sin\phi$ and where * stands in place of entries we will not use. Plugging this into equation (31) and setting $z_0 = 1$ we obtain

$$\mathrm{Tr}\, NV_{\mathrm{opt}}(P_2, R(\boldsymbol{u}, \phi), z_0 = 1)P_2 = \frac{1}{4}\left(\mathrm{Tr}\sqrt{P_2R(\boldsymbol{u}, \phi)P_2R(\boldsymbol{u}, \phi)^{\intercal}P_2}\right)^2$$

$$= \frac{1}{4}\left(\mathrm{Tr}\sqrt{[R(\boldsymbol{u}, \phi)]_{12}[R(\boldsymbol{u}, \phi)]_{12}^{\intercal}}\right)^2 = \frac{1}{4}\mathrm{Tr}[R(\boldsymbol{u}, \phi)]_{12}[R(\boldsymbol{u}, \phi)]_{12}^{\intercal} + \frac{1}{2}\left|\det[R(\boldsymbol{u}, \phi)]_{12}\right|$$

$$= \frac{1}{2}\left[f(c, u_z) + |f(c, u_z)|\right] + \frac{1}{4}(1-c^2)(1-u_z^2)^2,$$

where $[R(\boldsymbol{u}, \phi)]_{12}$ denotes the upper left 2×2 block of $R(\boldsymbol{u}, \phi)$, the function f is given by $f(c, u_z) := c^2 + c(1-c)(1-u_z^2) + u_z^2(1-c^2)$, and we used equation (13) to evaluate $\mathrm{Tr}\sqrt{\cdot}^2$

for a 2×2 matrix. Our original minimization problem

$$\min_{|\boldsymbol{u}|=1,\phi\in[0,2\pi],z_0\in[0,1]} \left(\mathrm{Tr}\sqrt{P_2 R(\boldsymbol{u},\phi)\mathfrak{H}^{-1}(z_0)R(\boldsymbol{u},\phi)^\intercal P_2}\right)^2 \tag{32}$$

was therefore reduced to

$$\min_{u_z\in[-1,1],c\in[-1,1]} \frac{1}{2}\left[f(c,u_z)+|f(c,u_z)|\right]+\frac{1}{4}(1-c^2)(1-u_z^2)^2, \tag{33}$$

which we performed numerically to obtain the value $\frac{1}{4}$.

The two-parameter relations equation (30) fully characterize the attainable region for two parameters, as seen from figure 4(b). As a partial characterization of the shape of the region attainable for all three parameters in figure 4(a) we prove the following[12]:

$$NV(\hat{\theta}_x)+NV(\hat{\theta}_y)+NV(\hat{\theta}_z)\geqslant 1, \tag{34}$$

i.e. that the plane $NV(\hat{\theta}_x)+NV(\hat{\theta}_y)+NV(\hat{\theta}_z)=1$ is a supporting plane of the attainable region, as can be seen in figure 4(a). As before, the minimum of $\mathrm{Tr}\ V_{\mathrm{opt}}(\mathbb{I},R,z_0)\mathbb{I}$ is obtained when $z_0=1$. There is no need to minimize over R as we can use the cyclicity of the trace to eliminate R with $R^\intercal$. We obtain

$$\mathrm{Tr}\ NV_{\mathrm{opt}}(\mathbb{I},R,z_0=1)\mathbb{I}=\left(\mathrm{Tr}\sqrt{R\mathfrak{H}^{-1}(z_0=1)R^\intercal}\right)^2=\frac{1}{4}\left(\mathrm{Tr}\sqrt{RP_2R^\intercal}\right)^2=\frac{1}{4}\left(\mathrm{Tr}\sqrt{P_2}\right)^2=1.$$

We conclude this section by mentioning that the same reasoning can be applied to the Holevo Cramér–Rao bound. Starting from equation (19) with a rotated QFI matrix and a rotated D matrix (equation (18)):

$$\mathrm{Tr}\ NVG\geqslant \mathrm{Tr}\ GR\mathfrak{H}(z_0)^{-1}R^\intercal+\frac{1}{2}\mathrm{Tr}\left|\sqrt{G}R\mathfrak{H}(z_0)^{-1}D\mathfrak{H}(z_0)^{-1}R^\intercal\sqrt{G}\right|,$$

and setting $G=\mathbb{I}$ we obtain

$$\begin{aligned}\mathrm{Tr}\ NV &\geqslant \mathrm{Tr}\ \mathfrak{H}(z_0)^{-1}+\frac{1}{2}\mathrm{Tr}\left|\mathfrak{H}(z_0)^{-1}D\mathfrak{H}(z_0)^{-1}\right| \\ &=\frac{1}{4}(3-z_0^2)+\frac{1}{2}\frac{8z_0}{16}\mathrm{Tr}\left|\begin{pmatrix}0&-1\\1&0\end{pmatrix}\right|=\frac{1}{4}(3-z_0^2+2z_0),\end{aligned}$$

where we used equation (24). This is minimized when $z_0=0$ and we obtain the bound[13]

$$NV(\hat{\theta}_x)+NV(\hat{\theta}_y)+NV(\hat{\theta}_z)\geqslant\frac{3}{4}, \tag{35}$$

which is a state-independent bound implied by the Holevo Cramér–Rao bound and therefore holds for collective measurements. It is saturated in the case of a maximally mixed state. In this case the commutation condition $\mathrm{Tr}\ \rho\left[L_i,L_j\right]=0$ is satisfied, which means that the Holevo

[12] see footnote 10.
[13] see footnote 10.

bound coincides with the SLD-QFI bound [41] and is attainable due to local asymptotic normality [47, 48]. This also shows that the state-independent trade-off surface for estimation using collective measurements is different from the one for separable measurements shown in figure 4, as with collective measurements equation (34) can be violated. It would be interesting to compute the state-independent trade-off surface implied by the Holevo Cramér–Rao bound. We leave this for future works.

4. The Holevo Cramér–Rao bound in the qutrit model

In this section, we compute the Holevo Cramér–Rao bound for a qutrit, i.e. a three level system. We use a model for which the Holevo Cramér–Rao bound is equal to the RLD bound and can, therefore, be computed by equation (19). As in the qubit case, the obtained bound exhibits both trivial and non trivial trade-offs between various parameters.

We compute the SLDs for a parameterization of a state of a 3-level system in terms of the Gell–Mann matrices.

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

The state is parameterized as follows:

$$\rho(\boldsymbol{\theta}) = \exp\left(-\mathrm{i}\sum_{i\in I}\theta_i\lambda_i\right)\rho_d(\theta_3,\theta_8)\exp\left(\mathrm{i}\sum_{i\in I}\theta_i\lambda_i\right),$$

where $I = \{1,2,4,5,6,7\}$ contains only indices of non diagonal λs and ρ_d is a diagonal state parameterized as[14]

$$\rho_d(\theta_3,\theta_8) = \frac{1}{3}\mathbb{I} + (\theta_3+\theta_3^0)\lambda_3 + (\theta_8+\theta_8^0)\lambda_8.$$

By choosing θ_3^0 and θ_8^0 we can specify any diagonal state $\rho_0 \coloneqq \rho_d(0,0)$.

[14] This parameterization is general enough for local estimation because for a given expansion $\rho = \rho_0 + \sum t^k\rho_k$, with t small enough, the following equation $\rho_0 + \sum t^k\rho_k = \exp\left(\mathrm{i}\sum t^k H_k\right)\left(\rho_0 + \sum t^k X_k\right)\exp\left(-\mathrm{i}\sum t^k H_k\right)$ admits a solution with H_k having zero entries on the diagonal and X_k diagonal.

$$\rho_0 = \begin{pmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & k_3 \end{pmatrix},$$

with $k_3 = 1 - k_1 - k_2$. The diagonal entries k_i are related to θ_3^0, θ_8^0 by

$$k_1 - k_2 = 2\theta_3^0, \quad 3(k_1 + k_2) - 2 = 2\sqrt{3}\theta_8^0.$$

The derivatives of ρ, evaluated at $\boldsymbol{\theta} = 0$ are

$$\begin{aligned} \frac{\partial \rho}{\partial \theta_i} &= -\mathrm{i}\,[\lambda_i, \rho_0], \quad i \in \{1,2,4,5,6,7\} \\ \frac{\partial \rho}{\partial \theta_j} &= \lambda_j, \quad j \in \{3,8\}. \end{aligned} \tag{36}$$

In reference [53] the SLDs for a three level system were computed in a more general setting. For simplicity, assume the state is full rank, and using the structure constants of $\mathfrak{su}(3)$ (given in reference [53]), simply verify that the SLDs are given by

$$\begin{aligned} L_1 &= -2\frac{k_1 - k_2}{k_1 + k_2}\lambda_2, \quad L_2 = 2\frac{k_1 - k_2}{k_1 + k_2}\lambda_1 \\ L_4 &= -2\frac{k_1 - k_3}{k_1 + k_3}\lambda_5, \quad L_5 = 2\frac{k_1 - k_3}{k_1 + k_3}\lambda_4 \\ L_6 &= -2\frac{k_2 - k_3}{k_2 + k_3}\lambda_7, \quad L_7 = 2\frac{k_2 - k_3}{k_2 + k_3}\lambda_6 \\ L_3 &= \begin{pmatrix} 1/k_1 & 0 & 0 \\ 0 & -1/k_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad L_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1/k_1 & 0 & 0 \\ 0 & 1/k_2 & 0 \\ 0 & 0 & -2/k_3 \end{pmatrix}. \end{aligned}$$

When all the k_i are different, the model is $\mathcal{D}$-invariant (see footnote 8; this also follows from the $\mathfrak{su}(3)$ structure constants). We compute the expectation values of the commutators of the SLDs $\mathrm{Tr}[L_i, L_j]\rho_0$ to obtain the matrix elements of D (equation (18)). The only non-zero elements are

$$\begin{aligned} |D_{12}| &= 8\frac{(k_1 - k_2)^3}{(k_1 + k_2)^2} \\ |D_{45}| &= 8\frac{(k_1 - k_3)^3}{(k_1 + k_3)^2} \\ |D_{67}| &= 8\frac{(k_2 - k_3)^3}{(k_2 + k_3)^2}. \end{aligned}$$

The quantum Fisher information matrix has only two non-zero entries off from the diagonal ($\mathfrak{H}_{83} = \mathfrak{H}_{38} \neq 0$). Combining these observations we can use equation (19) to understand the trade-offs which the Holevo Cramér–Rao bound exhibit in this model. We treat the matrices appearing in equation (19) as block diagonal. In $\mathfrak{H}^{-1}$ the only block which contains off diagonal terms corresponds to the parameters θ_3, θ_8. Since in this block, D is zero, we do not need to

consider its contribution. In the other blocks (corresponding to (θ_1, θ_2), (θ_4, θ_5) and (θ_6, θ_7)), the result of taking the absolute value of the restriction of D to this block conjugated with a diagonal positive matrix (the restriction of $\sqrt{G}\mathfrak{H}^{-1}$ to the same block), results in a functional dependence of the right-hand side of equation (19) which is a sum of three terms similar to equation (16). More precisely, for $G_{ij} = \delta_{ij} g_i$ with $g_i \geqslant 0$ and $\sum g_i = 1$

$$\begin{aligned}\mathrm{Tr}\; GV(\hat{\boldsymbol{\theta}}) &\geqslant \mathrm{Tr}\; G\mathfrak{H}^{-1} + \frac{1}{2}\mathrm{Tr}\left|\sqrt{G}\mathfrak{H}^{-1}D\mathfrak{H}^{-1}\sqrt{G}\right| \\ &= \mathrm{Tr}\; G\mathfrak{H}^{-1} + \frac{a^2}{2}|D_{12}|\sqrt{g_1 g_2} + \frac{b^2}{2}|D_{45}|\sqrt{g_4 g_5} + \frac{c^2}{2}|D_{67}|\sqrt{g_6 g_7},\end{aligned}$$

where a, b and c are the values of $\mathfrak{H}^{-1}$ in the corresponding blocks (which we do not compute explicitly as we just want to demonstrate the qualitative behavior). The functional dependence of the above on $\{g_i\}$ implies that non-trivial trade-off appears only between pairs of parameters corresponding to the x and y Pauli matrices within each of the 3 $\mathfrak{su}(2)$ sub-algebras of $\mathfrak{su}(3)$, and within each sub-algebra the trade-off is as the RLD bound in figure 2 (trivial trade-off with the diagonal element).

5. Discussion

This paper illustrated the fact that the unsaturability of the quantum Fisher information Cramér–Rao bound for multi parameter estimation gives rise to a rich variety of quantum uncertainty relations in the form of trade-off curves. Those trade-off curves relate to each other the pre-factors c_{ij} of the covariances $V_{ij} = c_{ij}/N$ of the optimal estimators for the unknown parameters $\{\theta_i\}$ in the limit when a large number of copies N of the state are available. This can be seen as a parameter estimation analogue of the quantum Chernoff and Hoeffding bounds [64] in quantum hypothesis testing, where trade-off curves are obtained for the error exponents α_i for the error of the first versus the second kind—scaling as $\exp(-\alpha_i N)$.

Trade-off curves bring into direct view the property which distinguishes quantum multi parameter estimation from its classical statistics counterpart—the unattainability of simultaneous optimal precision. This property was often discussed in the literature, however, we have never seen such trade-off curves plotted for the known tight bounds. In references [65, 66] the trade-off between two parameters is quantified by considering the set of allowed pairs $\left((NV_1\mathfrak{H}_{11})^{-1}, (NV_2\mathfrak{H}_{22})^{-1}\right)$, which for separable measurements is determined by equation (11). Reference [45] provides a comparison between the Holevo Cramér–Rao bound and the GM bound by comparing the bounds they put on the expected cost for *a single* (although state-dependent) cost matrix. In another work, reference [67], the authors present the difference between the regions of variances excluded by bounds on their arithmetic, geometric and harmonic means. What distinguishes our approach from the above works is that to obtain the trade-off curve we use the bound on the expected cost for a family of different costs all at once. This is best illustrated in figure 1 which shows how the trade-off curve is obtained as the point-wise maximum over a family of lines. We can also apply this in the reverse to obtain tight bounds on the expected cost given a convex region of attainable variances, as the latter is determined by its supporting hyper planes.

Trade-offs in quantum parameter estimation belong to the joint-measurement type of uncertainty relations. They show that when we wish to estimate certain parameters by performing a

measurement on a quantum state, increased precision in one parameter will typically come at the cost of increased uncertainty in other parameters.

Investigation of the trade-off surfaces implied by the GM bound led us to our main result: a state-independent uncertainty relation between the three parameters of a qubit system. We provided numerical evidence for this trade-off relation (figure 4(a)) and proved an additive bound equation (34) which forms part of the trade-off surface. In addition, we proved two-parameter additive uncertainty relations equation (30) which coincide with the uncertainty relation for state *preparation* proven in references [55, 62]. We showed that the Holevo Cramér–Rao bound also implies an additive uncertainty relation with a smaller lower bound than in equation (34). Our method for deriving state-independent trade-off surfaces from state-dependent bounds could be applied to the Holevo Cramér–Rao bound for a qubit to obtain a trade-off surface for estimation with collective measurements.

The attainability of the symmetric logarithmic derivative quantum Fisher information (SLD-QFI) bound, which exhibits *classical* (or trivial) trade-off, with collective measurements has recently been shown to be equivalent to the commutation condition $\mathrm{Tr}[L_i, L_j]\rho_0 = 0$, the vanishing of the expectation values of the commutators between all SLDs [41]. The degree to which this fails to be the case has been suggested in reference [58] as a measure of incompatibility between parameters. In section 2.3 we demonstrated this by relating the algebraic form of the Holevo Cramér–Rao bound to the strength of the resulting trade-off curve. Equations (16) and (19) show how as the expectation value $\mathrm{Tr}[L_i, L_j]\rho_0$ approaches zero, the corresponding trade-off curve becomes closer and closer to the trivial one. We have also provided two examples of systems—the qubit (section 3) and qutrit (section 4) models—where the commutation condition is satisfied only between some pairs of parameters, and demonstrated how this reflects in their trade-off surfaces.

The attainable bounds we dealt with in this paper pertain to two different measurement scenarios. The GM bound equation (12), is attainable for qubit ensembles ($d = 2$) with *separable* measurements, whereas the Holevo Cramér–Rao bound equation (17) is attainable for finite dimensional systems with *collective entangled* measurements. From the algebraic form of the GM bound equation (12) and the Holevo Cramér–Rao bound equation (17) the trade-off structure is not immediately visible. In figure 2 we plotted the trade-off surfaces for each of the bounds for the qubit case to show the qualitative difference between the two. The Holevo Cramér–Rao bound allows for higher precision and exhibits non-trivial trade-off only between the x and y parameters, whereas the GM bound—between all three parameters.

The attainability of the Holevo Cramér–Rao bound for finite dimensional systems relies on the theory of quantum local asymptotic normality. As described in [47], in the asymptotic limit the statistical model of a finite dimensional quantum system splits into a product of a classical Gaussian shift model corresponding to the diagonal elements of the density matrix, and independent harmonic oscillator models for the off-diagonal elements. The trade-off surfaces of the Holevo Cramér–Rao bound, which we described for the qubit (figure 2) and qutrit (section 4) systems, are exactly what one would expect to find in the corresponding asymptotic models. In both cases the parameters corresponding to the diagonal components behave like classical systems, i.e. they have trivial trade-off with any other parameter. In the three level system we observe the splitting of the off-diagonal parameters into independent pairs that have non-trivial trade-off within the pair, and trivial trade-off with elements of other pairs. The trade-off structure described in section 4 is therefore generic to finite dimensional systems when collective measurements can be implemented.

Finally, we studied the optimal single copy measurements in the qubit model. We showed that the strategy of measuring different SLD operators on parts of an ensemble of identical

states, which is optimal for the case of a coordinate system aligned with the state ρ_0, is far from optimal in the case of general coordinates. We further demonstrated that measuring the Pauli operators (rotated to the coordinate frame) achieves the optimal cost when all the cost is assigned to one parameter. Our numerical calculations figure 3 further showed that the rotated Pauli measurements are not very far from the optimal for general cost matrices.

Acknowledgments

The authors thank Lorenzo Maccone for the suggestion to look for a state-independent bound from state-dependent trade-off curves, and Borivoje Dakic for helpful discussions. We acknowledge the support of the Austrian Science Fund (FWF) through the Doctoral Program CoQuS and the project I-2526-N27 and the research platform TURIS. IK acknowledges support from the European Commission via Testing the Large-Scale Limit of Quantum Mechanics (TEQ) project (No. 766900).

ORCID iDs

Ilya Kull https://orcid.org/0000-0001-6490-7320
Philippe Allard Guérin https://orcid.org/0000-0002-5044-4799

References

[1] Heisenberg W 1927 *Z. Phys.* **43** 172
[2] Kennard E H 1927 *Z. Phys.* **44** 326
[3] Weyl H 1928 *Gruppentheorie and Quantenmechanik* (Leipzig: Hirzel)
[4] Robertson H P 1929 *Phys. Rev.* **34** 163
[5] Busch P, Heinonen T and Lahti P 2007 *Phys. Rep.* **452** 155
[6] Maassen H and Uffink J B M 1988 *Phys. Rev. Lett.* **60** 1103
[7] Landau H J and Pollak H O 1961 *Bell Syst. Tech.* J. **40** 65
[8] Uffink J B M and Hilgevoord J 1985 *Found. Phys.* **15** 925
[9] Coles P J, Berta M, Tomamichel M and Wehner S 2017 *Rev. Mod. Phys.* **89** 015002
[10] Kraus K 1974 *Foundations of Quantum Mechanics and Ordered Linear Spaces* (*Lect. Notes Phys.*) ed A Hartkämper and H Neumann (Berlin: Springer) vol 29 pp 206–29
[11] Davies E 1976 *Quantum Theory of Open Systems* (New York, NY: Academic)
[12] Kraus K, Böhm A, Dollard J D and Wootters W H (ed) 1983 *States, Effects, and Operations Fundamental Notions of Quantum Theory* (*Lect. Notes Phys.* vol 190) (Berlin: Springer)
[13] Martens H and de Muynck W M 1992 *J. Phys. A: Math. Gen.* **25** 4887
[14] Fuchs C A and Peres A 1996 *Phys. Rev.* A **53** 2038
[15] Ozawa M 2003 *Phys. Lett.* A **318** 21
[16] Ozawa M 2004 *Ann. Phys.* **311** 350
[17] Ozawa M 2005 *J. Opt. B: Quantum Semiclass. Opt.* **7** S672
[18] Hashagen A-L K and Wolf M M 2019 *Ann. Henri Poincaré* **20** 219
[19] Renes J M, Scholz V B and Huber S 2017 *Quantum* **1** 20
[20] Ishikawa S 1991 *Rep. Math. Phys.* **29** 257
[21] Raymer M G 1994 *Am. J. Phys.* **62** 986
[22] Appleby D M 1998 *Int. J. Theor. Phys.* **37** 1491
[23] Busch P 1986 *Phys. Rev.* D **33** 2253
[24] Busch P 1985 *Int. J. Theor. Phys.* **24** 63
[25] Busch P, Lahti P and Werner R F 2014 *J. Math. Phys.* **55** 042111
[26] Busch P, Lahti P and Werner R F 2013 *Phys. Rev. Lett.* **111** 160405
[27] Busch P, Lahti P and Werner R F 2014 *Rev. Mod. Phys.* **86** 1261

[28] Hilgevoord J and Uffink J 2016 *The Stanford Encyclopedia of Philosophy* ed E N Zalta (Stanford, CA: Metaphysics Research Lab, Stanford University) https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=qt-uncertainty&archive=win2016
[29] Ozawa M 2013 arXiv:1308.3540 [quant-ph]
[30] Appleby D 2016 *Entropy* **18** 174
[31] Helstrom C W 1969 *J. Stat. Phys.* **1** 231
[32] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)
[33] Braunstein S L and Caves C M 1994 *Phys. Rev. Lett.* **72** 3439
[34] Barndorff-Nielsen O E and Gill R D 2000 *J. Phys. A: Math. Gen.* **33** 4481
[35] Braunstein S L, Caves C M and Milburn G 1996 *Ann. Phys.* **247** 135
[36] Mandelstam L and Tamm I 1991 *Selected Papers* ed B Bolotovskii *et al* (Berlin: Springer)
[37] Watanabe Y, Sagawa T and Ueda M 2011 *Phys. Rev.* A **84** 042121
[38] Zhu H 2015 *Sci. Rep.* **5** 14317
[39] Shitara T, Kuramochi Y and Ueda M 2016 *Phys. Rev.* A **93** 032134
[40] Cramer H 1946 *Mathematical Methods of Statistics* (Princeton, NJ: Princeton University Press)
[41] Ragy S, Jarzyna M and Demkowicz-Dobrzański R 2016 *Phys. Rev.* A **94** 052108
[42] Liu J, Yuan H, Lu X-M and Wang X 2019 *J. Phys. A: Math. Theor.* **53** 023001
[43] Albarelli F, Barbieri M, Genoni M G and Gianani I 2019 arXiv:1911.12067 [quant-ph]
[44] Sidhu J S and Kok P 2019 arXiv:1907.06628 [quant-ph]
[45] Demkowicz-Dobrzanski R, Gorecki W and Guta M 2020 arXiv:2001.11742 [quant-ph]
[46] Hayashi M and Matsumoto K 2008 *J. Math. Phys.* **49** 102101
[47] Kahn J and Guţă M 2009 *Commun. Math. Phys.* **289** 597
[48] Yamagata K, Fujiwara A and Gill R D 2013 *Ann. Stat.* **41** 2197
[49] Nagaoka H 2005 *Asymptotic Theory of Quantum Statistical Inference* ed M Hayashi (Singapore: World Scientific)
[50] Hayashi M 2005 *Asymptotic Theory of Quantum Statistical Inference* ed M Hayashi (Singapore: World Scientific)
[51] Gill R D and Massar S 2000 *Phys. Rev.* A **61** 042312
[52] Matsumoto K 2002 *J. Phys. A: Math. Gen.* **35** 3111
[53] Ercolessi E and Schiavina M 2013 *Phys. Lett.* A **377** 1996
[54] Peterson W, Birdsall T and Fox W 1954 *Transactions of the IRE Professional Group on Information Theory* **4** 171
[55] Dammeier L, Schwonnek R and Werner R F 2015 *New J. Phys.* **17** 093046
[56] Jozsa R 1994 *J. Mod. Opt.* **41** 2315
[57] Suzuki J 2019 *Entropy* **21** 703
[58] Carollo A, Spagnolo B, Dubkov A A and Valenti D 2019 *J. Stat. Mech.* 094010
[59] Albarelli F, Tsang M and Datta A 2019 arXiv:1911.11036 [quant-ph]
[60] D'Ariano G M, Paris M G and Sacchi M F 2003 *Advances in Imaging and Electron Physics* vol 128 ed P W Hawkes (Amsterdam: Elsevier) pp 205–308
[61] Hou Z, Zhu H, Xiang G-Y, Li C-F and Guo G-C 2016 *npj Quant. Inf.* **2** 16001
[62] Busch P, Lahti P and Werner R F 2014 *Phys. Rev.* A **89** 012129
[63] Goldstein H, Poole C and Safko J 2001 *Classical Mechanics* 3rd edn (Reading, MA: Addison-Wesley)
[64] Audenaert K M R, Nussbaum M, Szkoła A and Verstraete F 2008 *Commun. Math. Phys.* **279** 251
[65] Szczykulska M, Baumgratz T and Datta A 2017 *Quantum Sci. Technol.* **2** 044004
[66] Vidrighin M D, Donati G, Genoni M G, Jin X-M, Kolthammer W S, Kim M S, Datta A, Barbieri M and Walmsley I A 2014 *Nat. Commun.* **5** 3532
[67] Lu X-M, Ma Z and Zhang C 2020 *Phys. Rev.* A **101** 022303