Tight bounds on the simultaneous estimation of incompatible parameters

Jasminder S. Sidhu,† Yingkai Ouyang,∗ Earl T. Campbell, and Pieter Kok
Department of Physics and Astronomy, The University of Sheffield, Sheffield, S3 7RH, UK
(Dated: December 20, 2019)

The estimation of multiple parameters in quantum metrology is important for a vast array of applications in quantum information processing. However, the unattainability of fundamental precision bounds for incompatible observables has greatly diminished the applicability of estimation theory in many practical implementations. The Holevo Cramér-Rao bound (hcrb) provides the most fundamental, simultaneously attainable bound for multi-parameter estimation problems. A general closed form for the hcrb is not known given that it requires a complex optimisation over multiple variables. In this work, we show that the hcrb can be solved analytically for two parameters. For more parameters, we generate a lower bound to the hcrb. Our work greatly reduces the complexity of determining the hcrb to solving a set of linear equations. We apply our results to magnetic field sensing. Our results provide fundamental insight and make significant progress towards the estimation of multiple incompatible observables.

I. INTRODUCTION

Physical quantities such as time, phase, and entanglement cannot be measured directly, but instead must be inferred through indirect measurements. An important category of such indirect measurements is parameter estimation. Quantum metrology describes the quantum mechanical framework that handles this estimation procedure. By recasting the problem as a statistical inference problem, parameter estimation can be associated with fundamental precision bounds. The key question in quantum metrology is what is the fundamental precision bound and how we can achieve it. Early applications of estimation theory focused on single parameter estimation such as phase measurements [1–3]. The ultimate precision bound for a single parameter is the quantum Cramér-Rao bound (qcrb), which was proved by Helstrom and Holevo [4–6]. Multi-parameter quantum metrology extends the single parameter case [7–9], and is of fundamental importance in understanding a variety of practical applications, such as Hamiltonian tomography [10], field sensing [11–13] and imaging [14–17], and distributed sensing [18–20]. A central problem is to determine the optimal measurement strategies that saturate the qcrb [21]. To achieve this, one must assume locally unbiased estimators [22], which is reasonable given large amounts of prior information [23, 24], and with many independent repetitions of the experiment [25]. Several reviews on the topic highlight recent progress in the field [26–28].

Each individual parameter we wish to estimate has an optimal measurement observable. However, when we wish to estimate two or more parameters simultaneously, the corresponding optimal observables may be incompatible. In this case, we can not achieve the optimal precision for each parameter individually. In this case the qcrb matrix bound is generally not simultaneously saturable for all parameters [29–31]. This motivates the search for tighter bounds that can be realised for practical applications of multi-parameter estimation theory. The Holevo Cramér Rao bound (hcrb) encapsulates the difficulties associated with incompatible observables [6]. Operationally, it is the maximum of all existing lower bounds for the error of unbiased measurements [6]. It represents the best precision attainable with collective measurements on an asymptotically large number of identical copies of a quantum state [32–35]. Despite its importance, the hcrb has seen limited use in quantum metrology so far. There are several reasons for this. First, the hcrb is difficult to evaluate given that it is defined through a complex optimisation over a set of observables. Second, implementing collective measurements is generally a difficult task. Nevertheless, applications of the hcrb in metrological tasks do exist. Suzuki found closed form results for parameter estimation with qubits [36], and explored connections between different types of metrological bounds in the special case of two parameter estimation theory. For pure states [7] and displacement estimation with Gaussian states [6], it has been shown that the hcrb is attained by single-copy measurements. The hcrb was also used as a tool to define the precision of state estimation for finite dimensional quantum systems [35]. Bradshaw et al. calculated the hcrb for a joint parameter estimation of a displacement operation on a pure two-mode squeezed probe [37]. Arguably, the hcrb is most relevant in multi-parameter estimation. An increasing number of true multi-parameter estimation protocols has been explored [38–41], and therefore the need for general, attainable bounds on multi-parameter quantum estimation is urgent. Recently, Albarelli et al. have investigated the numerical tractability of calculating the hcrb for the simultaneous estimation of multiple parameters [42]. For finite-dimensional systems, they recast the evaluation of the hcrb as a semi-definite program, which is an optimisation problem that can be efficiently implemented. To date, no general analytic expression for the hcrb is known.

In this paper, we find that it is possible to recast the hcrb as a quadratic program with linear constraints, thereby providing tight bounds for multi-parameter estimation problems. We find that whenever the hcrb is strictly larger than the qcrb, the two-parameter hcrb has a surprisingly simple analytic form that does not require any numerical optimisation. Our analytical solution for the optimal observables that saturate the hcrb allows one to establish analytically the minimum penalty due to

†Electronic address: jsmrsidhu@gmail.com
∗Electronic address: y.ouyang@sheffield.ac.uk
the incompatibility of the observables. Specifically, we generalise attainability constraints for simultaneous multi-parameter estimation problems where the commonly used Cramér-Rao bounds can not be saturated due to incompatibility. The analytic two-parameter hcrb can be considered a generalised quantum uncertainty relation [43]. For more than two parameters, our method does not provide tight bounds but still outperforms the qcrb.

We provide an overview of multi-parameter quantum estimation in section II. We derive the analytic results for the hcrb for two parameters in section III, and extend the treatment to arbitrary number of parameters. In section IV, we discuss the application of the bound to magnetometry. We conclude in section V.

II. MULTI-PARAMETER QUANTUM ESTIMATION

Quantum estimation theory provides fundamental bounds to the estimation precision of physical parameters and the optimal measurements that saturate these limits [21]. We are interested in estimating multiple parameters simultaneously. The prototypical scheme requires that the vector of parameters \( \theta = (\theta_1, \ldots, \theta_d)^\top \in \mathbb{R}^d \) be imprinted on a quantum state \( \rho(\theta) \). Denoting \( \mathbb{H}_d \) as the set of all Hermitian matrices in the Hilbert space of dimension \( d \), we can see that \( \rho(\theta) \) is a positive semidefinite matrix in \( \mathbb{H}_d \) with unit trace. We define measurement operators via a positive operator valued measure (POVM)

\[
\Pi = \left\{ \Pi_\omega \geq 0, \omega \in \Omega \bigg| \sum_{\omega \in \Omega} \Pi_\omega = \mathbb{1}_D \right\},
\]

where \( \mathbb{1}_D \) denotes the identity operator, and \( \Omega \) is the set of measurement outcomes. The outcomes of such a measurement can be used in a function called the estimator \( \hat{\theta} \), which gives an estimate of the parameters. A general estimation scheme requires access to multiple identical copies of the quantum probe state. A separable measurement can be individually applied to each copy of the state to obtain estimates of each parameter separately, whereas a collective measurement can be applied jointly on all copies of the state to acquire a simultaneous estimate of all parameters. The ultimate precision bound is the one that is asymptotically achieved by a sequence of the best collective measurements as the number of copies tends to infinity \([9, 32–34, 44–48]\).

The performance of the estimator \( \hat{\theta} \) under any measurement can be quantified in terms of its mean square estimate matrix (msem)

\[
\Sigma_\theta \left( \Pi, \hat{\theta} \right) = \sum_{\omega \in \Omega^N} p(\omega|\theta) \left( \tilde{\theta}(\omega) - \theta \right) \left( \tilde{\theta}(\omega) - \theta \right)^\top,
\]

where the probability of measurement outcomes is provided by Born’s rule \( p(\omega|\theta) = \text{Tr}[\rho(\theta) \Pi_\omega] \), and \( N \) is the number of independently repeated measurements. The set of estimators are said to be locally unbiased if for all \( \omega \in \Omega \)

\[
\sum_{\omega \in \Omega^N} \left( \tilde{\theta}_j(\omega) - \theta_j \right) p(\omega|\theta) = 0, \quad \sum_{\omega \in \Omega^N} \tilde{\theta}_j(\omega) \frac{\partial p(\omega|\theta)}{\partial \theta_k} = \delta_{jk}.
\]

Under these conditions, the msem is equivalent to the covariance matrix of parameter estimates, and is lower bounded through generalisations of the Cramér-Rao bound from classical statistics

\[
\Sigma_\theta \left( \Pi, \hat{\theta} \right) \geq \mathcal{F} (\rho(\theta), \Pi)^{-1},
\]

where \( \mathcal{F} \) is the classical Fisher information matrix \([5, 6]\). The Fisher information characterises the msem for the best classical data manipulation given a measurement strategy in the asymptotic limit \([25]\). Well known quantum generalisations include the symmetric logarithmic derivative (SLD), \( L_j \in \mathbb{H}_N \) which generates the real symmetric quantum Fisher information (QFIM) \( \mathcal{I}_R^{\text{SL}} \)

\[
\mathcal{I}_R^{\text{SL}} = \text{Re} \left[ \text{Tr}[\rho(\theta)L_j L_k] \right].
\]

This is referred to as the SLD QFIM. Similarly, the right logarithmic derivative (RLD), \( R_j \in \mathbb{H}_N \) induces the complex Hermitian RLD QFIM \( \mathcal{I}_R^{\text{RL}} = \text{Tr}[\rho(\theta) R_j R_k] \)

\[
\mathcal{I}_R^{\text{RL}} = \text{Im} \left[ \text{Tr}[\rho(\theta)L_j L_k] \right] + \| \sqrt{W} \text{Im}(\mathbb{Z})^{-1} \sqrt{W} \|_1.
\]

Here \( W \) is a positive definite weight matrix, and \( \| \cdot \|_1 \) denotes the trace norm of a matrix \([53]\). Throughout, \( \text{Re}(\cdot) \) and \( \text{Im}(\cdot) \) of a matrix denotes taking the real and imaginary part of each matrix element. Nagaoka investigated in detail the relationship between these bounds \([54]\). The central problem in quantum estimation theory is the minimisation of these scalar bounds over the family of probability of distributions defined by quantum measurements.

Helstrom demonstrated that the SLD QCRB is tight for single parameter estimation under the locally unbiased condition of Eq. (3), which is weaker than the unbiased condition. In this case, the qcrb is asymptotically attained through adaptive measurements \([55, 56]\). This attainability does not generally extend to multiple parameter estimations. Specifically, measuring one parameter may affect the precision in the measurement of another parameter, and a trade-off between the msem associated with each parameter exists for any collective measurement \([8]\). Intuitively, any incompatibility among the parameters \( \theta \) prohibits the simultaneous optimal estimation of all parameters. The RLD QCRB was introduced as the minimum of the weighted sum of the mse for each parameter under the unbiased condition. Its achievability was shown for Gaussian states \([6]\), qubits \([33, 57]\), and qudit states \([48]\). However, the RLD QCRB is not always attainable since the optimal estimators derived from the RLD may not correspond to physical POVMs \([58]\).

The problem with saturability of the multiparameter bound was noted by Holevo, who provided the most general quantum extension to the classical Cramér-Rao bound. He introduced
a tighter bound, the hcrb $C_h(\theta)$, defined as the maximum among all existing lower bounds for the mse of unbiased measurements for the estimation of a set of parameters [59]

$$\text{Tr}[W \Sigma_\theta (\Pi, \hat{\theta})] \geq C_h(\theta) \geq \max \{ C_S(\theta) , C_R(\theta) \}. \quad (7)$$

Helstrom [5] and Holevo [59] demonstrated that $C_h(\theta)$ is attainable if the locally unbiased equality constraints in Eq. (3) are satisfied. Consider the Hermitian matrix

$$Z(X) = \begin{pmatrix} \text{Tr}[\rho X_1^2] & \text{Tr}[\rho X_1 X_2] \\ \text{Tr}[\rho X_2 X_1] & \text{Tr}[\rho X_2^2] \end{pmatrix}, \quad (8)$$

with $X_j = \int d\theta (\hat{\delta}_j - \theta_j) \Pi(\theta)$ Hermitian operators. The matrix $Z(X)$ majorises both the inverse of the rld and sld Fisher information matrix [33]. Helstrom and Holevo also showed that if $\Sigma_\theta \geq W$, then $\text{Tr}[\Sigma_\theta] \geq \text{Tr}[W] + \| \text{Im}W \|$. Together with the rld cr inequality in Eq. (6), the Holevo bound in Eq. (7) can be written as the solution to the following minimisation [54]

$$\begin{aligned}
&\text{minimize } \left\{ \begin{array}{c}
X_1, X_2 \\
\text{subject to } \\
\text{Tr}[\rho X_1] = 0, \\
\text{Tr}[\rho X_2] = \delta_{jk},
\end{array} \right. \\
&\text{Tr}[W \text{Re}Z] + \| W \text{Im}Z \|_1,
\end{aligned} \quad (9)$$

where $\hat{\delta}_j \equiv \partial / \partial \theta_j$ and the constraints are the local unbiased conditions of Eq. (3). By minimising over only the first term in the objective function of Eq. (9), we obtain the sld qcrb [54]

$$C_S(\theta)(\rho, W) = \text{Tr}[W(I^S)^{-1}] = \min_{X_1, X_2} \text{Tr}[W \text{Re}Z]. \quad (10)$$

This shows that the hcrb is indeed a tighter bound than the sld qcrb, since the second term in Eq. (9) is non-negative [60]. Further, note that the hcrb is not defined explicitly in terms of a closed form for a given statistical model. This is in contrast to the classical case, where the Fisher information can be readily determined from a given statistical model. Recently, Tsang demonstrated that the hcrb is upper bounded by a quantity that is three times the sld-qcrb, such that [61]

$$\max \{ C_S(\theta) , C_R(\theta) \} \leq C_h(\theta) \leq 3C_S(\theta). \quad (11)$$

In this paper, we provide an analytic form for the hcrb.

The hcrb is the best asymptotically achievable bound under the conditions stated in Refs [32–34, 48]. Both inequalities in Eq. (7) can be tight [60]. For instance, consider the skew-symmetric matrix $\text{Im}(\text{Tr}[L_j L_k \rho \theta])$. Then

$$\text{Im}(\text{Tr}[L_j L_k \rho \theta]) = 0, \quad (12)$$

we have $C_h(\rho, \theta, W) = C_S(\rho, \theta, W)$ [8]. This condition is referred to as the weak commutativity criterion [62], and when it is fulfilled the qcrb is a good proxy for the hcrb. In the next section, we show how we can use methods from optimisation theory to address the minimisation over several Hermitian operators in the case where the weak commutativity criterion is not fulfilled.

## III. HOLEVO CRAMÉR-RAO BOUND

In this section, we present algorithms for calculating the Holevo Cramér-Rao bound. To facilitate the derivation, we construct a bound on $C_h^*$ such that

$$C_h = \max \{ C_h^*, C_S \}, \quad (13)$$

and from Eq. (7) we see that $C_h^* \geq C_R$ whenever $C_h$ is a tighter bound than $C_S$ (in other words, whenever $C_R$ is relevant). We first derive an analytic expression for $C_h^*$ for two parameters in section III A. We demonstrate that the optimisation problem can be solved using the method of Lagrange multipliers. This reduces the problem of evaluating $C_h^*$ to that of solving two sets of linear equations. In section III B, we derive a lower bound on $C_h^*$ for more than two parameters.

### A. Two-parameter setting

We first consider the hcrb for two parameters $\theta = (\theta_1, \theta_2)^T$. To obtain an analytic expression, we must define the weight matrix $W$ for the scalar bound. For simplicity, we use the identity weight matrix. We determine $C_h^*$ for two parameter estimation schemes using optimisation theory [63]. We want to solve the minimisation in Eq. (9), which is convex but not quadratic. We will manipulate Eq. (9) into a quadratic form in the variables $X_1$ and $X_2$. Once this is done, such an optimisation problem can be solved analytically using the method of Lagrange multipliers.

Choosing $Y = X_1 + iX_2$, Eq. (9) can be written as an optimisation program (see appendix B)

$$\begin{aligned}
&\text{minimize } \left\{ \begin{array}{c}
Y, t \\
\text{subject to } \\
\text{Tr}[YP^\dagger] \leq t, \\
\text{Tr}[Y^\dagger P Y] \leq t, \\
\text{Tr}[\rho Y] = 0, \\
\text{Tr}[\partial \rho Y] = 1, \\
\text{Tr}[\partial \rho Y] = i.
\end{array} \right.
\end{aligned} \quad (14)$$

Note that by considering both the real and imaginary parts of the above equality constraints, the actual number of real-valued equality constraints is six, which is consistent with the number of equality constraints corresponding to the minimisation in Eq. (9). Here $Y$ is optimised over all complex matrices of dimension $D$, and is in general not a Hermitian matrix. By mapping $Y$ and $t$ into a real vector $x$, we cast this optimisation program into the standard form of

$$\begin{aligned}
&\min_{x} \{ f(x) : c_i(x) \leq 0, h_i(x) = 0 \}, \\
&\quad \left\{ \begin{array}{c}
f(x) \\
c_i(x) \leq 0 \\
h_i(x) = 0
\end{array} \right.
\end{aligned} \quad (15)$$

where $f(x)$ is a real linear objective function, while $h_i(x)$ and $c_i(x)$ are the corresponding equality and inequality constraints functions that must also be real. Eq. (14) is a convex program, since its equality constraints are linear and its inequality constraints are quadratic and convex. To check whether we can use optimality conditions from optimisation theory, we check whether Slater’s constraint qualification holds. This amounts to checking that all the inequality constraints in Eq. (14) can strictly hold. Since $t$ can be arbitrarily large, this indeed is the case. The optimality conditions for a continuous optimisation
Algorithm 1 \( \{\text{hcrb},X_1,X_2\} = C_H^*(\rho,\partial_1\rho,\partial_2\rho) \)

1: \( A_1 \leftarrow \rho/2 \)
2: \( A_2 \leftarrow \partial_1\rho/2 \)
3: \( A_3 \leftarrow \partial_2\rho/2 \)
4: \( A_4 \leftarrow -iA_1 \)
5: \( A_5 \leftarrow -iA_2 \)
6: \( A_6 \leftarrow -iA_3 \)
7: \( b \leftarrow [0,1,0,0,0,1]^T \)
8: \( Q_1, Q_2 \leftarrow \text{size 6 zeros matrix} \)
9: for all \( j,k \) from 1 to 6 do
10: \( [Q_1]_{jk} = \text{Tr}[A^j_k\rho^{-1}A_k] \)
11: \( [Q_2]_{jk} = \text{Tr}[A_k\rho^{-1}A_k^j] \)
12: \( R_1 \leftarrow \text{Re}(Q_1) \)
13: \( R_2 \leftarrow \text{Re}(Q_2) \)
14: Solve for \( y = [y_1,\ldots,y_6]^T \) in \( 2R_1y + b = 0 \)
15: Solve for \( z = [z_1,\ldots,z_6]^T \) in \( 2R_2z + b = 0 \)
16: \( v_1 \leftarrow -y^TQ_1y - b^Ty \)
17: \( v_2 \leftarrow -z^TQ_2z - b^Tz \)
18: if \( v_1 > v_2 \) then
19: \( \text{hcrb} \leftarrow v_1 \)
20: \( A \leftarrow y_1A_1 + \ldots + y_6A_6 \)
21: \( Y \leftarrow -A^\dagger\rho^{-1}A^1 \)
22: else
23: \( \text{hcrb} \leftarrow v_2 \)
24: \( A \leftarrow z_1A_1 + \ldots + z_6A_6 \)
25: \( Y \leftarrow -\rho^{-1}A^1 \)
26: \( X_1 \leftarrow (Y + Y^\dagger)/2 \)
27: \( X_2 \leftarrow (Y - Y^\dagger)/(2i) \)
28: Return \( \{\text{hcrb},X_1,X_2\} \)

FIG. 1: Pseudocode for generating the analytic form of the two-parameter hcrb and its associated optimal measurement observables \( X_1 \) and \( X_2 \). Note that this algorithm depends only on the state \( \rho \) and its two derivatives \( \partial_1\rho \) and \( \partial_2\rho \).

Due to the duality principle in optimisation theory [63], we may equivalently view the optimisation by considering the Lagrange dual function \( g(\lambda,z) = \inf_x L(x,\lambda,z) \) of Eq. (16). Since the Lagrangian \( L \) is quadratic in \( x \), the Lagrange dual can be found analytically by an unconstrained minimisation of the Lagrangian with respect to \( x \) for fixed values of the dual variables \( \lambda \) and \( z \) [63]. Due to the structure of the Lagrangian in Eq. (17), the Lagrange dual is never unbounded from below whenever \( u + v = 1 \). Hence, maximising the Lagrange dual function corresponds to an unconstrained maximisation problem. Since the Lagrange dual is also a quadratic function in terms of its dual variables \( z \), it can be easily maximised exactly with respect to \( z \).

Note that our Lagrange dual is not a quadratic function with respect to \( \lambda = (u,v) \). This means that we cannot solve it analytically as it stands. Instead, we find a lower bound to \( C_H \) that is associated with \( (u,v) = (0,1) \) or \( (u,v) = (1,0) \). In those two cases, the Lagrangian in Eq. (17) simplifies. In these cases, we may no longer have \( C_H \) but we do have a lower bound \( C_H^* \leq C_H \). However, \( C_H = C_H^* \) if \( C_H^* > C_B \), as stated in Eq. (13). This arises as a necessary consequence of removing the absolute value in the objective function of the optimisation program in Eq. (9), which makes the optimisation problem continuous. Hence, we obtain the maximum of two analytical formulas, which is itself analytical. When \( C_H^* \) is equal to \( C_B \), the Lagrange dual becomes a lower bound to \( C_H \) and bears no physical significance. As shown in appendix B, the resulting analytic expression for \( C_H^* \) can be determined through finding a \( z \in \mathbb{R}^6 \) that solves

\[
2\text{Re}(Q_j)z + b = 0, \quad j = 1,2
\]

where \( Q_j \) has the matrix elements

\[
[Q_1]_{jk} = \text{Tr}[A^j_k\rho^{-1}A_k] \quad \text{and} \quad [Q_2]_{jk} = \text{Tr}[A_k\rho^{-1}A_k^j].
\]

The matrices \( Q_j \) are full rank when the derivatives \( \partial_1\rho \) and \( \partial_2\rho \) are linearly independent. We collect the result of this optimisation in the following theorem:

**Theorem 1.** Let \( \partial_1\rho \) and \( \partial_2\rho \) be linearly independent. Then \( C_H^* \) for two parameters is given by

\[
C_H^* = \max_{j=1,2} \left( \frac{b^T\text{Re}(Q_j)^{-1}b}{2} - \frac{b^T\text{Re}(Q_j)^{-1}Q_j\text{Re}(Q_j)^{-1}b}{4} \right).
\]
This theorem is the main result of the paper: it gives a procedure for finding the analytic form of the HCRB for two-parameter estimation. Fig. 1 shows the pseudocode for this procedure.

Next, we need to establish whether $C_{h} > C_{S}$. As shown in appendix B.2, when $C_{h} > C_{S}$ the KKT conditions require that either $(u, v) = (0, 1)$ or $(u, v) = (1, 0)$. Hence, there are two choices for $Y$ that minimise the Lagrangian in Eq. (17), corresponding to the two choices for $(u, v)$:

\begin{align}
(u, v) &= (1, 0) : \quad Y = -A^{T} \rho^{-1}, \\
(u, v) &= (0, 1) : \quad Y = -\rho^{-1} A^{T},
\end{align}

(22, 23)

corresponding to the choice $Q_{1}$ and $Q_{2}$, respectively. Then, using Eq. (18) and the optimised values for $Z$, we construct the analytic form for the observables $X_{1}$ and $X_{2}$. Since the optimisation is only over 6 parameters, it is straightforward to find an analytic form of $C_{h}$. The procedure is shown algorithmically in Fig. 1.

From the expressions for $X_{1}$ and $X_{2}$, we can check whether the assumption that $C_{h} > C_{S}$ holds. From duality theory, if we substitute $X_{1}$ and $X_{2}$ into the objective function of Eq. (9), we must obtain a value that coincides with either the $(u, v) = (1, 0)$ value or the $(u, v) = (0, 1)$ value. If not, then the assumption that $C_{h} > C_{S}$ must be violated, and we have in fact $C_{h} = C_{S}$, corresponding to

\[|\text{Tr}[\rho X_{1} X_{2}] - \text{Tr}[\rho X_{2} X_{1}]| = 0,\]

(24)

for the optimal $X_{1}$ and $X_{2}$. This procedure gives either $C_{h}^{*}$ or $C_{S}$ as optimal. In the latter case, we have to find $C_{S}$ separately by calculating the SLD QPIM.

### B. Lower bound in the multi-parameter setting

For more than two parameters, we can also use the method of Lagrange multipliers to find $C_{h}^{*}$. However, this method is considerably more involved than the two-parameter case. In the two-parameter case, we could obtain a simple quadratic expression for $\text{Re} \text{Tr}[Z] + \|\text{Im} Z\|_{1}$ that appears in the objective function of Eq. (9). However, for the corresponding generalisation to multiple parameters, $\text{Re} \text{Tr}[Z] + \|\text{Im} Z\|_{1}$ is no longer a quadratic form in the variables $X_{j}$. For example, for three parameters $Z$ takes the form

\[Z(X) = \begin{pmatrix}
\text{Tr}[\rho X_{1}^{2}] & \text{Tr}[\rho X_{1} X_{2}] & \text{Tr}[\rho X_{1} X_{3}] \\
\text{Tr}[\rho X_{2} X_{1}] & \text{Tr}[\rho X_{2}^{2}] & \text{Tr}[\rho X_{2} X_{3}] \\
\text{Tr}[\rho X_{3} X_{1}] & \text{Tr}[\rho X_{3} X_{2}] & \text{Tr}[\rho X_{3}^{2}]
\end{pmatrix}.
\]

(25)

The trace norm of $\text{Im} Z$ is related to the eigenvalues of $\text{Im} Z$, and the eigenvalues of a $3 \times 3$ matrix involve a cubic equation. This renders evaluating the trace norm incompatible with our methodology. To address this, we obtain a lower bound to $\|\text{Im} Z\|_{1}$ that allows $\text{Re} \text{Tr}[Z] + \|\text{Im} Z\|_{1}$ to be written as a quadratic form. This yields an optimisation problem whose optimal value is a lower bound for $C_{h}^{*}$, and which is given by

\[\min \{ t : \text{Tr}[\rho X_{j}] = 0, \text{Tr} [\partial_{j} \rho X_{k}] = \delta_{jk}, V_{\alpha} \leq t \}.
\]

(26)

The minimisation is performed over Hermitian matrices $X_{1}, \ldots, X_{d}$ and $t$. The indices $j, k$ run over $\{1, \ldots, d\}$. The inequality constraint $V_{\alpha}$ is a function of a binary string $\alpha$ such that

\[V_{\alpha} = \frac{1}{2} \sum_{j=1}^{d} \text{Tr}[(X_{j} + (-1)^{\alpha_{j}} i X_{j+1}) \rho (X_{j} + (-1)^{\alpha_{j}} i X_{j+1})^{\dagger}],
\]

(27)

with $X_{d+1} = X_{1}$. The inequality constraints $V_{\alpha}$ arise from the structure of our lower bound on the trace norm of $\text{Im} Z$ (see appendix C). By substituting $Y_{j} = X_{j} + i X_{j+1}$, we can write the matrices $X_{j}$ in terms of the matrices $Y_{j}$, as before. We now interpret the $Y_{j}$ as arbitrary complex matrices of size $n$, and impose Hermiticity conditions for the corresponding $X_{j}$ matrices.

The Lagrangian of such an optimisation problem is a function of the complex matrices $\{Y_{1}, \ldots, Y_{d}\}$, and also a function of its Lagrange multipliers. Its Lagrange multipliers are given by the non-negative multipliers $\xi \in \mathbb{R}^{2d}$ for the inequality constraints, $\zeta \in \mathbb{R}^{d(d+1)}$ for the equality constraints, and Hermitian multipliers $\xi_{1}, \ldots, \xi_{d}$ for the Hermitian constraints. Most importantly, the inequality constraints can be satisfied strictly, so Slater’s constraint qualification holds, and we can use the KKT to determine the optimality conditions for Eq. (27).

We minimise the Lagrangian constructed from the optimisation problem in Eq. (26). Since the Lagrangian is a convex quadratic form in the variables $Y_{1}, \ldots, Y_{d}$, it can be minimised exactly. When this is done, we obtain the Lagrange dual function, which only depends on the Lagrange multipliers $\zeta$ and $\xi_{1}, \ldots, \xi_{d}$. The Lagrange dual function always gives a lower bound for the primal optimisation problem.

While the Lagrange dual is quadratic in $\zeta$ and $\xi_{1}, \ldots, \xi_{d}$, it is not quadratic in $\xi$. By minimising the Lagrangian over $\tau$ and using the KKT conditions, we conclude as before that the sum of the components in $\zeta$ is 1. We obtain a lower bound for the Lagrange dual by maximising over a discrete set of feasible Lagrange multipliers $\zeta$, which corresponds to the tightness of the constraints $V_{\alpha} \leq t$. Thus, we created a quadratic optimisation problem for three or more parameters that leads to a lower bound on $C_{h}$. However, there is no guarantee that this lower bound is tight.

Next, we study the Lagrange dual function. By carefully choosing $\zeta$, the Lagrangian is quadratic in $\{Y_{1}, \ldots, Y_{d}\}$ and can be minimised individually for each $Y_{j}$. The coefficients for $Y_{j}$ in the Lagrangian are given by $\Gamma_{j}$, where

\[\Gamma_{j} = \sum_{k=1}^{d} T_{k,j} \left( \sum_{l=0}^{d} Z_{l,k} \rho_{l} + i \xi_{l} \right),
\]

(28)

where $\rho_{0} = \rho$, $\rho_{j} = \partial_{j,\rho}$ for $j = 1, 2$, and $T_{k,j}$ are matrix elements that relate the $Y_{j}$ to the $X_{k}$. Then the optimal value for the Lagrange multipliers can be obtained by maximising the Lagrange dual functions

\[g_{\alpha} = -\sum_{j=1}^{d} Z_{i,j} - \frac{d}{2} \sum_{j=1}^{d} \frac{\delta_{0,\alpha_{j}} \text{Tr}[\Gamma_{j} \rho^{-1} \Gamma_{j}^{\dagger}]}{\delta_{1,\alpha_{j}} \text{Tr}[\Gamma_{j}^{\dagger} \rho^{-1} \Gamma_{j}]}.
\]

(29)
Algorithm 2 \( v = g_{\alpha}(Z, \xi_1, \ldots, \xi_d) \)
\begin{enumerate}
\item \( S \leftarrow \text{size} \ d \ \text{zeros matrix} \)
\item for all \( j = 1 \) to \( d \) do
\item \( S_{jj} = 1 \)
\item if \( j \neq d \) then
\item \( k \leftarrow j + 1 \)
\item else
\item \( k \leftarrow 1 \)
\item \( S_{jk} = i \)
\item \( T \leftarrow S^{-1} \)
\item \( v \leftarrow \sum_{\ell=j}^{d} |Z|_{jj} \)
\item \( v \leftarrow v - \text{Tr}[\Gamma_j \rho^{-1} \Gamma_j^\dagger]/2 \)
\item \( v \leftarrow v - \text{Tr}[\Gamma_j^\dagger \rho^{-1} \Gamma_j]/2 \)
\end{enumerate}
Return \( v \)

Algorithm 3 \( \{\text{lowerhcrb}, X_1, \ldots, X_d\} = C_{\theta}(\rho, \partial_1 \rho, \ldots, \partial_d \rho) \)
\begin{enumerate}
\item \( \rho_0 \leftarrow \rho \)
\item \( S \leftarrow \text{size} \ d \ \text{zeros matrix} \)
\item for all \( j = 1 \) to \( d \) do
\item \( \rho_j \leftarrow \partial_j \rho \)
\item \( S_{jj} = 1 \)
\item if \( j \neq d \) then
\item \( k \leftarrow j + 1 \)
\item else
\item \( k \leftarrow 1 \)
\item \( S_{jk} = i \)
\item \( T \leftarrow S^{-1} \)
\item for all \( j = 1 \) to \( d \) do
\item \( Y_j \leftarrow -\rho^{-1} \Gamma_j \)
\item for all \( k = 1 \) to \( d \) do
\item \( X_k \leftarrow \sum_{\ell=0}^{d} T_{k,j} Y_j \)
\end{enumerate}
Return \( \{\text{lowerhcrb}, X_1, \ldots, X_d\} \)

IV. APPLICATION: MAGNETOMETRY

We use the technique introduced in this paper to consider magnetic field sensing, which has important technological applications in navigation, position tracking, and imaging [64]. We apply our method of finding the hcrb to the estimation of a magnetic field \( \mathbf{B} = (B_x, B_y, B_z) \) in three dimensions. Quantum magnetometry is an important application of quantum metrology, and is essential for detecting defects and realising compact magnetic resonance imaging scanners [65]. Estimating each component individually allows us to attain the quantum limit [21], and this has been demonstrated in several studies [66, 67]. However, in many practical applications, knowledge of multiple parameters is required simultaneously, and we must consider joint estimation strategies.

The three parameters of interest are \( \theta = (\theta_1, \theta_2, \theta_3)^T \) appear in the single spin Hamiltonian \( H_j(\theta) = \theta \cdot S_j \), where \( S_j \) is the spin operator for the \( j \)th spin, and we included all coupling constants in \( \theta \). Local depolarising noise, described by the

![FIG. 4: Estimation precision of two directional components of the a magnetic field of size 2 to 10 qubits.](image-url)

FIG. 4: Estimation precision of two directional components of the a magnetic field of size 2 to 10 qubits.

**Algorithm 2**

Algorithm 1

**Algorithm 2**

Algorithm 3

**Algorithm 3**

**Theorem 2.** Let \( d \geq 3 \), \( z \in \mathbb{R}^{d(d+1)} \) and \( \xi_1, \ldots, \xi_d \) are Hermitian matrices. Then \( C_H \) is lower-bounded by

\[
C_H > C_{\alpha}^* = \max_{\alpha \in (0,1)^d} \max_{Z, \xi_1, \ldots, \xi_d} g_{\alpha},
\]

where \( g_{\alpha} \) is given by Eq. (29).

This optimisation problem can be solved exactly using a single step of Newton’s method. It requires the input state \( \rho \) and its derivatives \( \partial_j \rho \). The algorithm to implement this lower bound is illustrated in Fig. 3.
single spin CPTP map

\[ D_{GJ}[\rho] = (1 - g)\rho + g \frac{1}{2} I. \]  

provides a general description for a noisy environment, where \( g \) denotes the depolarisation magnitude and takes values between 0 and 1. The parameters \( \theta \) are imprinted on the probe state via the unitary evolution \( \hat{U} = \exp[-i\hat{H}_f(\theta)] \). For our example, we assume that the magnetic field in the \( z \)-direction is known, and we therefore wish to estimate the two parameters \( B_x \) and \( B_y \). We choose an identity weight matrix to equally prioritise each parameter into a weighted scalar mean square error. We consider three families of \( n \)-spin probe states, namely the traditional\( g_n \) states for single-parameter estimation, the modified 3D-GHZ states introduced by Baumgratz and Datta [13], and the\( g_n \) states introduced by Ouyang in the context of quantum error correction [68].

First, the 3D-GHZ state can be written as

\[ |\psi_{n}^{3D-GHZ}\rangle = \frac{1}{N} \sum_{j=1}^{3} (|\phi_+^j\rangle^{\otimes n} + |\phi_-^j\rangle^{\otimes n}), \]  

where \( n \) is the total number of spins, \( N \) is the normalisation constant of the state and \( |\phi_+^j\rangle \) are the eigenvectors corresponding to the \( +1 \) eigenvectors of the \( j \)-th spin matrix. The evolved state then becomes

\[ \rho(\theta) = \hat{U}(\theta) D_{GJ}[|\psi_{n}^{3D-GHZ}\rangle \langle \psi_{n}^{3D-GHZ}|] \hat{U}(\theta)^\dagger. \]  

We illustrate how the hCqRB and the qCqRB change with the number of probe qubits \( n \) for different depolarising channel strengths \( g \) in Fig. 4. We observe that the hCqRB is indeed tighter than the qCqRB. Both variance bounds increase with an increasing depolarising probability of the depolarising channel, as expected. The 3D-GHZ state attains the Heisenberg precision scaling for the noiseless case.

Second, we consider the class of\( g_n \) states that are robust to a constant amount of erasure and dephasing [69]:

\[ |\varphi_{w}\rangle = \frac{1}{2} \sum_{j=0}^{2} \sqrt{\binom{2}{j}} |D_{GJ}^w\rangle, \]  

where, for every \( w = 0, \ldots, n \), the Dicke state \( |D_{GJ}^w\rangle \) is a uniform superposition over all computation basis states \( |x_1\rangle \otimes \cdots \otimes |x_n\rangle \) with Hamming weight \( w \). Since \( n = 2G \), where \( G \) is related to the number of bit-flip errors that can be corrected, we present results for the\( g_n \) states for which \( n \) is even. These are shown in Fig. 5, and compared with traditional\( g_n \) states and 3D-GHZ states. The traditional\( g_n \) states give a worse estimation for larger qubit number at constant depolarisation rate, as is well-known. The\( g_n \) states perform similarly to the 3D-GHZ states.

Finally, we use our formalism to determine the optimal \( n \)-qubit observables \( X_1 \) and \( X_2 \) that attain the hCqRB using the 3D-GHZ states. Unlike the single qubit estimation case, analytic solutions to these observables are challenging and the dimension of \( X_j \) scales as \( 2^n \). Instead, we numerically determine their structure, shown in Fig. 6. We plot the real and imaginary parts of the matrices \( X_1 \) and \( X_2 \), and the Hermiticity of the observables is clearly observed.

V. CONCLUSIONS AND DISCUSSIONS

Quantum metrology promises practical near term quantum technologies. Experimental developments in sensing are demonstrating early theoretical results and advancements in estimation theory. On the theoretical front, one prominent limitation that remains is the estimation of multiple non-compatible observables. Specifically, the optimal strategy to define the fundamental limits to precision estimates and their attainab-
ity is not known. Efforts to estimate multiple non-compatible observables have largely been focused on approaching the fundamental quantum Cramér-Rao bound (qcrb). This has led to efforts to devise non trivial measurement schemes that approach the qcrb. An alternative approach is to focus on the tighter Holevo Cramér-Rao bound (hcrb), which is physically attainable. However, the hcrb is difficult to evaluate since it involves a difficult optimisation over two observables. This has limited its application in quantum estimation theory.

In this paper, we have made significant progress in analytically solving the hcrb for two-parameter estimation problems, and providing bounds for larger number of parameters. In the two-parameter case, we reduce the complexity of the optimisation procedure to that of solving a set of linear equations, which can be easily solved using most numerical software packages. We also provide analytic expressions for the optimal positive operator valued measurements (povm). Our results readily apply to a large range of physical applications. This will provide deeper insight into the role of quantum measurements in quantum sensing, and help continue the drive of realising quantum technologies.

We illustrate an application of our results by considering the estimation of a magnetic field using noisy multi-qubit probe states. A recent numerical study by Albarelli et al. demonstrated the necessity of using the hcrb over the qcrb, based on a violation of the weak commutation condition [42]. Here, we provide further insight into the role that the hcrb plays in quantum estimation theory. We provide conditions for when this bound is tighter than the sld qcrb (or the Helstrom bound) and provide the corresponding optimal measurement observables.

There are several clear extensions of our work that can be readily addressed. The first would be to use the analytic expressions that we derive to provide further insight into more protocols in estimation theory. We hope that this will help to drive the wave for experimental validation. A second line of work would consider an extension of the Holevo bound to parameters with arbitrary choice of weight matrices. In this work we have considered unit weight matrices, which was motivated through placing equal importance to each parameter. A more general weight matrix would provide a more general bound. A final line of work would consider the optimal implementation of the general measurements that were derived in this work. This would provide an immediate access to the tighter hcrb through experimental implementation.

VI. ACKNOWLEDGMENTS

J.S.S and P.K. acknowledge the support of EPSRC via the Quantum Communications Hub (EP/M013472/1). Y.O. and E.C. acknowledge the support of EPSRC through grant number EP/M024261/1.

Appendix A: Matrix calculus

We prove some elementary facts about matrix calculus that we use repeatedly in our analysis of the turning points of the Lagrangian functions that occur throughout the manuscript.

We begin by defining some notations. Since a complex matrix of size $n$ is a map from $\mathbb{C}^n$ to $\mathbb{C}^n$, we use $L(\mathbb{C}^n)$ to denote the set of all complex matrices of size $n$. Here, the notation $L(\mathbb{C}^n)$ reflects the fact that a matrix is a linear mapping that is an automorphism on $\mathbb{C}^n$. At times we are interested in matrices that are also Hermitian, which means that they are equal to their complex conjugates. In this scenario, we use $\mathbb{H}_n$ to denote the set of all complex matrices that are also Hermitian. Clearly for instance, $\mathbb{H}_n$ is a strict subset of $L(\mathbb{C}^n)$.

Now let $f : L(\mathbb{C}^n) \to \mathbb{C}$ denote a function that maps a complex matrix to a complex scalar. If $f(Y)$ is differentiable at $Y$ in the direction $H$, we use

$$\nabla_{Y,H} f(Y) = \lim_{h \to 0} \frac{f(Y + hH) - f(Y)}{h} \quad (A1)$$

to denote the Fréchet derivative of $f(Y)$ in the direction $H$. In the above formula, $h$ is a real infinitesimal parameter. Properties of these Fréchet derivatives continues to be an active area of research [70], and they have also been recently used in quantum information theory [71].

In this paper, we are interested in matrix functions that are either linear or quadratic in the matrix variable $Y$. This leads us to analyse the Fréchet derivatives given by the following lemma.

Lemma 3. Let $Y, H \in L(\mathbb{C}^n)$. Then

$$\nabla_{Y,H} \text{Tr}[AY] = \text{Tr}[AH] \quad (A2)$$

$$\nabla_{Y,H} \text{Tr}[AY^\dagger] = \text{Tr}[AH^\dagger] \quad (A3)$$

$$\nabla_{Y,H} \text{Tr}[YAY^\dagger] = \text{Tr}[AY^\dagger H + YAH^\dagger] \quad (A4)$$

$$\nabla_{Y,H} \text{Tr}[Y^\dagger AY] = \text{Tr}[Y^\dagger AH + AYH^\dagger]. \quad (A5)$$

Proof. The proof of the above results from direct application of the definition of the Fréchet derivative for the first two equations. For the last two equations, we also use the cyclic property of the trace.

We are often faced with the unconstrained minimisation of a quadratic form, and we show in the following lemma what the optimal solution to these optimisation problems are.

Lemma 4. Let $A \in L(\mathbb{C}^n)$ and let $\rho$ be a full rank matrix in $\mathbb{H}_n$. Then

$$\min_{Y \in L(\mathbb{C}^n)} \left( \text{Tr}[Y\rho Y^\dagger] + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger] \right) = -\text{Tr}[A^\dagger \rho^{-1} A] \quad (A6)$$

$$\min_{Y \in L(\mathbb{C}^n)} \left( \text{Tr}[Y^\dagger \rho Y] + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger] \right) = -\text{Tr}[A \rho^{-1} A^\dagger]. \quad (A7)$$

with optimal solutions given by $Y = -A^\dagger \rho^{-1}$ and $Y = -\rho^{-1} A^\dagger$ respectively.

Proof. We first prove (A6) and (A7). The corresponding objective functions that are to be minimised are convex and differentiable, so it suffices to find when their Fréchet derivatives
are equal to zero for any direction $H$. For this, we use Lemma 3, from which we find that we must have $\rho Y^\dagger + A = 0$ and $Y^\dagger \rho + A = 0$ respectively. Making use of the fact that $\rho$ is invertible whenever it has full rank, we multiply both sides of the equations, and find that the optimal $Y$s are given by $Y = -A^\dagger \rho^{-1}$ and $Y = -\rho^{-1}A^\dagger$ respectively. Substituting this back into the objective functions gives the result. \hfill \Box

Appendix B: hcrb for two parameters

We explicitly derive the hcrb for the two-parameter case. In the two-parameter setting, the hcrb with a weight matrix $W$ is given by the optimisation problem

$$
\begin{align*}
\text{minimize} & \quad \text{Tr}[W\Re Z] + \|W\Im Z\|_1, \\
\text{subject to} & \quad \text{Tr}[\rho X_i] = 0, \\
& \quad \text{Tr}[^*] \rho X_k = 0,
\end{align*}
$$

(B1)

where $X_i$ are constrained to be Hermitian matrices in $\mathbb{H}_N$, and $Z$ is a matrix given by

$$
Z = \begin{bmatrix}
\text{Tr}[\rho X_1^2] & \text{Tr}[\rho X_1 X_2] \\
\text{Tr}[\rho X_2 X_1] & \text{Tr}[\rho X_2^2]
\end{bmatrix}.
$$

(B2)

Note that $W$ is always taken to be a positive definite matrix. For simplicity, we only consider the scenario where $W$ is the identity matrix.

1. Reformulation of the optimisation problem

The optimisation problem (9) can be solved analytically primarily from our ability to rewrite the objective function as a quadratic function in the optimisation variables $X_1$ and $X_2$. The method of Lagrange multipliers when applied to problems with quadratic objective functions and linear equality constraints is well-known to be exactly solvable, for example in theory of portfolio optimisation in finance [72]. A similar argument will allow us to solve (9) using this method.

We begin by showing why the objective function is quadratic. To see this, we first note that the diagonal terms of $Z$ are positive numbers because $X_1$ and $X_2$ are Hermitian and $X(\cdot)^\dagger X$ is a completely positive map. Second, the positivity of the diagonal entries of $Z$ implies that

$$
\text{Re}\text{Tr}[Z] = \text{Tr}[Z] = \text{Tr}[X_1 \rho Y^\dagger] + \text{Tr}[X_2 \rho Y^\dagger].
$$

Third, the positivity of the diagonal entries of $Z$ implies that the trace norm of $\Im Z$ can be explicitly evaluated. This is because the diagonal entries of $\Im Z$ must be zero. Since $X_1$, $X_2$ and $\rho$ are Hermitian matrices, it follows that

$$
\Im Z = \frac{1}{2i} \begin{pmatrix} 0 & w \\ -w & 0 \end{pmatrix},
$$

where $w = \text{Tr}[\rho X_1 X_2] - \text{Tr}[X_2 X_1 \rho]$ is an imaginary number. The eigenvalues of $\Im Z$ are therefore $\pm w/2$, which implies that the trace norm of $\Im Z$ is max{$iw, -iw$}. From this, we get

$$
\begin{align*}
\text{Re}\text{Tr}[Z] + iw = & \text{Tr}\left[(X_1 + iX_2)\rho(X_1 + iX_2)^\dagger\right] \\
\text{Re}\text{Tr}[Z] - iw = & \text{Tr}\left[(X_1 - iX_2)\rho(X_1 - iX_2)^\dagger\right].
\end{align*}
$$

(B3, B4)

Now let us make the substitution $Y = X_1 + iX_2$. In this scenario, we can rewrite the equality constraints in (B1) as

$$
\begin{align*}
\text{Tr}[\rho Y] &= 0, \\
\text{Tr}[\partial_1 \rho Y] &= 1, \\
\text{Tr}[\partial_2 \rho Y] &= i.
\end{align*}
$$

(B5)

Hence the optimisation problem (9) can be written as

$$
\begin{align*}
\text{minimize} & \quad t, \\
\text{subject to} & \quad \text{Tr}[Y \rho Y^\dagger] \leq t, \\
& \quad \text{Tr}[Y^\dagger \rho Y] \leq t, \\
& \quad \text{Tr}[\partial_1 \rho Y] = 1, \\
& \quad \text{Tr}[\partial_2 \rho Y] = i.
\end{align*}
$$

(B6)

Note that the optimisation problem (B6) is a linear optimisation problem with convex quadratic and linear constraints. When the equality constraints are satisfied, the quadratic terms in the inequality constraints are non-negative, and by setting $t$ to be arbitrarily large, we can see that the inequality constraints in (B6) can always be strictly satisfied. Since (B6) is also a convex optimisation problem because of its linear objective function and convex constraint functions, the Slater constraint qualification holds with respect to (B6). This implies that the first order Karush-Kuhn-Tucker (kkt) conditions suffices to determine the optimality of (B6).

2. Analysing the Lagrangian

The kkt conditions are stated in terms of the Lagrangian of (B6). The column vector of Lagrange multipliers corresponding to the equality constraints is

$$
\mathbf{z} = (y_1, z_{1,1}, z_{2,1}, y_2, z_{1,2}, z_{2,2}).
$$

(B7)

The Lagrangian of (B6) is

$$
\begin{align*}
L(Y, t, u, v, \mathbf{z}) = & \quad t + u\text{Tr}[Y \rho Y^\dagger] - ut + v\text{Tr}[Y^\dagger \rho Y] - vt \\
& \quad + y_1\text{Re}\text{Tr}[\rho Y] + y_2\text{Im}\text{Tr}[\rho Y] \\
& \quad + z_{1,1}(\text{Re}\text{Tr}[\partial_1 \rho Y] - 1) + z_{1,2}\text{Im}\text{Tr}[\partial_1 \rho Y] \\
& \quad + z_{2,1}\text{Re}\text{Tr}[\partial_2 \rho Y] + z_{2,2}(\text{Im}\text{Tr}[\partial_2 \rho Y] - 1),
\end{align*}
$$

(B8)

where $u, v$ are non-negative Lagrange multipliers corresponding to the inequality constraints.

There are four types of kkt conditions. First is the stationarity of the derivative of the Lagrangian with respect to the primal variables. Second is complementary slackness, which
states that the product of the constraint functions \[73\] and their corresponding Lagrange multipliers is always zero. Third is the feasibility of the primal variables, and fourth is feasibility of the dual variables. Our strategy is to show that these KKT optimality conditions hold.

Now we use the fact that

\[
\begin{align*}
\mathrm{Re} \text{Tr}[\rho Y] &= \frac{\text{Tr}[\rho Y] + \text{Tr}[\rho Y^\top]}{2} \quad (B9) \\
\mathrm{Im} \text{Tr}[\rho Y] &= \frac{\text{Tr}[\rho Y] - \text{Tr}[\rho Y^\top]}{2i} \quad (B10) \\
\mathrm{Re} \text{Tr}[\partial_j \rho Y] &= \frac{\text{Tr}[\partial_j \rho Y] + \text{Tr}[\partial_j \rho Y^\top]}{2} \quad (B11) \\
\mathrm{Im} \text{Tr}[\partial_j \rho Y] &= \frac{\text{Tr}[\partial_j \rho Y] - \text{Tr}[\partial_j \rho Y^\top]}{2i} \quad (B12)
\end{align*}
\]

Using this, it follows that

\[
L(Y, t, u, v, z) = t(1 - u - v) - b^\top z + u \text{Tr}[Y \rho Y^\top] + v \text{Tr}[Y^\top \rho Y] + \text{Tr}[AY] + \text{Tr}[A^\top Y^\top], \quad (B13)
\]

where \(b\) is the column vector \((0, 1, 0, 0, 0, 1)\) and

\[
A = y_1 A_1 + z_{1,1} A_2 + z_{2,1} A_3 + y_2 A_4 + z_{1,2} A_5 + z_{2,2} A_6, \quad (B14)
\]

where

\[
A_1 = \frac{\rho}{2}, \quad A_2 = \frac{\partial_1 \rho}{2}, \quad A_3 = \frac{\partial_2 \rho}{2}, \quad (B15)
\]

and \(\{A_4, A_5, A_6\} = -i\{A_1, A_2, A_3\}\).

Before we proceed to derive the Lagrange dual function, we note the following.

1. We prove that the optimal \(t\) must be strictly positive from the positive definiteness of \(\rho\). From the positive definiteness of \(\rho\), \(t\) is equal to zero if and only if \(Y = 0\), but this would violate the feasibility constraints. Hence \(t\) cannot be equal to zero.

2. We prove that for optimal dual variables, we must have \(u + v = 1\). If the dual variables \(u\) and \(v\) are such that \(u + v < 1\), then the optimal \(t\) is 0 which is a contradiction. If \(u + v > 1\), the optimal \(t\) is arbitrarily large which makes the Lagrangian unbounded from below, which is also a trivial scenario. This implies that the term in the Lagrangian \(t(1 - u - v)\) for optimal values of the primal and dual variables must evaluate to zero. Hence for the optimal solution, we must have \(u + v = 1\) and \(t > 0\).

3. The KKT conditions require that the complementary slackness conditions hold for the inequality constraints in (B6). This means that

\[
\begin{align*}
{u} \big(\text{Tr}[Y \rho Y^\top] - t\big) &= 0 \\
{v} \big(\text{Tr}[Y^\top \rho Y] - t\big) &= 0 \quad (B16)
\end{align*}
\]

If \(\|\text{Im} Z\|_1 > 0\), exactly one of the constraints corresponding to \(u\) and \(v\) must be tight, and complementary slackness implies that the optimal \((u, v)\) must be either \((u, v) = (1, 0)\) or \((u, v) = (0, 1)\). This corresponds to the scenario where the QCRR is not equal to the hCRB. If \(\|\text{Im} Z\|_1 = 0\), \((u, v) = (1, 0)\) and \((u, v) = (0, 1)\) do not necessarily optimize the value of the Lagrange dual, and in general provide a lower bound to the Lagrange dual.

### 3. Deriving the Lagrange dual functions

When \((u, v) = (1, 0)\), the Lagrangian evaluates to

\[
L(Y, t, 0, 0, z) = -b^\top z + \text{Tr}[Y \rho Y^\top] + \text{Tr}[AY] + \text{Tr}[A^\top Y^\top], \quad (B17)
\]

where \(b = (0, 1, 0, 0, 0, 1)^\top\). Since \(\rho\) is full rank, \(\rho\) is invertible. Using Lemma 4, the above is minimised with respect to \(Y\) when \(Y = -A^\top \rho^{-1} A\). In this scenario, the Lagrange dual function of (B6) evaluated with \((u, v) = (1, 0)\) is

\[
g(1, 0, z) = -\text{Tr}[A^\top \rho^{-1} A] - b^\top z. \quad (B18)
\]

Similarly when \((u, v) = (0, 1)\), the Lagrangian evaluates to

\[
L(Y, t, 0, 1, z) = -b^\top z + \text{Tr}[Y^\top \rho Y] + \text{Tr}[AY] + \text{Tr}[A^\top Y^\top], \quad (B19)
\]

and is minimised when \(Y = -\rho^{-1} A^\top\) with an optimal value of \(-\text{Tr}[A \rho^{-1} A^\top]\). In this scenario, the Lagrange dual function of (B6) evaluated with \((u, v) = (1, 0)\) is

\[
g(0, 1, z) = -\text{Tr}[A \rho^{-1} A^\top] - b^\top z. \quad (B20)
\]

The Lagrange dual functions \(g(1, 0, z)\) and \(g(0, 1, z)\) can be rewritten in terms of the matrices \(Q_1\) and \(Q_2\) where in the Dirac bra-ket notation, we have

\[
Q_1 = \sum_{j,k=1,\ldots,6} \text{Tr}[A_j \rho^{-1} A_k] |j\rangle \langle k| \]

\[
Q_2 = \sum_{j,k=1,\ldots,6} \text{Tr}[A_j \rho^{-1} A_k^\top] |j\rangle \langle k|. \quad (B21)
\]

Here \(|j\rangle\) denotes a column vector and \(|k\rangle\) denotes a row vector. The Lagrange dual function that we consider are thus

\[
g(1, 0, z) = -z^\top Q_1 z - b^\top z \]

\[
g(0, 1, z) = -z^\top Q_2 z - b^\top z. \quad (B22)
\]

### 4. Optimal solutions

When hCRB is strictly greater than QCRR, the hCRB is thus equal to the maximum of the optimal value of both the Lagrange dual functions (B22), and is thus

\[
\max_{j=1,2} \max_{z \in \mathbb{R}^6} (-z^\top Q_j z - b^\top z). \quad (B23)
\]

Since \(Q_j\) are derived from the Lagrange dual, we have the promise that \(Q_j\) are positive-semidefinite matrices.
Using the spectral decomposition of $Q_j$, we can write

$$Q_j = \sum_{i=1}^{6} \lambda_i u_i u_i^\dagger,$$

where $\lambda_i \geq 0$ and $u_i$ are eigenvectors of $Q_j$. Then, by denoting $\langle \cdot, \cdot \rangle$ as the inner product on a complex Euclidean space, we can write

$$z^T Q z = \sum_{i=1}^{6} \lambda_i \langle z, u_i \rangle \langle u_i, z \rangle$$

$$= \sum_{i=1}^{6} \lambda_i |\langle z, u_i \rangle|^2,$$ (B24)

and we can see that $z^T Q_j z$ is always a real number for every real vector $z$.

Since $z^T Q_j z$ is always real, and we know that $z$ is real, then we must have

$$\text{Re}(z^T Q_j z) = z^T \text{Re}(Q_j) z.$$

This implies that it suffices to consider the minimization

$$\min_{z \in \mathbb{R}^b} \text{Re}(Q_j) z + b^T z,$$

from which it now follows that

$$2 \text{Re}(Q_j) z + b = 0$$

is the correct optimality condition to consider. Thus, when $\text{Re}(Q_j)$ is full rank, we can write the hcrb as

$$\max_{j=1,2} \left( \frac{b^T (\text{Re}(Q_j))^{-1} b}{2} - \frac{b^T \text{Re}(Q_j)^{-1} Q_j \text{Re}(Q_j)^{-1} b}{4} \right).$$ (B25)

Interestingly, when $\rho$ is full rank, the matrices $\text{Re}(Q_j)$ are also full rank. We demonstrate this in the next subsection.

5. Full-rankness of $Q$

The analytic solution to the hcrb requires $\text{Re}(Q_j)$ to have full rank such that the solution can be determined. In this subsection, we demonstrate that the full-rankness of the probe state $\rho$ entails the full-rankness of these matrices. Since the regularity conditions of estimation theory require the state to be full-rank, our solution to the hcrb always exists.

Notice that the matrices $Q_j$ defined in Eq. (B21) can be written

$$Q_1 = \begin{pmatrix} H & -iH \\ iH & H \end{pmatrix}, \quad Q_2 = \begin{pmatrix} H & iH \\ -iH & H \end{pmatrix},$$ (B26)

where $H$ is Gram matrix defined as follows. We consider the Hilbert-Schmidt inner-product $\langle X, Y \rangle = \text{Tr}[X^\dagger Y]$. We define the operators

$$B_1 = \rho^{-1/2} A_1 = \rho^{-1/2} \rho/2 = \rho^{1/2}/2,$$ (B27)

$$B_2 = \rho^{-1/2} A_2 = \rho^{-1/2} \delta_1 \rho/2,$$ (B28)

$$B_3 = \rho^{-1/2} A_3 = \rho^{-1/2} \delta_2 \rho/2.$$ (B29)

Then, we have that $H$ is a Gram matrix with respect to this set of operators

$$H_{i,j} = \langle B_i, B_j \rangle.$$ (B30)

As a Gram matrix, it is positive semi-definite. Furthermore, we know that $H$ will be full-rank if and only if the set $\{B_1, B_2, B_3\}$ is linearly independent. We note that $A_1$ cannot be written as a sum of $A_2$ and $A_3$ (since $A_1$ has nonzero trace whereas $A_2$ and $A_3$ are traceless). Also by a trace-argument, if $\{A_1, A_2, A_3\}$ are linearly dependent, we must have that $A_2$ is proportional to $A_3$. But if $A_2$ and $A_3$ are proportional, then it is really one-parameter problem and not a two-parameter problem. Hence, $\{A_1, A_2, A_3\}$ are linearly independent. If we assume $\rho$ is full-rank, then $\{B_1, B_2, B_3\} = \rho^{-1/2} \{A_1, A_2, A_3\}$ is also a linearly independent set. So full-rankness of $\rho$ entails full-rankness of $H$.

We are now interested in the real part of the $Q_j$ matrices, looking for solutions of $2 \text{Re}(Q_j) z + b = 0$. Considering $j = 2$

$$\text{Re}(Q_2) = \begin{pmatrix} \text{Re}(H) & -i\text{Im}(H) \\ i\text{Im}(H) & \text{Re}(H) \end{pmatrix}.$$ (B31)

By performing elementary row operations by taking a linear combination of rows, followed by elementary column operators by taking a linear combination of columns, we get

$$\text{Re}(Q_2) \rightarrow \begin{pmatrix} H & iH \\ \text{Im}(H) & \text{Re}(H) \end{pmatrix} \rightarrow \begin{pmatrix} H & 0 \\ \text{Im}(H) & H^* \end{pmatrix},$$ (B32)

where we used $\text{Re}(H) + i\text{Im}(H) = H$. Since both rows are linearly independent, $\text{Re}(Q_2)$ is also always full-rank. Therefore, we have that if the state is full-rank, then so too is the matrix $\text{Re}(Q_j)$.

Appendix C: Lower bound in the multi-parameter setting

By restricting ourselves to the identity weight matrix, recall that the hcrb is the optimal value of the following optimisation problem over the Hermitian matrices $X_j$ in $\mathbb{H}_N$ given by

$$\begin{align*}
\text{minimize} & \quad \text{Tr}[\text{Re}Z] + \|\text{Im}Z\|_1, \\
\text{subject to} & \quad \text{Tr}[\rho X_j] = 0, \\
& \quad \text{Tr}[\delta_j \rho X_k] = \delta_{jk}.
\end{align*}$$ (C1)

For $j, k = \{1, \ldots, d\}$, let

$$w_{j,k} = \text{Tr}[\rho X_j X_k] - \text{Tr}[\rho X_k X_j].$$ (C2)
1. Deriving a lower bound for the objective function

In general for a $d$-parameter estimation problem, we have

$$\text{Tr} [\text{Re} Z] = \sum_{j=1}^d \text{Tr} [X_j \rho X_j^+]$$

$$\text{Im} Z = \frac{1}{2d} \sum_{1 \leq j < k \leq d} w_{j,k} \langle |j\rangle \langle k| - |k\rangle \langle j| \rangle.$$  \hspace{1cm} (C3)

Note that $\text{Im} Z$ is always a skew-Hermitian matrix. For example, when $d = 3$, we have

$$\text{Im} Z = \begin{pmatrix} 0 & w_{1,2} & w_{1,3} \\ -w_{1,2} & 0 & w_{2,3} \\ -w_{1,3} & -w_{2,3} & 0 \end{pmatrix}. \hspace{1cm} (C5)$$

We use several ansatz’s to obtain a lower bound for the trace norm of $\text{Im} Z$ utilising the fact that

$$||\text{Im} Z||_1 = \max \{\text{Tr}[U \text{Im} Z] : -I_d \leq U \leq I_d\}. \hspace{1cm} (C6)$$

Here, the inequality for the matrix indicates Loewner ordering of matrices. Instead of optimising over all possible $U$, we can optimise $U$ over the finite set

$$\mathcal{U} = \left\{ \sum_{j=0}^{d-1} |i\rangle \langle i| \oplus s |(-1)^s \rangle : x_r \in \{0, 1\}, s = 1, 2, \ldots, d - 1 \right\}, \hspace{1cm} (C7)$$

where $\oplus$ denotes addition modulo $d$ over the ring $\mathbb{Z}_d$, $s$ denotes the shift, and $\{ |j\rangle : j = 0, \ldots, d-1 \}$ is an orthonormal basis. Clearly, the singular values of the matrices in $\mathcal{U}$ are all equal to 1, and hence $||\mathcal{U}||_{s_\infty} \leq 1$, where $\| \cdot \|_{s_\infty}$ denotes the Schatten infinity norm. Hence we can use $\mathcal{U}$ to get the lower bound

$$||\text{Im} Z||_1 \geq \max \{\text{Tr}[\text{Im} Z U] : U \in \mathcal{U}\}, \hspace{1cm} (C8)$$

which implies that for every $s = \{1, \ldots, \lfloor (d - 1)/2 \rfloor \}$ we have

$$||\text{Im} Z||_1 \geq \frac{1}{2} \sum_{(a,b) \in E_{d,s}} |w_{a,b}| \hspace{1cm} (C9)$$

where

$$E_{d,s} = \{(j + 1, (j \oplus s) + 1) : j = 0, \ldots, d - 1\} \hspace{1cm} (C10)$$

and each $|w_{a,b}|$ is the maximum of

$$i\text{Tr}[X_b \rho X_a^+] - i\text{Tr}[X_a \rho X_b^+]$$

and

$$-i\text{Tr}[X_b \rho X_a^+] + i\text{Tr}[X_a \rho X_b^+].$$

2. Recasting the optimisation problem

Now let us observe that $\text{Tr} [\text{Re} Z] + \sum_{(a,b) \in E_{d,s}} |w_{a,b}|/2$ can be written as a quadratic form in the optimisation variables $X_1, \ldots, X_d$. To see this, note for instance that

$$\text{Tr} [\text{Re} Z] + \frac{i}{2} \sum_{(a,b) \in E_{d,s}} w_{a,b} \hspace{1cm} (C11)$$

$$= \sum_{j=1}^d \text{Tr} [X_j \rho X_j^+] + \frac{i}{2} \left( \text{Tr} [X_2 \rho X_1^+] + \cdots + i\text{Tr} [X_d \rho X_1^+] \right)$$

$$- \frac{1}{2} \left( \text{Tr} [X_1 \rho X_2^+] + \cdots + i\text{Tr} [X_d \rho X_1^+] \right)$$

$$= \frac{1}{2} \sum_{(a,b) \in E_{d,s}} \text{Tr} [(X_a + iX_b) \rho (X_a + iX_b)^{\dagger}]. \hspace{1cm} (C12)$$

The last equality in Eq. (C11) arises because every $j$ can be a component of $(a, b)$ in exactly two ways, as the tuples $(a, b)$ correspond to edges in a cycle graph. To see how this works explicitly in the $d = 3$ and $s = 1$ scenario, note that

$$\text{Tr} [\text{Re} Z] + \frac{i}{2} \sum_{(a,b) \in E_{d,s}} w_{a,b}$$

$$= \text{Tr} [X_1 \rho X_1^+] + \text{Tr} [X_2 \rho X_2^+] + \text{Tr} [X_3 \rho X_3^+]$$

$$+ \frac{i}{2} (\text{Tr} [X_2 \rho X_1^+] + \text{Tr} [X_3 \rho X_2^+] + \text{Tr} [X_1 \rho X_3^+] - \text{Tr} [X_1 \rho X_2^+] + \text{Tr} [X_2 \rho X_3^+] + \text{Tr} [X_3 \rho X_1^+])$$

$$= \frac{1}{2} \text{Tr} [(X_1 + iX_2) \rho (X_1 + iX_2)^{\dagger}] + \frac{1}{2} \text{Tr} [(X_2 + iX_3) \rho (X_2 + iX_3)^{\dagger}] + \frac{1}{2} \text{Tr} [(X_3 + iX_1) \rho (X_3 + iX_1)^{\dagger}].$$

Now, given a binary vector $\alpha = (\alpha_1, \ldots, \alpha_d)$, let us define

$$V_\alpha = \frac{1}{2} \sum_{(a,b) \in E_{d,s}} \text{Tr} [(X_a + (-1)^{\alpha_a} iX_b) \rho (X_a + (-1)^{\alpha_b} iX_b)^{\dagger}]. \hspace{1cm} (C13)$$

Then we can rewrite (C1) as an optimisation over Hermitian matrices $X_1, \ldots, X_d$ where

$$\minimize_{X_1, \ldots, X_d} \max_{\alpha \in \{0,1\}^d} V_\alpha, \hspace{1cm} \text{subject to} \hspace{1cm} \text{Tr} [\rho X_k] = 0, \hspace{1cm} \text{Tr} [\partial_j \rho X_k] = \delta_{jk}. \hspace{1cm} (C14)$$
We can rewrite with an introduction of an auxiliary variable \( \alpha \in \mathbb{R} \) so that (C14) is equivalent to

\[
\begin{align*}
\text{minimize} & \quad t, \\
\text{subject to} & \quad \text{Tr}[\rho X_j] = 0, \\
& \quad \text{Tr}[\partial_j \rho X_k] = \delta_{jk}, \\
& \quad V_\alpha \leq t, \\
& \quad \alpha \in \{0, 1\}^d.
\end{align*}
\]  

(C15)

This minimisation problem can be numerically checked for consistency with the optimisation in Eq. (C1).

3. Diagonalising the quadratic forms

At this point, we want to define new variables \( Y_1, \ldots, Y_d \) so that we can for example write

\[
V_\alpha = \frac{1}{2} \sum_{j=1}^d \left( \delta_{0,\alpha_j} \text{Tr}[Y_j \rho Y_j^\dagger] + \delta_{1,\alpha_j} \text{Tr}[Y_j^\dagger \rho Y_j] \right).
\]  

(C16)

In general, we define the variables \( Y_1, \ldots, Y_d \) so that they depend linearly on the variables \( X_1, \ldots, X_d \) according to the system of linear equations

\[
Y = SX,
\]  

(C17)

where

\[
S = \begin{pmatrix} S_{1,1} & \cdots & S_{1,d} \\ \vdots & \ddots & \vdots \\ S_{d,1} & \cdots & S_{d,d} \end{pmatrix}, \quad Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_d \end{pmatrix}, \quad X = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix}.
\]  

(C18)

It is important to also express the variables \( X_1, \ldots, X_d \) in terms of \( Y_1, \ldots, Y_d \), and this means that we need the matrix

\[
S = \begin{pmatrix} S_{1,1} & \cdots & S_{1,d} \\ \vdots & \ddots & \vdots \\ S_{d,1} & \cdots & S_{d,d} \end{pmatrix}
\]  

(C19)

to be invertible. When \( S \) is full rank, \( T = S^{-1} \) exists. One possible choice of \( S \) is where \( S_{i,j} = 1 \) and \( S_{j,j+1} = i \) for all \( j = 1, \ldots, d - 1 \), and \( S_{d,d} = 1, S_{d,1} = i \), and all other matrix elements of \( S \) are zero. For instance, when \( d = 3 \), we use

\[
S = \begin{pmatrix} 1 & i & 0 \\ 0 & 1 & i \\ i & 0 & 1 \end{pmatrix}.
\]  

(C20)

However such a choice of \( S \) need not have be full rank. It is easy to see that \( S \) is full rank whenever its dimension is not a multiple of 4.

**Proposition 5.** Let \( d \) be a positive integer, that is at least 2 and is not a multiple of 4. Then \( S \) has full rank.

**Proof.** To prove this, we use pairwise orthogonality of the vectors \( (1, i), (1, -i), (1, -1), (1, 1) \). By Gaussian elimination and the properties of \( i \), we can easily arrive at the result. \( \square \)

When \( d \) is not a multiple of 4, we denote

\[
T = \begin{pmatrix} T_{1,1} & \cdots & T_{1,d} \\ \vdots & \ddots & \vdots \\ T_{d,1} & \cdots & T_{d,d} \end{pmatrix},
\]  

(C21)

and it follows that for every \( j = 1, \ldots, d \), we have

\[
X_k = \sum_{\ell=1}^d T_{k,\ell} Y_\ell.
\]  

(C22)

From (C22), assuming the Hermiticity of \( X_k \), we can recast the equality constraints in (C15) as

\[
c_{0,k}(Y) = \frac{1}{2} \sum_{\ell=1}^d \left( T_{k,\ell} \text{Tr}[\rho Y_\ell] + T_{k,\ell}^* \text{Tr}[\rho Y_\ell^\dagger] \right) = 0,
\]  

(C23)

\[
c_{j,k}(Y) = \frac{1}{2} \sum_{\ell=1}^d \left( T_{k,\ell} \text{Tr}[\rho_j Y_\ell] + T_{k,\ell}^* \text{Tr}[\rho_j Y_\ell^\dagger] \right) - \delta_{j,k} = 0.
\]  

(C24)

By choosing our explicit form of the matrix \( A \) strategically, we can write the quadratic inequality constraints with \( Y_\ell \) in the form as given by (C16).

Since the variables \( Y_\ell \) are non-Hermitian in general, we need to impose additional constraints, namely the fact that the corresponding \( X_k \) are Hermitian. The Hermiticity of \( X_k \) implies from (C22) that

\[
\sum_{\ell=1}^d \left( T_{k,\ell} Y_\ell - T_{k,\ell}^* Y_\ell^\dagger \right) = 0.
\]  

(C25)

The left side of (C25) is in general an antihermitian matrix, and to make it Hermitian, we multiply both sides by \( i \) to get

\[
H_k(Y) = i \sum_{\ell=1}^d \left( T_{k,\ell} Y_\ell - T_{k,\ell}^* Y_\ell^\dagger \right) = 0.
\]  

(C26)

With all these constraints, we recast the optimisation problem (C15) as the following optimisation problem.

\[
\begin{align*}
\text{minimize} & \quad t, \\
\text{subject to} & \quad c_{0,k}(Y) = 0, \\
& \quad c_{j,k}(Y) = 0, \\
& \quad \frac{1}{2} \sum_{\ell=1}^d \left( \delta_{0,\alpha_j} \text{Tr}[\rho Y_\ell Y_\ell^\dagger] + \delta_{1,\alpha_j} \text{Tr}[Y_\ell^\dagger \rho Y_\ell] \right) \leq t, \\
& \quad H_k(Y) = 0.
\end{align*}
\]  

(C27)

4. Analysis on the Lagrangian

Here we consider the constraints in (C27) over \( j, k = 1, \ldots, d \) and \( \alpha_1, \ldots, \alpha_d = 0, 1 \), which gives us a total of \( d(d+1) \) regular equality constraints, \( d \) matrix equality constraints, and \( 2^d \)
regular inequality constraints. The Lagrangian corresponding to (C27) can then be written as

\[
L_d = t + \sum_{j=0}^{d} \sum_{k=1}^{d} z_{j,k} c_{j,k}(Y) + \sum_{k=1}^{d} \text{Tr}[\xi_k H_k(Y)] \\
+ \frac{1}{2} \sum_{\alpha \in \{0,1\}^d} v_{\alpha} \sum_{t=1}^{d} \left( \delta_{0,\alpha_t} \text{Tr}[Y_{t}\rho Y_{t}^\dagger] + \delta_{1,\alpha_t} \text{Tr}[Y_{t}^\dagger \rho Y_{t}] \right) \\
- \sum_{\alpha \in \{0,1\}^d} \alpha_t. \tag{C28}
\]

Here, the Lagrange multipliers \(z_{j,k}\) are real numbers while the Lagrange multipliers \(v_{\alpha}\) are non-negative numbers. The Lagrange multipliers \(\xi_k\) are Hermitian matrices in \(\mathbb{H}_N\). Note that the multiparameter Lagrangian is a quadratic form in \(Y\), and as such, can be minimised using Lemma 4. Before we do so, we consider the minimisation of the Lagrangian with respect to the primal variable \(t\).

If the Lagrangian multiplier \(v_{\alpha}\) do not all sum to one, by picking \(t\) to either approach positive or negative infinity, the Lagrangian \(L_d\) becomes unbounded. Hence the optimal multipliers \(v_{\alpha}\) must sum to one. By picking a discrete set of values of \(v_{\alpha}\), where \(v_{\alpha}\) is equal to zero to all but one value of \(\alpha\), and maximising the Lagrange dual function for each of these cases, we can obtain our lower bound to the multi-parameters ncRB.

Hence without loss of generality, there is some value of the binary vector \(\alpha\) for which the effective Lagrangian that we need to consider is

\[
L_{d,\alpha} = \sum_{j=0}^{d} \sum_{k=1}^{d} z_{j,k} c_{j,k}(Y) + \sum_{k=1}^{d} \text{Tr}[\xi_k H_k(Y)] \\
+ \frac{1}{2} \sum_{t=1}^{d} \left( \delta_{0,\alpha_t} \text{Tr}[Y_{t}\rho Y_{t}^\dagger] + \delta_{1,\alpha_t} \text{Tr}[Y_{t}^\dagger \rho Y_{t}] \right). \tag{C29}
\]

Now let \(\rho_0 = \rho\), and let

\[
\Gamma_{\ell} = \sum_{k=1}^{d} T_{k,\ell} \left( \sum_{j=0}^{d} z_{j,k} \rho_j + i \xi_k \right). \tag{C30}
\]

Then we can rewrite the terms on the first line on the right side of (C29) as

\[
L_{d,\alpha} = - \sum_{j=1}^{d} z_{j,j} + \frac{1}{2} \sum_{t=1}^{d} \left( \text{Tr}[\Gamma_{t}\rho \Gamma_{t}^\dagger] + \text{Tr}[\Gamma_{t}^\dagger \rho Y_{t}^\dagger] \right) \\
+ \delta_{0,\alpha_t} \text{Tr}[(1/ \rho)^{-1} \Gamma_{t}] + \delta_{1,\alpha_t} \text{Tr}[\Gamma_{t}^\dagger \rho^{-1} \Gamma_{t}]. \tag{C31}
\]

Then, given that \(\rho\) is a Hermitian full rank matrix, we can use Lemma 4 to get the corresponding Lagrange dual to be

\[
g_{\alpha} = \min_{Y} L_{d,\alpha} \\
= - \sum_{j=1}^{d} z_{j,j} + \frac{1}{2} \sum_{t=1}^{d} \left( \delta_{0,\alpha_t} \text{Tr}[(1/ \rho)^{-1} \Gamma_{t}] + \delta_{1,\alpha_t} \text{Tr}[\Gamma_{t}^\dagger \rho^{-1} \Gamma_{t}]. \tag{C32}
\]

Our lower bound to the ncRB is thus

\[
\max_{\alpha \in \{0,1\}^d} \max \left\{ g_{\alpha} : z_{j,k} \in \mathbb{R}, \xi_k \in \mathbb{H}_N \right\}, \tag{C33}
\]

where \(j = 0, \ldots, d\), \(k = 1, \ldots, d\). Any feasible value of \(g_{\alpha}\) yields a lower bound to the ncRB.

---

[1] S. L. Braunstein and C. M. Caves, *Phys. Rev. Lett.* **72**, 3439 (1994).

[2] V. Giovannetti, S. Lloyd, and L. Maccone, *Nat. Photon.* **5**, 222 (2011).

[3] G. Tóth and I. Apellaniz, *J. Phys. A: Math. Theoret.* **47**, 424006 (2014).

[4] C. Helmstrom, *IEEE Transactions on Information Theory* **14**, 234 (1968).

[5] C. W. Helmstrom, *Quantum Detection and Estimation Theory* (Academic Press Inc., 1976).

[6] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, 1st ed. (Springer, 2011).

[7] K. Matsumoto, *J. Phys. A: Math. Gen.* **35**, 3111 (2002).

[8] M. Hayashi, *Asymptotic Theory of Quantum Statistical Inference* (World Scientific Publishing, 2008).

[9] R. D. Gill and M. Guta, arXiv.org (2012), 1112.2078v2.

[10] D. Burgarth and A. Ajoy, *Phys. Rev. Lett.* **110**, 090402 (2017).

[11] M. Tsang, H. M. Wiseman, and C. M. Caves, *Phys. Rev. Lett.* **106**, 090401 (2011).

[12] S. Z. Ang, G. I. Harris, W. P. Bowen, and M. Tsang, *New J. Phys.* **15**, 103028 (2013).

[13] T. Baumgratz and A. Datta, *Phys. Rev. Lett.* **116**, 030801 (2016).

[14] M. Tsang, *Phys. Rev. Lett.* **102**, 253601 (2009).

[15] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, *Phys. Rev. Lett.* **111**, 070403 (2013).

[16] C. N. Gagatsos, D. Branford, and A. Datta, *Phys. Rev. A* **94**, 042342 (2016).

[17] L. Perzè, M. A. Cianzini, N. Spagnolo, P. C. Humphreys, A. Datta, I. A. Walmsley, M. Barbieri, F. Sciarrino, and A. Smerzi, *Phys. Rev. Lett.* **119**, 130504 (2017).

[18] T. J. Proctor, P. A. Knott, and J. A. Dunningham, “Networked quantum sensing,” (2017), unpublished, arXiv:1702.04271.

[19] J. S. Sidhu and P. Kok, *Phys. Rev. A* **95**, 063829 (2017).

[20] J. S. Sidhu and P. Kok, *ArXiv* (2018), arXiv:1802.01601 [quant-ph].

[21] M. G. A. Paris, *Int. J. Quantum Inf.* **7**, 125 (2009).

[22] D. A. S. Fraser, *J. R. Stat. Soc. Ser. B* **26**, 46 (1964).

[23] M. J. W. Hall and H. M. Wiseman, *Phys. Rev. X* **2**, 041006 (2012).

[24] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński,
