Tight Bounds on the Simultaneous Estimation of Incompatible Parameters

Jasminder S. Sidhu[1,2,*] Yingkai Ouyang,[1,†] Earl T. Campbell,[1,3] and Pieter Kok[1]

[1]*Department of Physics and Astronomy, The University of Sheffield, Sheffield, S3 7RH, United Kingdom*
[2]*SUPA Department of Physics, The University of Strathclyde, Glasgow, G4 0NG, United Kingdom*
[3]*AWS Center for Quantum Computing, Pasadena, California 91125, USA*

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 greatly diminishes the applicability of estimation theory in many practical implementations. The Holevo Cramér-Rao bound (HCRB) provides the most fundamental, simultaneously attainable bound for multiparameter estimation problems. A general closed form for the HCRB is not known given that it requires a complex optimization over multiple variables. In this work, we develop an analytic approach to solving the HCRB for two parameters. Our analysis reveals the role of the HCRB and its interplay with alternative bounds in estimation theory. 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 that even numerically permits a quadratic speedup over previous state-of-the-art approaches. We apply our results to compare the performance of different probe states in magnetic field sensing and characterize the performance of state tomography on the code space of noisy bosonic error-correcting codes. The sensitivity of state tomography on noisy binomial code states can be improved by tuning two coding parameters that relate to the number of correctable phase and amplitude damping errors. Our work provides fundamental insights and makes significant progress toward 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 the fundamental precision bound is 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 lower bound (QCRB), which was proved by Helstrom and Holevo [4–6]. Multiparameter 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 [14–18], and distributed sensing [19–22]. A central problem is to determine the optimal measurement strategies that saturate the QCRB [23]. To achieve this determination, one must assume locally unbiased estimators [24], which is reasonable given large amounts of prior information [25,26] and with many independent repetitions of the experiment [27]. Several reviews on the topic highlight recent progress in the field [28–30].

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 cannot achieve the optimal precision for each parameter individually. In this case, the QCRB matrix bound is generally not simultaneously saturable for all parameters [31–33]. This constraint motivates the search for tighter bounds that can be realized for practical applications of multiparameter estimation theory. The Holevo Cramér Rao bound (HCRB) encapsulates the

*[*]jsmdrsidhu@gmail.com
[†]y.ouyang@sheffield.ac.uk

difficulties associated with incompatible observables [6]. It represents the best precision attainable with collective measurements on an asymptotically large number of identical copies of a quantum state [34–37].

Despite its importance, the HCRB has seen limited use in quantum metrology so far. There are several reasons for this limited use. First, the HCRB is difficult to evaluate given that it is defined through a complex optimization 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 finds closed form results for parameter estimation with qubits [38] and explores 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 is shown that the HCRB is attained by single-copy measurements. The HCRB is also used as a tool to define the precision of state estimation for finite-dimensional quantum systems [37]. Bradshaw, Assad, and Lam calculate the HCRB for a joint parameter estimation of a displacement operation on a pure two-mode squeezed probe [39].

Arguably, the HCRB is most relevant in multiparameter estimation. An increasing number of true multiparameter estimation protocols is being explored [40–43], and, therefore, the need for general, attainable bounds on multiparameter quantum estimation is urgent. Recently, Albarelli, Friel, and Datta have investigated the numerical tractability of calculating the HCRB for the simultaneous estimation of multiple parameters [44]. For finite-dimensional systems, they recast the evaluation of the HCRB as a semidefinite program, which is an optimization 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 multiparameter estimation problems. We develop an analytical approach to solving the two-parameter HCRB and provide expressions on when the analytical solution is tight. Our analytical solution for the optimal observables that can saturate the HCRB allows one to establish analytically the minimum penalty due to the incompatibility of the observables. Specifically, we generalize attainability constraints for simultaneous multiparameter estimation problems where the commonly used Cramér-Rao bounds cannot be saturated due to incompatibility. The analytic two-parameter HCRB can be considered a generalized quantum uncertainty relation [45]. For more than two parameters, our method does not provide tight bounds but still outperforms the QCRB.

A. Summary of results

The HCRB is defined as a constrained minimization problem over measurement observables. By recasting the definition as a quadratic program with linear constraints, we find exact solutions to this minimization and determine the optimal observables. Our method to solve this minimization relies on the notion of duality in optimization theory, where the primal problem is transformed to its dual problem. Through the duality gap, we are able to quantify the minimum penalty of estimating incompatible observables.

In this article, we introduce three new algorithms that derive bounds to the HCRB for two and more parameters. First, we determine upper and lower bounds to the HCRB for two-parameter estimation problems that are not always tight. This algorithm leads to simple analytic expressions that are straightforwardly determined for probe states with full rank. The salient feature of this framework, from which the simplification is inherited, is that only the boundary values for the Lagrange dual variables are considered. Second, this method is extended to determine upper and lower bounds to the HCRB for more than two parameters. Finally, we return to the two-parameter HCRB to develop tight bounds. For this algorithm, we lift the constraint on the values for the Lagrange dual variables to explore the full generality permitted by our method. Our analysis for this exploration shows that the HCRB is a general solution to a Sylvester equation in the measurement observables and recovers the standard Lyapunov symmetric logarithmic derivative QCRB solution when the weak commutativity criterion is violated. This algorithm can be implemented numerically using a Bartels-Stewart algorithm for linear equation solvers and offers a quadratic speedup in run time over state-of-the-art semidefinite programming approaches.

Table I provides a high-level summary of these algorithms, along with any assumptions made. Our results provide a significant extension of the capabilities of previous approaches and clarify the role of the HCRB in the estimation of incompatible observables.

B. Outline of paper

We begin in Sec. II by providing an overview of multiparameter quantum estimation. In Sec. III, we introduce the four new algorithms for analytic and numerical results to the HCRB for two parameters and an arbitrary number of parameters. We detail connections between alternative precision bounds and significantly extend the capabilities of previous approaches in the literature. Section IV discusses applications of our results to magnetometry and explores how bosonic quantum codes can bestow resilience of parameter estimates against noise beyond practical control. These applications demonstrate the strengths of our results and extend deep connections between quantum metrology and quantum error-correcting codes. Finally, conclusions and interesting extensions to our results are provided in Sec. V.

TABLE I.   Algorithms and bounds to the HCRB for two and more parameters. Bounds are analytic, numerical, or a hybrid of analytical and numerical, as indicated in the first column. There is a trade-off between the assumptions taken for each algorithm and its complexity. Here, D is the dimension of the probe state, and ϵ a measure of how close the HCRB bound is to optimal. The final row provides comparative details for the SDP approach in Ref. [44].

Nature of HCRB bound	Assumptions	Algorithm details
Algorithm 1: analytic two-parameter bound	Full-rank ρ, linearly independent ρ_1 and ρ_2, and analytic form for Q matrix	Provides upper $\mathscr{U}$ and lower $\mathscr{L}$ bounds, need not be tight
Algorithms 2 and 3: hybrid multiparameter bound	Full-rank ρ and analytic form for Q matrix	Provides upper $\mathscr{U}$ and lower $\mathscr{L}$ bounds, need not be tight
Algorithm 4: hybrid two-parameter bound	Full-rank ρ, spectral decomposition of ρ. Full rankness of intermediate. $\mathcal{Q}$ matrix is full rank and takes at most τ time to compute.	Analytic bounds for $u \in [0, 1]$. Tight bounds certifiably attained by numerically varying u to maximize $\mathscr{L}_u$. Computes in $\mathcal{O}[\text{polylog}(1/\epsilon)\tau D^0]$.
Numerical two-parameter bound using Eq. (37)	Full-rank ρ	Computes in $\mathcal{O}[\text{polylog}(1/\epsilon)D^{2.376}]$ time using Bartels-Stewart algorithm or $\mathcal{O}[\text{polylog}(1/\epsilon)D^3]$ time using Gaussian elimination
SDP numerical algorithm	Arbitrary ρ, ρ_1, and ρ_2	Computes in $\mathcal{O}[\text{polylog}(1/\epsilon)D^{2\times2.376}]$ time, or $\mathcal{O}[\text{polylog}(1/\epsilon)D^6]$ time using Gaussian elimination

II. MULTIPARAMETER QUANTUM ESTIMATION

Quantum estimation theory provides fundamental bounds to the estimation precision of physical parameters and the optimal measurements that saturate these limits [23]. We are interested in estimating multiple parameters simultaneously. The prototypical scheme requires that the vector of parameters $\boldsymbol{\theta} = (\theta_1, ..., \theta_d)^\top \in \mathbb{R}^d$ be imprinted on a quantum state $\rho(\boldsymbol{\theta})$. Denoting $\mathbb{H}_D$ as the set of all Hermitian matrices in the Hilbert space of dimension D, we can see that $\rho(\boldsymbol{\theta})$ is a positive semidefinite matrix in $\mathbb{H}_D$ with unit trace. We define measurement operators via a positive operator valued measure (POVM):

$$\boldsymbol{\Pi} = \left\{ \Pi_\omega \geq 0, \omega \in \Omega \,\middle|\, \sum_{\omega\in\Omega}\Pi_\omega = \mathbb{1}_D \right\}, \quad (1)$$

where $\mathbb{1}_D$ denotes the identity operator and Ω is the set of measurement outcomes. The outcomes of such a measurement can be used in a function called the estimator $\check{\boldsymbol{\theta}}$, 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 [9,34–36,46–50].

The performance of the estimator $\check{\boldsymbol{\theta}}$ under any measurement can be quantified in terms of its mean square error (MSE) matrix:

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\boldsymbol{\Pi}, \check{\boldsymbol{\theta}}) = \sum_{\boldsymbol{\omega}\in\Omega^N} p(\boldsymbol{\omega}|\boldsymbol{\theta})[\check{\boldsymbol{\theta}}(\boldsymbol{\omega}) - \boldsymbol{\theta}][\check{\boldsymbol{\theta}}(\boldsymbol{\omega}) - \boldsymbol{\theta}]^\top, \quad (2)$$

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

$$\sum_{\boldsymbol{\omega}\in\Omega^N} [\check{\theta}_j(\boldsymbol{\omega}) - \theta_j]p(\boldsymbol{\omega}|\boldsymbol{\theta}) = 0,$$
$$\sum_{\boldsymbol{\omega}\in\Omega^N} \check{\theta}_j(\boldsymbol{\omega})\partial_k p(\boldsymbol{\omega}|\boldsymbol{\theta}) = \delta_{jk}, \quad (3)$$

where $\partial_k \equiv \partial/\partial\theta_k$. Under these conditions, the MSE matrix is equivalent to the covariance matrix of parameter estimates and is lower bounded through generalizations of the Cramér-Rao bound from classical statistics:

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\boldsymbol{\Pi}, \check{\boldsymbol{\theta}}) \geq \mathscr{F}[\rho(\boldsymbol{\theta}), \boldsymbol{\Pi}]^{-1}, \quad (4)$$

where $\mathscr{F}$ is the classical Fisher information matrix [5,6]. The Fisher information characterizes the MSE matrix for the best classical data manipulation given a measurement strategy in the asymptotic limit [27]. A well-known quantum generalization includes the symmetric logarithmic derivative (SLD) $L_j \in \mathbb{H}_N$, which is implicitly defined through $2\partial_j\rho = \{L_j, \rho\}$ and generates the real symmetric quantum Fisher information matrix (QFIM) $\mathscr{I}_{jk}^{\mathsf{S}} = \text{Re}[\text{Tr}[\rho L_j L_k]]$ [4,51]. This matrix is referred to as the SLD QFIM. Notice that, for ease of notation, we drop the explicit dependence of the state on the vector of parameters $\boldsymbol{\theta}$. Similarly, the right logarithmic derivative (RLD) $R_j \in \mathbb{H}_N$, defined through $\partial_j\rho = \rho R_j$, induces the complex Hermitian RLD QFIM $\mathscr{I}_{jk}^{\mathsf{R}} = \text{Tr}[\rho R_j R_k]$ [52,53].

There is a fundamental trade-off on how well each parameter in $\boldsymbol{\theta}$ can be simultaneously estimated. Hence, a meaningful multiparameter estimation protocol minimizes the weighted sum of parameter estimate variances. For this estimation, a size d positive definite square weight matrix W is chosen to define the weighted mean square error (WMSE) $\mathrm{Tr}[W\boldsymbol{\Sigma_\theta}]$. Holevo proves an equivalence between a matrix inequality and its corresponding scalar inequality, which allows the WMSE to be optimally minimized. In particular, for any real symmetric V and W and Hermitian M, the inequality $V \geq M$ implies that $\mathrm{Tr}[WV] \geq \mathrm{Tr}[\mathrm{Re}[WM]] + \mathrm{Tr}|\sqrt{W}\mathrm{Im}[M]\sqrt{W}|$ (Lemma 6.6.1 in Ref. [6]), where $\mathrm{Re}[\cdot]$ and $\mathrm{Im}[\cdot]$ denote the real and imaginary part, respectively, of each matrix element and $\mathrm{Tr}|\cdot|$ denotes the sum of the absolute values of the eigenvalues of a matrix [54]. Since the sum of the absolute values of the eigenvalues of a matrix is, in fact, the sum of the singular values of a matrix, the function $\mathrm{Tr}|\cdot|$ is equivalent to the more commonly used trace norm $\|\cdot\|_1$. Given that the covariance matrix is always real and symmetric, we identify the matrix V with $\boldsymbol{\Sigma_\theta}$. Hence, we can write the WMSE as

$$\mathrm{Tr}[W\boldsymbol{\Sigma_\theta}] \geq \mathrm{Tr}[W\mathrm{Re}[M]] + \|\sqrt{W}\mathrm{Im}[M]\sqrt{W}\|_1. \quad (5)$$

Notice that the scalar cost function in Eq. (5) appropriately assigns individual priority weights to different parameters. For a given weight matrix and Hermitian matrix M, we want to minimize the scalar WMSE to derive better parameter estimates.

Now, we identify M with the inverse of the family of definitions for the quantum Fisher information matrices to generate different lower bounds on the scalar WMSE cost function. Specifically, the matrices $\mathscr{I}^{\mathsf{S}}$ and $\mathscr{I}^{\mathsf{R}}$ generate the following scalar cost functions on the SLD QCRB and RLD QCRB:

$$C_{\mathsf{S}}(\boldsymbol{\theta}) = \mathrm{Tr}[W[\mathscr{I}^{\mathsf{S}}]^{-1}], \quad (6)$$

$$C_{\mathsf{R}}(\boldsymbol{\theta}) = \mathrm{Tr}[W\mathrm{Re}[\mathscr{I}^{\mathsf{R}}]^{-1}] + \|\sqrt{W}\mathrm{Im}[\mathscr{I}^{\mathsf{R}}]^{-1}\sqrt{W}\|_1, \quad (7)$$

respectively. Nagaoka investigates in detail the relationship between these bounds [55]. The central problem in quantum estimation theory is the minimization of these scalar bounds over the family of probability of distributions defined by quantum measurements.

The SLD and RLD QCRB do not always provide the best bounds to parameter estimates. For example, the attainability of the SLD QCRB does not generally hold for multiple parameter estimations [8]. Intuitively, any incompatibility among the parameters $\boldsymbol{\theta}$ prohibits the simultaneous optimal estimation of all parameters. Correspondingly, the RLD QCRB is not always attainable, since the optimal estimators derived from the RLD may not correspond to physical POVMs [56].

The problem with saturability of the multiparameter bound is noted by Holevo, who provides the most general quantum extension to the classical Cramér-Rao bound, called the Holevo Cramér-Rao bound (HCRB). Specifically, if a vector of Hermitian observables $\boldsymbol{X} = (X_1, \ldots, X_d)$ satisfies the locally unbiased conditions $\mathrm{Tr}[\rho X_j] = 0$ and $\mathrm{Tr}[\partial_j \rho X_k] = \delta_{jk}$, its covariance matrix $Z(\boldsymbol{X})$ with matrix elements $[Z(\boldsymbol{X})]_{jk} = \mathrm{Tr}[\rho X_j X_k]$ satisfies the inequalities [5,6]

$$Z(\boldsymbol{X}) \geq [\mathscr{I}^{\mathsf{S}}]^{-1} \quad \text{and} \quad Z(\boldsymbol{X}) \geq [\mathscr{I}^{\mathsf{R}}]^{-1}. \quad (8)$$

From this result, it is clear that identifying M in Eq. (5) with the Hermitian matrix $Z(\boldsymbol{X})$, such that

$$\mathrm{Tr}[W\boldsymbol{\Sigma_\theta}] \geq \mathrm{Tr}[W\mathrm{Re}[Z(\boldsymbol{X})]] + \|\sqrt{W}\mathrm{Im}[Z(\boldsymbol{X})]\sqrt{W}\|_1, \quad (9)$$

we have a tighter bound on the scalar WMSE than either of the bounds in Eq. (6) or Eq. (7). By optimizing the objective function in Eq. (9) subject to appropriate unbiasedness constraints on X, we obtain the tightest bound on the WMSE. This optimization defines the HCRB $C_{\mathsf{H}}(\boldsymbol{\theta})$, which explicitly is the minimum of the following minimization problem [55]:

$$\begin{aligned}
\underset{X_1, \ldots, X_d}{\text{minimize}} \quad & \mathrm{Tr}[W\mathrm{Re}Z(\boldsymbol{X})] + \|\sqrt{W}\mathrm{Im}[Z(\boldsymbol{X})]\sqrt{W}\|_1, \\
\text{subject to} \quad & \mathrm{Tr}[\rho X_j] = 0 \\
& \mathrm{Tr}[\partial_j \rho X_k] = \delta_{jk}.
\end{aligned} \quad (10)$$

The HCRB is the best asymptotically attainable precision with global, unbiased measurements of a set of parameters. By minimizing over only the first term in the objective function of Eq. (10), we obtain the SLD QCRB [55]

$$C_{\mathsf{S}}(\boldsymbol{\theta}) = \mathrm{Tr}[W[\mathscr{I}^{\mathsf{S}}]^{-1}] = \min_{X_1, \ldots, X_d} \mathrm{Tr}[W\mathrm{Re}Z]. \quad (11)$$

This result shows that the HCRB is a tighter bound than the SLD QCRB, since the second term in Eq. (10) is non-negative [33]. In fact, the HCRB is more informative than both the scalar SLD and RLD QCRBs, satisfies the inequality [57]

$$\mathrm{Tr}[W\boldsymbol{\Sigma_\theta}] \geq C_{\mathsf{H}}(\boldsymbol{\theta}) \geq \max\{C_{\mathsf{S}}(\boldsymbol{\theta}), C_{\mathsf{R}}(\boldsymbol{\theta})\}, \quad (12)$$

and gives the best asymptotically attainable precision with global, unbiased measurements of a set of parameters. Specifically, Helstrom [5] and Holevo [57] demonstrate that $C_{\mathsf{H}}(\boldsymbol{\theta})$ is attainable if the locally unbiased equality constraints in Eq. (3) are satisfied.

We note that the HCRB is not defined explicitly in terms of a closed form for a given statistical model. This constraint is in contrast to the classical case, where the Fisher information can be readily determined from a given

statistical model. Recent efforts focus on determining upper bounds to the HCRB [58–60]. Specifically, the HCRB is upper bounded by a quantity that is twice the SLD QCRB, such that [58]

$$\max\{C_S(\boldsymbol{\theta}), C_R(\boldsymbol{\theta})\} \le C_H(\boldsymbol{\theta}) \le 2C_S(\boldsymbol{\theta}). \tag{13}$$

In this paper, we provide an analytic solution to the HCRB and provide conditions on when it is tight.

The HCRB is the best asymptotically achievable bound under the conditions stated in Refs. [34–36,50]. Both inequalities in Eq. (12) can be tight [61]. For instance, consider the skew-symmetric matrix $\text{Im}\{\text{Tr}[L_j L_k \rho(\boldsymbol{\theta})]\}$. When

$$\text{Im}\{\text{Tr}[L_j L_k \rho(\boldsymbol{\theta})]\} = 0, \tag{14}$$

we have $C_H(\boldsymbol{\theta}) = C_S(\boldsymbol{\theta})$ [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 optimization theory to address the minimization 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 to calculate bounds on the Holevo Cramér-Rao bound C_H. We first derive simple analytic upper and lower bounds for C_H for two parameters in Sec. III A. We show how these bounds are generated by studying the optimization problem using the method of Lagrange multipliers. This method has the advantage of reducing the complexity involved in evaluating bounds on C_H to that of solving two sets of linear equations. In Sec. III B, we focus on deriving lower bounds on C_H for more than two parameters. At the expense of additional analysis, our formalism can be extended to also provide tight analytic solutions to the HCRB. We demonstrate this extension in Sec. III C, where we provide a complete exposition of analytic bounds on the two-parameter HCRB and provide conditions for when the bounds are tight.

A. Simple bounds in the two-parameter setting

We first consider the HCRB for two parameters $\boldsymbol{\theta} = (\theta_1, \theta_2)^\top$. To obtain simple analytic bounds to the HCRB, we must define the weight matrix W for the scalar bound. For simplicity, we use the identity weight matrix and determine upper and lower bounds to the two-parameter HCRB using optimization theory [63]. We want to solve the minimization in Eq. (10), which is convex but not quadratic. Hence, we first manipulate Eq. (10) into a quadratic form in the variables X_1 and X_2. Then, such an optimization problem can be studied analytically using the method of Lagrange multipliers.

Choosing $Y = X_1 + iX_2$, Eq. (10) can be written as an optimization program (see Appendix B):

$$\underset{Y,t}{\text{minimize}}\, t,$$

subject to $\text{Tr}[Y\rho Y^\dagger] \le t, \quad \text{Tr}[Y^\dagger \rho Y] \le t,$

$$\text{Tr}[\rho Y] = 0, \quad \text{Tr}[\partial_1 \rho Y] = 1, \quad \text{Tr}[\partial_2 \rho Y] = i. \tag{15}$$

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 minimization in Eq. (10). Here, Y is optimized over all complex matrices of dimension D and is, in general, not a Hermitian matrix. By mapping Y and t into a real vector $\mathbf{x}$, we cast this optimization program into the standard form of

$$\underset{\mathbf{x}}{\min}\{f(\mathbf{x}) : c_i(\mathbf{x}) \le 0, h_i(\mathbf{x}) = 0\}, \tag{16}$$

where $f(\mathbf{x})$ is a real linear objective function, while $h_i(\mathbf{x})$ and $c_i(\mathbf{x})$ are the corresponding equality and inequality constraint functions, respectively, that must also be real. Equation (15) 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 optimization theory, we check whether Slater's constraint qualification holds, which amounts to checking that all the inequality constraints in Eq. (15) can strictly hold. Since t can be arbitrarily large, this result indeed is the case. The optimality conditions for a continuous optimization program are best stated in terms of the Lagrangian of Eq. (15), given by

$$\mathsf{L}(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{z}) = f(\mathbf{x}) + \sum_{i=1}^{2} \lambda_i c_i(\mathbf{x}) + \sum_{i=1}^{6} z_i h_i(\mathbf{x}), \tag{17}$$

where the coefficients $\lambda_i \ge 0$ and $z_i \in \mathbb{R}$ are Lagrange multipliers for the inequality and equality constraints, respectively. Since Eq. (16) is a convex program and Slater's constraint qualification holds, the first-order Karush-Kuhn-Tucker (KKT) conditions of stationarity, primal and dual feasibility, and complementary slackness are necessary and sufficient [63] to determine the optimality of Eq. (15).

For our problem, we have dual variables $\boldsymbol{\lambda} = (\lambda_1, \lambda_2) = (u, v)^\top$ and $\mathbf{z} = (z_1, ..., z_6)^\top$, which are vectors of Lagrange multipliers. The primal variables are Y and t, and the Lagrangian is given by

$$\mathsf{L}(Y, t, u, v, \mathbf{z}) = t(1 - u - v) - \mathbf{b}^\top \mathbf{z} + u\text{Tr}[Y\rho Y^\dagger]$$
$$+ v\text{Tr}[Y^\dagger \rho Y] + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger]. \tag{18}$$

Here, $\mathbf{b} = (0, 1, 0, 0, 0, 1)^\top$ is a column vector that encodes the equality constraints in Eq. (15), constructed

in Appendix B 2. The operator A is a linear superposition of ρ and its derivatives:

$$A = z_1 A_1 + \cdots + z_6 A_6, \tag{19}$$

where

$$
\begin{aligned}
A_1 &= \frac{1}{2}\rho, & A_4 &= -iA_1, \\
A_2 &= \frac{1}{2}\partial_1 \rho, & A_5 &= -iA_2, \\
A_3 &= \frac{1}{2}\partial_2 \rho, & A_6 &= -iA_3.
\end{aligned}
\tag{20}
$$

Because of the duality principle in optimization theory [63], we may equivalently view the optimization by considering the Lagrange dual function $g(\lambda, \mathbf{z}) = \inf_{\mathbf{x}} L(\mathbf{x}, \lambda, \mathbf{z})$ of Eq. (17). Since the Lagrangian L is quadratic in $\mathbf{x}$, the Lagrange dual can be found analytically by an unconstrained minimization of the Lagrangian with respect to $\mathbf{x}$ for fixed values of the dual variables λ and $\mathbf{z}$ [63]. Because of the structure of the Lagrangian in Eq. (18), the Lagrange dual is never unbounded from below whenever $u + v = 1$. Hence, maximizing the Lagrange dual function corresponds to an unconstrained maximization problem. Since the Lagrange dual is also a quadratic function in terms of its dual variables $\mathbf{z}$, it can be easily maximized exactly with respect to $\mathbf{z}$.

Note that our Lagrange dual is not a quadratic function with respect to $\lambda = (u, v)$. To bound C_H, it suffices to evaluate the Lagrangian for feasible values of (u, v) that satisfy $u + v = 1$. Two such values are the boundary values $(u, v) = \{(0, 1), (1, 0)\}$, for which the Lagrangian in Eq. (18) is greatly simplified. For each case, we first determine the stationary point of the resulting Lagrangian with respect to Y, where Y has an implicit dependence on $\mathbf{z}$, and then perform a maximization over $\mathbf{z}$. By evaluating the primal and dual objective functions, we obtain simple analytic two-sided bounds for C_H. Specifically, an analytic lower bound $\mathscr{L}$ to the HCRB is determined through finding $\mathbf{z} \in \mathbb{R}^6$ that solves

$$2\mathrm{Re}(Q_j)\mathbf{z} + \mathbf{b} = 0, \quad j = 1, 2, \tag{21}$$

where Q_j has the matrix elements

$$[Q_1]_{ik} = \mathrm{Tr}[A_i^\dagger \rho^{-1} A_k] \quad \text{and} \quad [Q_2]_{ik} = \mathrm{Tr}[A_i \rho^{-1} A_k^\dagger]. \tag{22}$$

Details for this situation are delegated to Appendix B. The matrices $\mathrm{Re}(Q_j)$ are full rank when the derivatives $\partial_1 \rho$ and $\partial_2 \rho$ are linearly independent. Armed with these dual variables $\mathbf{z}$, we collect the result of this optimization in the following theorem.

Theorem 1.—Let $\partial_1 \rho$ and $\partial_2 \rho$ be linearly independent. With Q_j defined in Eq. (22) and the matrices $A_1, \ldots, A_6$ given in Eq. (20), the HCRB C_H for two parameters satisfies the inequality

$$\max_{j=1,2}\{l_j\} = \mathscr{L} \le C_H \le \mathscr{U} = \min_{j=1,2}\{\max\{l_j, m_j\}\},$$

where

$$l_j = \frac{1}{4}\mathbf{b}^\top \mathrm{Re}(Q_j)^{-1}\mathbf{b}, \tag{23}$$

$$m_j = \sum_{a,b=1}^{6} \mathrm{Tr}[\rho^{-2}A_a \rho A_b^\dagger] z_{a,j} z_{b,j}, \tag{24}$$

and

$$z_{a,j} = -\frac{1}{2}\{[\mathrm{Re}(Q_j)^{-1}]_{a2} + [\mathrm{Re}(Q_j)^{-1}]_{a6}\}. \tag{25}$$

For a detailed proof of this theorem, consult Appendixes B 2–B 5. Theorem 1 gives a simple procedure for finding analytic upper and lower bounds to two-parameter HCRB. Notice that the complexity of determining these bounds is commensurate with linear equation solvers that are used in determining the Lagrange dual variables.

Algorithm 1 $\{\mathscr{L}, \mathscr{U}, X_1, X_2\} = \text{SIMPLE}\ (\rho, \partial_1\rho, \partial_2\rho)$

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: $\mathbf{b} \leftarrow [0, 1, 0, 0, 0, 1]^\top$
8: $Q_1, Q_2 \leftarrow$ size 6 zeros matrix
9: **for all** $j, k = 1$ to 6 **do**
10: $\quad [Q_1]_{jk} = \mathrm{Tr}[A_j^\dagger \rho^{-1} A_k]$
11: $\quad [Q_2]_{jk} = \mathrm{Tr}[A_j \rho^{-1} A_k^\dagger]$
12: Solve for $\mathbf{y} = [y_1, \ldots, y_6]^T$ in $2\mathrm{Re}(Q_1)\mathbf{y} + \mathbf{b} = 0$
13: Solve for $\mathbf{z} = [z_1, \ldots, z_6]^T$ in $2\mathrm{Re}(Q_2)\mathbf{z} + \mathbf{b} = 0$
14: $v_1 \leftarrow -\mathbf{y}^\top \mathrm{Re}(Q_1)\mathbf{y} - \mathbf{b}^\top \mathbf{y}$
15: $v_2 \leftarrow -\mathbf{z}^\top \mathrm{Re}(Q_2)\mathbf{z} - \mathbf{b}^\top \mathbf{z}$
16: **if** $v_1 > v_2$ **then**
17: $\quad \mathscr{L} \leftarrow v_1$
18: $\quad A \leftarrow y_1 A_1 + \cdots + y_6 A_6$
19: $\quad Y \leftarrow -A^\dagger \rho^{-1}$
20: $\quad \mathscr{U} \leftarrow \max\{\mathrm{Tr}[Y\rho Y^\dagger], \mathrm{Tr}[Y^\dagger \rho Y]\}$
21: **else**
22: $\quad \mathscr{L} \leftarrow v_2$
23: $\quad A \leftarrow z_1 A_1 + \cdots + z_6 A_6$
24: $\quad Y \leftarrow -\rho^{-1} A^\dagger$
25: $\quad \mathscr{U} \leftarrow \max\{\mathrm{Tr}[Y\rho Y^\dagger], \mathrm{Tr}[Y^\dagger \rho Y]\}$
26: $X_1 \leftarrow (Y + Y^\dagger)/2$
27: $X_2 \leftarrow (Y - Y^\dagger)/(2i)$
28: Return $\{\mathscr{L}, \mathscr{U}, X_1, X_2\}$

FIG. 1. Pseudocode to determine simple bounds to the two-parameter HCRB and its associated optimal measurement observables X_1 and X_2. Note that this algorithm depends only on the state ρ and its two derivatives $\partial_1 \rho$ and $\partial_2 \rho$.

This procedure makes these bounds readily accessible for general two-parameter applications. Figure 1 shows the pseudocode for this procedure.

Next, we establish how to construct the observables that saturate these bounds. Specifically, there are two choices for Y that minimize the Lagrangian in Eq. (18), corresponding to the two choices for (u, v):

$$(u, v) = (1, 0): \quad Y = -A^\dagger \rho^{-1}, \qquad (26)$$

$$(u, v) = (0, 1): \quad Y = -\rho^{-1} A^\dagger, \qquad (27)$$

that correspond to the choice Q_1 and Q_2, respectively. Then, using Eq. (19) and the optimized values for $\mathbf{z}$, we construct the analytic form for the observables X_1 and X_2. Since the matrices Q_j, $j \in \{1, 2\}$, are only six-dimensional matrices, determining $\text{Re}(Q_j)$ is easy, and, hence, it is straightforward to find analytic bounds to the two-parameter HCRB. The procedure is shown algorithmically in Fig. 1.

Finally, we note that if our lower bound to C_H is strictly larger than C_S, then we know that the skew-symmetric matrix $\text{Im}\{\text{Tr}[L_j L_k \rho(\boldsymbol{\theta})]\}$ cannot be equal to zero, and the weak commutativity criterion does not hold.

B. Lower bound in the multiparameter setting

For more than two parameters, we can also use the method of Lagrange multipliers to bound the HCRB. 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 $\text{ReTr}[Z] + \|\text{Im}Z\|_1$ that appears in the objective function of Eq. (10).

Algorithm 2 $v = g_\alpha(Z, \xi_1, \ldots, \xi_d)$

1: $S \leftarrow$ size d zeros matrix
2: **for all** $j = 1$ to d **do**
3:     $\rho_j \leftarrow \partial_j \rho$
4:     $S_{jj} = 1$
5:     **if** $j \neq d$ **then**
6:         $k \leftarrow j + 1$
7:         $S_{jk} = i$
8:     **else**
9:         **if** d is a multiple of 4 **then**
10:             $S_{d1} = -i$
11:             $\alpha_d \leftarrow 1 - \alpha_d$
12:         **else**
13:             $S_{d1} = i$
14: $T \leftarrow S^{-1}$
15: $v \leftarrow -\sum_{j=1}^d [Z]_{jj}$
16: **for all** $j = 1$ to d **do**
17:     $\Gamma_j \leftarrow \sum_{k=1}^d T_{kj} \left(i\xi_k + \sum_{l=0}^d [Z]_{lk} \rho_l \right)$
18:     **if** $\alpha_j = 0$ **then**
19:         $v \leftarrow v - \text{Tr}[\Gamma_j \rho^{-1} \Gamma_j^\dagger]/2$
20:     **else**
21:         $v \leftarrow v - \text{Tr}[\Gamma_j^\dagger \rho^{-1} \Gamma_j]/2$
22: **Return** v

FIG. 2. Pseudocode to generate the Lagrange dual functions defined in Eq. (33).

However, for the corresponding generalization to more parameters, $\text{ReTr}[Z] + \|\text{Im}Z\|_1$ is no longer a quadratic form in the variables X_j. For example, for three parameters, Z takes the form

$$Z(X) = \begin{pmatrix} \text{Tr}[\rho X_1^2] & \text{Tr}[\rho X_1 X_2] & \text{Tr}[\rho X_1 X_3] \\ \text{Tr}[\rho X_2 X_1] & \text{Tr}[\rho X_2^2] & \text{Tr}[\rho X_2 X_3] \\ \text{Tr}[\rho X_3 X_1] & \text{Tr}[\rho X_3 X_2] & \text{Tr}[\rho X_3^2] \end{pmatrix}. \quad (28)$$

The trace norm of $\text{Im}Z$ is related to the eigenvalues of $\text{Im}Z$, and the eigenvalues of a 3×3 matrix involve a cubic equation. This relation renders evaluating the trace norm incompatible with our methodology. To address this incompatibility, we obtain a lower bound to $\|\text{Im}Z\|_1$ that allows $\text{ReTr}[Z] + \|\text{Im}Z\|_1$ to be written as a quadratic form. As shown in Appendix C, this form yields an optimization problem whose optimal value is a lower bound to the HCRB and which is given by

$$\min\{t : \text{Tr}[\rho X_j] = 0, \text{Tr}[\partial_j \rho X_k] = \delta_{jk}, V_\alpha \leq t\}, \quad (29)$$

where the minimization is performed over t and the Hermitian matrices $X_1, \ldots, X_d$, with $j, k = \{1, \ldots, d\}$, and the inequality constraint V_α is a function of a binary string α such that

$$V_\alpha = \frac{1}{2} \sum_{j=1}^d \text{Tr}[[X_j + (-1)^{\alpha_j} i X_{j+1}] \rho [X_j + (-1)^{\alpha_j} i X_{j+1}]^\dagger], \quad (30)$$

with $X_{d+1} = X_1$. The inequality constraints V_α arise from the structure of our lower bound on the trace norm of $\text{Im}Z$ (see Appendix C). By substituting $Y_j = \sum_{k=1}^d S_{jk} X_j$, where

$$S = \begin{cases} \sum_{j \in \mathbb{Z}_d} (|j\rangle\langle j| + i|j\rangle\langle j \oplus 1|) & d \neq 0 \ (\text{mod } 4) \\ \sum_{j \in \mathbb{Z}_d} (|j\rangle\langle j| + (-1)^{\delta_{j,d}} i|j\rangle\langle j \oplus 1|) & \text{otherwise}, \end{cases} \quad (31)$$

we can write the matrices X_j in terms of the matrices Y_j, as before. We next 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 optimization 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 $\mathbf{v} \in \mathbb{R}^{2^d}$ for the inequality constraints, $\mathbf{z} \in \mathbb{R}^{d(d+1)}$ 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. (30). We minimize the Lagrangian constructed from the optimization problem in Eq. (29). Since the Lagrangian is a convex quadratic form in the variables $Y_1, \ldots, Y_d$, it can be minimized exactly. When this minimization is done, we obtain the Lagrange dual function, which depends only on the Lagrange multipliers $\mathbf{v}$, $\mathbf{z}$, and $\xi_1, \ldots, \xi_d$. The Lagrange dual function always gives a lower bound for the primal optimization problem.

While the Lagrange dual is quadratic in $\mathbf{z}$ and $\xi_1, \ldots, \xi_d$, it is not quadratic in $\mathbf{v}$. By minimizing the Lagrangian over t and using the KKT conditions, we conclude as before that the sum of the components in $\mathbf{v}$ is 1. We obtain a lower bound for the Lagrange dual by maximizing over a discrete set of feasible Lagrange multipliers $\mathbf{v}$, which corresponds to the tightness of the constraints $V_\alpha \leq t$. Thus, we create a quadratic optimization 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 $\mathbf{v}$, the Lagrangian is quadratic in $\{Y_1, \ldots, Y_d\}$ and can be minimized individually for each Y_j. The coefficients for Y_j in the Lagrangian are given by Γ_j, where

$$\Gamma_j = \sum_{k=1}^{d} T_{k,j} \left(\sum_{l=0}^{d} Z_{l,k} \rho_l + i\xi_k \right), \qquad (32)$$

where $\rho_0 = \rho$, $\rho_j = \partial_j \rho$ for $j = \{1, \ldots, d\}$, and $T_{k,j}$ are matrix elements that relate the Y_j to the X_k. Specifically, T is the matrix inverse of S. Then, the optimal value for the Lagrange multipliers can be obtained by maximizing the Lagrange dual functions

$$g_\alpha = -\sum_{j=1}^{d} z_{j,j} - \sum_{j=1}^{d} \frac{\delta_{0,\bar{\alpha}_j} \mathrm{Tr}[\Gamma_j \rho^{-1} \Gamma_j^\dagger] + \delta_{1,\bar{\alpha}_j} \mathrm{Tr}[\Gamma_j^\dagger \rho^{-1} \Gamma_j]}{2}$$

$$(33)$$

with respect to the scalar variables $z_{j,k}$ and the Hermitian variables ξ_j, where $\bar{\boldsymbol{\alpha}} = \boldsymbol{\alpha}$ when d is not a multiple of 4, and when d is a multiple of 4, then $\bar{\boldsymbol{\alpha}}$ differs from $\boldsymbol{\alpha}$ by simply flipping the last bit. Our lower bound to C_{H} in the multiparameter setting is then given by the following theorem.

Theorem 2.—Let $d \geq 3$, $\mathbf{z} \in \mathbb{R}^{d(d+1)}$, and $\xi_1, \ldots, \xi_d$ are Hermitian matrices. Then,

$$C_{\mathsf{H}} \geq \max_{\boldsymbol{\alpha} \in \{0,1\}^d} \max_{\mathbf{z}, \xi_1, \ldots, \xi_d} g_\alpha,$$

where g_α is given by Eq. (33).

This optimization problem can be solved exactly using a single step of Newton's method. It requires the input state ρ

Algorithm 3 $\{\mathscr{L}, X_1, \ldots, X_d\} = \mathrm{MULTI}\,(\rho, \partial_1 \rho, \ldots, \partial_d \rho)$

```
1:  ρ₀ ← ρ
2:  S ← size d zeros matrix
3:  for all j = 1 to d do
4:      ρⱼ ← ∂ⱼρ
5:      Sⱼⱼ = 1
6:      if j ≠ d then
7:          k ← j + 1
8:          Sⱼₖ = i
9:      else
10:         if d is a multiple of 4 then
11:             S_{d1} = -i
12:         else
13:             S_{d1} = i
14: T ← S⁻¹
15: for all binary vectors α = (α₁,...,α_d)  do
16:     Solve max gα(Z, ξ₁,...,ξ_d) w.r.t. the real matrix Z.
        Z: rows labelled from 0 to d, columns labelled from 1 to d
        ξ₁,...,ξ_d: complex size d Hermitian matrices
17:     Let (Z_opt, ξ_{opt,1},...,ξ_{opt,d}) denote the optimal solution
18:     vα ← gα(Z_opt, ξ_{opt,1},...,ξ_{opt,d}) (See Algorithm 2)
19: 𝓛 ← 0
20: for all binary vectors α = (α₁,...,α_d)  do
21:     if d is a multiple of 4 then
22:         α_d ← 1 − α_d
23:     if vα > 𝓛 then
24:         𝓛 ← vα
25:         for all j = 1 to d do
26:             Γⱼ ← Σ_{k=1}^d T_{k,j} (iξ_{opt,k} + Σ_{l=0}^d [Z_opt]_{l,k}ρ_l)
27:             if αⱼ = 0  then
28:                 Yⱼ ← −Γⱼ†ρ⁻¹
29:             else
30:                 Yⱼ ← −ρ⁻¹Γⱼ†
31: for all k = 1 to d do
32:     X_k ← Σ_{j=1}^d T_{k,j}Yⱼ
33: Return {𝓛, X₁,...,X_d}
```

FIG. 3. Pseudocode to generate a lower bound to the HCRB for multiple parameters. The pseudocode to generate the Lagrange dual functions is shown in Fig. 2

and its derivatives $\partial_j \rho$. The algorithm to implement this lower bound is illustrated in Fig. 3.

C. Tight two-parameter bounds

Notice that, for Theorem 1, we constrained the values of the Lagrange multiplier u to two values. This constraint does not provide the most general case, and, as a result, the analysis can generate observables that are not always optimal. That is, the corresponding upper and lower bounds are not always tight. By lifting this restriction, we expand the analysis to explore the full generality of our formalism to generate tight bounds to the estimation of incompatible observables. As we observe in this section, this expansion is necessary to develop an intuition into multiparameter quantum estimation that is captured by the construction of the HCRB. To achieve this expansion, we revisit the two-parameter scenario. Specifically, for fixed u, we minimize the Lagrangian and find the optimal observables that attain these stationary points. In doing so, for every feasible value of u, we obtain an upper and lower bound on the HCRB. Since the lower bound is a concave and smooth function,

then optimization theory guarantees a solution to both the HCRB and the observables that attain it.

Recall the Lagrangian in Eq. (18), which for $u + v = 1$ becomes

$$\mathsf{L}(Y, u, \mathbf{z}) = -\mathbf{b}^\top \mathbf{z} + u\mathrm{Tr}[Y\rho Y^\dagger] + (1 - u)\mathrm{Tr}[Y^\dagger \rho Y]$$
$$+ \mathrm{Tr}[AY] + \mathrm{Tr}[A^\dagger Y^\dagger]. \quad (34)$$

As we show in Appendix D, this Lagrangian is minimized when Y is chosen such that

$$uY\rho + (1 - u)\rho Y = -A^\dagger. \quad (35)$$

Notice that Y, and hence the optimal observables, is the solution to a Sylvester equation. When $\mathrm{Im}Z = 0$, we know that the observables X_1 and X_2 commute and the weak commutativity criterion is preserved. Then, Eq. (35) reduces to solving a Lyapunov equation, where Y generates the well-explored SLD. This case corresponds to

$$|\mathrm{Tr}[\rho X_1 X_2] - \mathrm{Tr}[\rho X_2 X_1]| = 0, \quad (36)$$

for the optimal X_1 and X_2, which recovers the weak commutativity condition. In this way, the solution to Y in Eq. (35) provides the most general definition to quantum logarithmic derivatives for multiple incompatible parameters $\boldsymbol{\theta}$. From our definition of A, it defines exactly how the optimal observables depend on dynamics in both parameters.

Similar to analytic solutions to the SLD, given the spectral decomposition of the state $\rho = \sum_j p_j |e_j\rangle\langle e_j|$, we can analytically solve Eq. (35) to obtain the Y that minimizes the Lagrangian:

$$Y = -\sum_{j,k}[up_k + (1 - u)p_j]^{-1}\langle e_j|A^\dagger|e_k\rangle|e_j\rangle\langle e_k|. \quad (37)$$

By taking the trace, it is clear to see that Y has a zero expectation value. This result recovers the unbiasedness condition on the observables as required. Using $Y = X_1 + iX_2$ and the definition for Y in Eq. (37), we can write an analytic solution for the observables that saturate the HCRB in terms of the optimal Lagrange multipliers $\mathbf{z}$. Specifically, defining the statistical admixture $\varrho^{(s)} = \sum_{l=0}^{2} z_{l+s}\rho_l$, then in the eigenbasis of ρ, we have

$$[X_1]_{jk} = \frac{i(p_j - p_k)(1 - 2u)[\varrho^{(4)}]_{jk} - (p_j + p_k)[\varrho^{(1)}]_{jk}}{4[up_k + (1 - u)p_j][up_j + (1 - u)p_k]}, \quad (38)$$

$$[X_2]_{jk} = -\frac{i(p_j - p_k)(1 - 2u)[\varrho^{(1)}]_{jk} + (p_j + p_k)[\varrho^{(4)}]_{jk}}{4[up_k + (1 - u)p_j][up_j + (1 - u)p_k]}, \quad (39)$$

with $[\varrho^{(s)}]_{jk} = \langle e_j|\varrho^{(s)}|e_k\rangle$. The Hermiticity of the state and its derivatives guarantees the Hermiticity of these observables such that $[X_l]_{jk} = [X_l]_{kj}^*$. Equations (38) and (39) show exactly how each observable depends on the dynamics of each parameter.

With access to the spectral decomposition of ρ, we can also find analytic expressions to the HCRB. The master algorithm in Fig. 4 concisely clarifies this procedure.

Algorithm 4 $\{\mathscr{L}, \mathscr{U}, X_1, X_2\} = \text{MASTER}(\rho, \partial_1\rho, \partial_2\rho)$

1: Find $u \in [0, 1]$ that maximizes
    $\mathscr{L}_u$ that is returned from $\text{HCRBU}(u, \rho, \partial_1\rho, \partial_2\rho)$
2: $\{\mathscr{L}_u, \mathscr{U}_u, Y\} = \text{HCRBU}(u, \rho, \partial_1\rho, \partial_2\rho)$
3: $X_1 \leftarrow (Y + Y^\dagger)/2$
4: $X_2 \leftarrow (Y - Y^\dagger)/(2i)$
5: Return $\{\mathscr{L}_u, \mathscr{U}_u, X_1, X_2\}$

Procedure: $\{\mathscr{L}_u, \mathscr{U}_u, Y\} = \text{HCRBU}(u, \rho, \partial_1\rho, \partial_2\rho)$

1: $D \leftarrow$ size of ρ
2: $\{p_j, |e_j\rangle : j = 1, \ldots, D\} \leftarrow$ eigenvalues and eigenvectors of ρ
3: $|e_j\rangle \leftarrow |e_j\rangle/\langle e_j|e_j\rangle$
4: $\gamma_1 \leftarrow -\mathbf{1}_D/2$
5: $\gamma_2 \leftarrow \sum_{j,k=1}^D (up_k + (1-u)p_j)^{-1}\langle e_j|(\partial_1\rho/2)|e_k\rangle|e_j\rangle\langle e_k|$
6: $\gamma_3 \leftarrow \sum_{j,k=1}^D (up_k + (1-u)p_j)^{-1}\langle e_j|(\partial_2\rho/2)|e_k\rangle|e_j\rangle\langle e_k|$
7: $G_1 \leftarrow \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & \mathrm{Tr}[\gamma_2\rho\gamma_2^\dagger] & \mathrm{Tr}[\gamma_2\rho\gamma_3^\dagger] \\ 0 & \mathrm{Tr}[\gamma_3\rho\gamma_2^\dagger] & \mathrm{Tr}[\gamma_3\rho\gamma_3^\dagger] \end{pmatrix}$
8: $G_2 \leftarrow \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & \mathrm{Tr}[\gamma_2^\dagger\rho\gamma_2] & \mathrm{Tr}[\gamma_2^\dagger\rho\gamma_3] \\ 0 & \mathrm{Tr}[\gamma_3^\dagger\rho\gamma_2] & \mathrm{Tr}[\gamma_3^\dagger\rho\gamma_3] \end{pmatrix}$
9: $G_3 \leftarrow \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & \mathrm{Tr}[\partial_1\rho\gamma_2] & \mathrm{Tr}[\partial_1\rho\gamma_3] \\ 0 & \mathrm{Tr}[\partial_2\rho\gamma_2] & \mathrm{Tr}[\partial_2\rho\gamma_3] \end{pmatrix}$
10: $\mathcal{Q}_k \leftarrow \begin{pmatrix} G_k & (-1)^k iG_k \\ (-1)^{k-1}iG_1 & G_k \end{pmatrix}, k = 1, 2.$
11: $\mathcal{Q}_3 \leftarrow \frac{1}{2}\begin{pmatrix} G_3 + G_3^* & i(G_3 - G_3^*) \\ -i(G_3 - G_3^*) & G_3 + G_3^* \end{pmatrix}$
12: $\mathcal{Q} \leftarrow u\mathrm{Re}(\mathcal{Q}_1) + (1 - u)\mathrm{Re}(\mathcal{Q}_2) + \mathrm{Re}(\mathcal{Q}_3)$
13: trigger $\leftarrow 0$
14: KernelDim $\leftarrow 6 - \mathrm{rank}(\mathcal{Q})$
15: $\{\mathbf{ker}_{\text{idx}}\}_{\text{idx}=1:\text{KernelDim}} \leftarrow$ orthonormal basis of $\mathcal{Q}$'s kernel.
16: **for all** idx $= 1 : $ KernelDim **do**
17:     trigger $\leftarrow$ trigger $+ \|\mathbf{b}^\top\mathbf{ker}_{\text{idx}}\|_1$
18: **if** trigger > 0 **then**
19:     $\mathscr{L}_u, \mathscr{U}_u \leftarrow \infty$
20:     $Y \leftarrow$ matrix of NaN of size D
21: **else**
22:     $\mathbf{b} \leftarrow [0, 1, 0, 0, 0, 1]^\top$
23:     **if** $\mathrm{rank}(\mathcal{Q}) = 6$ **then**
24:         Solve for $\mathbf{z} = [z_1, \ldots, z_6]^\top$ in $2\mathcal{Q}\mathbf{z} = \mathbf{b}$
25:     **else**
26:         Find $\mathbf{z} = [z_1, \ldots, z_6]^\top$ that maximizes $-\mathbf{b}^\top\mathbf{z} + \mathbf{z}^\top\mathcal{Q}\mathbf{z}$
27:     $\mathscr{L}_u \leftarrow -\mathbf{b}^\top\mathbf{z} + \mathbf{z}^\top\mathcal{Q}\mathbf{z}$
28:     $\mathscr{U}_u \leftarrow \max(\mathbf{z}^\top\mathcal{Q}_1\mathbf{z}, \mathbf{z}^\top\mathcal{Q}_2\mathbf{z})$
29:     $Y \leftarrow z_1\gamma_1 + z_2\gamma_2 + z_3\gamma_3 + i(z_4\gamma_1 + z_5\gamma_2 + z_6\gamma_3)$
30: Return $\{\mathscr{L}_u, \mathscr{U}_u, Y\}$

FIG. 4. Master algorithm to generate the analytic form of the two-parameter HCRB.

With this procedure, Theorem 3 concretely demonstrates how to construct tight bounds on the HCRB and is central to our result.

Theorem 3.—Let $\mathcal{Q}_1$, $\mathcal{Q}_2$, and $\mathcal{Q}_3$ be matrices defined in the master algorithm in Fig. 4, and let $\mathbf{b} = (0, 1, 0, 0, 0, 1)^\top$ be a column vector. Let $0 < u < 1$ and $\mathcal{Q} = u\mathrm{Re}\mathcal{Q}_1 + (1 - u)\mathrm{Re}\mathcal{Q}_2 + \mathrm{Re}\mathcal{Q}_3$ be a negative definite matrix. Then, when $\mathcal{Q}$ is full rank for all $0 < u < 1$, the HCRB is bounded through $\mathscr{L}_u \leq C_\mathsf{H} \leq \mathscr{U}_u$, where

$$\mathscr{L}_u = -\frac{1}{4}\mathbf{b}^\top \mathcal{Q}^{-1}\mathbf{b}, \tag{40}$$

$$\mathscr{U}_u = \frac{1}{4}\max\{\mathbf{b}^\top \mathcal{Q}^{-1}\mathcal{Q}_1\mathcal{Q}^{-1}\mathbf{b}, \mathbf{b}^\top \mathcal{Q}^{-1}\mathcal{Q}_2\mathcal{Q}^{-1}\mathbf{b}\} \tag{41}$$

with equality on both sides attained at the stationary point of the lower bound with respect to u when $(d/du)(\mathbf{b}^\top \mathcal{Q}^{-1}\mathbf{b}) = [\mathbf{b}^\top \mathcal{Q}^{-1}(d\mathcal{Q}/du)\mathcal{Q}^{-1}\mathbf{b}] = 0$.

We refer the reader to Appendix D for a complete proof of this theorem. For applications where the spectral decomposition of the state is not known, the Sylvester equation (37) can be efficiently solved numerically using a variant of the Bartels-Stewart algorithm [64].

Before concluding, we clarify an important subtlety. Our analysis assumes that the probe state is fixed. However, there are multiparameter sensing applications that permit full control over the probe states used. In this case, it is possible to extend our formalism to determine the optimal probe state. To see how, note that the HCRB is a biconvex function of the probe state ρ and the observable X. We have already determined the optimal observable corresponding to a chosen state: $C_{\mathsf{H},\rho}(X) = C_\mathsf{H}(\rho, X)$. Conversely, fixing X amounts to solving a convex problem in ρ to determine the optimal state corresponding to the choice in X: $C_{\mathsf{H},X}(\rho) = C_\mathsf{H}(\rho, X)$. Based on this relation, we can implement an efficient iterative biconvex program that alternatively updates the state and optimal observables by fixing one and solving the corresponding convex optimization problem [65].

IV. APPLICATIONS

Quantum metrology has applications in both spin and bosonic systems. We demonstrate the broad applicability of our results by showing how our bounds work in each of the two settings. First, for spin systems, one natural problem to consider is that of estimating the different components of a magnetic field. When the total magnetic field is known, there are only two independent components of a magnetic field to estimate, and such a problem can be tackled directly using our analytical approach for two-parameter estimation. In particular, our simple approach using the algorithm in Fig. 1 already gives interesting insights into the problem of quantum magnetometry on various types of noisy probe states.

Second, for bosonic systems, a key obstacle in determining the ultimate precision limits on parameter estimation is the infinite dimension of such systems. We show that, using our analytical approach, this obstacle can be overcome. Specifically, we use our tight analytical bounds (Fig. 4) to calculate the precision bounds on estimating the incompatible components of a logical Bloch vector of a pure bosonic code state when mixed with a thermal state.

A. Magnetometry

We use our simple two-parameter bounds to consider magnetic field sensing, which has important technological applications in navigation, position tracking, and imaging [66]. 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 [67]. Estimating each component individually allows us to attain the quantum limit [23], and this result is demonstrated in several studies [68,69]. However, in many practical applications, knowledge of multiple parameters is required simultaneously, and we must consider joint estimation strategies.

The three parameters of interest $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)^\top$ appear in the single spin Hamiltonian $\hat{H}_j(\boldsymbol{\theta}) = \boldsymbol{\theta} \cdot \mathbf{S}_j$, where $\mathbf{S}_j$ is the spin operator for the jth spin. Local depolarizing noise, described by the single spin completely positive trace-preserving map

$$\mathcal{D}_g[\rho] = (1 - g)\rho + g\frac{\mathbb{1}_2}{2}, \tag{42}$$

provides a general description for a noisy environment, where g denotes the depolarization magnitude and takes values between 0 and 1. The parameters $\boldsymbol{\theta}$ are imprinted on the probe state via the unitary evolution $\hat{U} = \exp[-i\hat{H}_j(\boldsymbol{\theta})]$. 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 prioritize each parameter into a weighted scalar mean square error. We consider three families of n-spin probe states, namely, the traditional Greenberger-Horne-Zeilinger (GHZ) states for single-parameter estimation, the modified 3D GHZ states introduced by Baumgratz and Datta [13], and the GNU states introduced by Ouyang in the context of quantum error correction [70].

First, the 3D GHZ state can be written as

$$|\psi_n^{\mathrm{3D\,GHZ}}\rangle = \frac{1}{\mathcal{N}}\sum_{j=1}^{3}(|\phi_j^+\rangle^{\otimes n} + |\phi_j^-\rangle^{\otimes n}), \tag{43}$$

where n is the total number of spins, $\mathcal{N}$ is the normalization constant of the state, and $|\phi_j^\pm\rangle$ are the eigenvectors corresponding to the ± 1 eigenvectors of the jth spin matrix. The evolved state then becomes
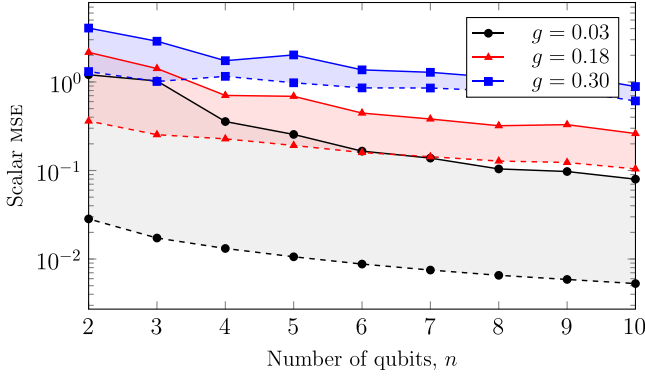
FIG. 5. Estimation precision of two directional components of a magnetic field B_x and B_y with an increasing number of spins in a depolarizing environment parameterized by g, using modified 3D GHZ states. The corresponding Hamiltonian is $\sum_{j=1}^{n} \hat{H}_j(\boldsymbol{\theta})$ and has no interaction terms. With the identity weight matrix, the dashed lower lines illustrate the QCRB and the solid lines the upper bound $\mathscr{U}$ to the HCRB as given by Theorem 1.

$$\rho(\boldsymbol{\theta}) = \hat{U}(\boldsymbol{\theta})\mathcal{D}_g^{\otimes n}[|\psi_n^{\text{3D GHZ}}\rangle\langle\psi_n^{\text{3D GHZ}}|]\hat{U}(\boldsymbol{\theta})^{\dagger}. \quad (44)$$

We illustrate how the upper bound to the HCRB and the QCRB changes with the number of probe qubits n for different depolarizing channel strengths g in Fig. 5. We observe that the upper bound to the HCRB is indeed tighter than the QCRB. Both variance bounds increase with an increasing depolarizing probability of the depolarizing channel, as expected. The 3D GHZ state attains the Heisenberg precision scaling for the noiseless case.

Second, we consider the class of GNU states that are robust to a constant amount of erasure and dephasing [71]:

$$|\varphi_1\rangle = \frac{1}{2}\sum_{j=0}^{2}\sqrt{\binom{2}{j}}|D_{Gj}^n\rangle. \quad (45)$$

Here, for every $w = 0, \ldots, n$, the Dicke state $|D_w^n\rangle$ 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 GNU states for which n is even. These are shown in Fig. 6 and compared with traditional GHZ states and 3D GHZ states. The traditional GHZ states give a worse estimation for a larger qubit number at a constant depolarization rate, as is well known. The GNU 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 HCRB 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. 7. We plot the real
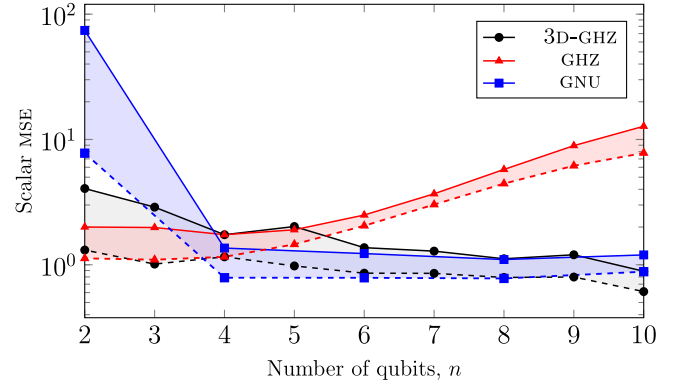

FIG. 6. Using Theorem 1, we depict two-sided bounds on the HCRB for estimating the elements of a magnetic field in a depolarizing channel. All plots with depolarizing strength $g = 0.3$. This plot compares the performance of 3D GHZ states with n-qubit GHZ states and the permutation-invariant GNU states. We observe an interesting crossover point between the tightest bound generated by the n-qubit GHZ states and 3D GHZ states, with an increasing number of qubits. The GNU states are defined over the even number of qubit numbers generates the lowest Holevo bound for a small number of qubits.

and imaginary parts of the matrices X_1 and X_2, and the Hermiticity of the observables is clearly observed.

B. Bosonic quantum codes

Our formalism for the HCRB allows performance characterization of fault-tolerant quantum codes in the
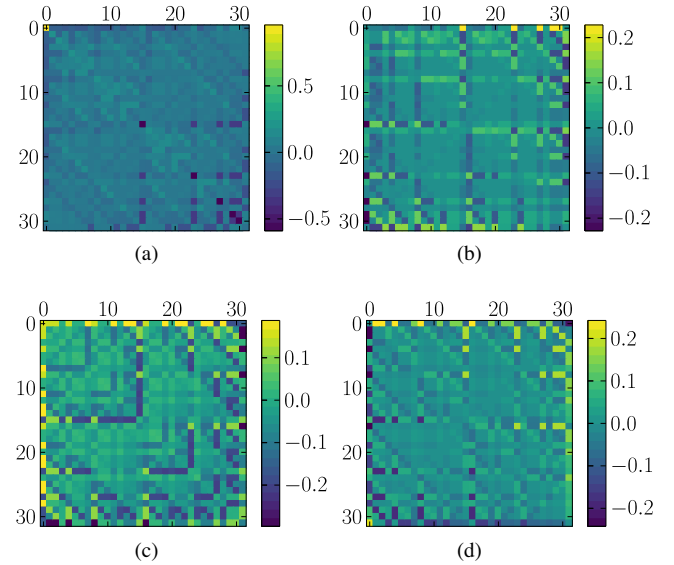

FIG. 7. Using Algorithm 1, we obtain a heat map for the measurement observables X_1 and X_2 for the depolarized five-qubit 3D GHZ state, under a depolarizing strength $g = 0.3$. We plot the real and imaginary parts separately. The Hermiticity of these observables is clearly illustrated. (a) $\text{Re}(X_1)$ (b) $\text{Im}(X_1)$ (c) $\text{Re}(X_2)$ (d) $\text{Im}(X_2)$.

context of quantum metrology. In this section, we apply our master algorithm in Fig. 4 to explore how bosonic error-correcting codes can improve characterization of logically encoded states in the presence of noise.

Susceptible quantum information can be safeguarded from decoherence by storing it in quantum error-correcting codes (QECCs), which in the case of continuous variable (bosonic) quantum systems are subspaces of infinite-dimensional Hilbert space. The working principle of QECCs is to project states with errors with high probability onto correctible subspaces labeled by the error syndromes and dynamically evolve the projected state back to the original code space. When these codes are well chosen, they can correct against errors that are introduced in physically realistic noise models. While bosonic codes on multiple modes that correct against displacement errors [72,73] and photon loss [74–77] exist, a key attraction of bosonic codes is that they can be used even on a single mode. For example, to protect codes against photon loss and phase errors on a single mode, one can use codes gapped in the Fock basis [78,79] or a single-mode Gottesman-Kitaev-Preskill code for displacement errors [72]. For a complete exposition of fault-tolerant quantum computing and error-correcting codes, the reader is directed to Refs. [80–82].

We focus on bosonic codes, which, unlike two-level systems, have infinite energy levels per mode. We define a logically encoded state that is parameterized by the coordinates $\boldsymbol{\theta} = (x, \phi)^\top$ with $x = \cos(\theta/2)$. We evaluate the HCRB for this bivariate estimation scheme by using Theorem 3 to evaluate upper and lower bounds to the HCRB for fixed u. We then tune the parameters of the binomial codes to effect improvements to estimates of $\boldsymbol{\theta}$. We consider binomial codes that protect code words against number-shift and phase errors. In particular, we analyze the ultimate limits of estimating the complex coefficients of a pure binomial code state in the presence of thermal noise.

The logical code words for the binomial code are supported on a bounded number of Fock states through

$$|0_L\rangle = \sum_{\substack{j \geq 0 \\ j \text{ even}}} 2^{-[(n-1)/2]} \sqrt{\binom{n}{j}} |Gj\rangle, \tag{46}$$

$$|1_L\rangle = \sum_{\substack{j \geq 0 \\ j \text{ odd}}} 2^{-[(n-1)/2]} \sqrt{\binom{n}{j}} |Gj\rangle, \tag{47}$$

where $G, n \in \mathbb{R}$ are related to the number of correctable number-shift and phase errors, respectively [83]. From Eq (7) in Ref. [78], one requires $G \geq G_{\text{bin}} + L_{\text{bin}} + 1$ to correct G_{bin} gain and L_{bin} loss errors and $n - 1 \geq \max\{L_{\text{bin}}, G_{\text{bin}}, 2D_{\text{bin}}\}$ to correct D_{bin} phase errors. For fixed G and n, we construct the logical state $\rho_L = |\psi_L\rangle\langle\psi_L|$ with

$$|\psi_L\rangle = x|0_L\rangle + \sqrt{1 - x^2}e^{i\phi}|1_L\rangle, \tag{48}$$

where $x \in [-1, 1]$ and $\phi \in \mathbb{R}$. In the noisy scheme, we thermalize this logical pure state through $\rho = \lambda_{\text{th}}\rho_{\text{th}} + (1 - \lambda_{\text{th}})\rho_L$, where

$$\rho_{\text{th}} = \frac{1}{1 - e^{-\beta}} \sum_{k=0}^{\infty} e^{-k\beta}|k\rangle\langle k| \tag{49}$$

is a thermal state with temperature β. Since $\partial_j\rho = (1 - \lambda_{\text{th}})\partial_j\rho_L$, $j = \{x, \phi\}$, and ρ_L is supported only on the Fock states $|0\rangle, |G\rangle, ..., |Gn\rangle$, the state derivatives $\partial_j\rho$ are supported only on the Fock states $|0\rangle, |G\rangle, ..., |Gn\rangle$. Using this property, we can determine that, in the calculation of the HCRB, we need to consider only the evaluation of ρ on the support of the Fock states $|0\rangle, |G\rangle, ..., |Gn\rangle$. Denoting such a state as τ, we can write

$$\tau = \sum_{j,k=0}^{n} |j\rangle\langle k|\langle Gj|\rho|Gk\rangle \tag{50}$$

and observe that it has spectral decomposition

$$\tau = \sum_{k=0}^{n} t_k|\tau_k\rangle\langle\tau_k| \quad \text{with} \quad |\tau_k\rangle = \sum_{j=0}^{n} \tau_{k,j}|j\rangle. \tag{51}$$

The key implication is that τ is now an effective size $(n + 1)$ matrix and, unlike ρ, does not have infinite dimensions. Now define

$$|T_k\rangle = \sum_{j=0}^{n} \tau_{k,j}|Gj\rangle. \tag{52}$$

From the form of our noise model, τ is a full-rank matrix, because it is a convex combination of a positive definite matrix and a positive semidefinite matrix. The positive definite matrix arises from a truncation of the thermal state on the Fock states $|0\rangle, |G\rangle, ..., |Gn\rangle$, and the positive semidefinite matrix arises from ρ_L. Since τ is a full-rank matrix, it follows that the spectral decomposition of ρ is

$$\rho = \sum_{k=0}^{n} t_k|T_k\rangle\langle T_k| + \lambda_{\text{th}} \sum_{k=0}^{n-1}\sum_{j=1}^{g-1} \frac{e^{-\beta(Gk+j)}}{1 - e^{-\beta}}|Gk+j\rangle\langle Gk+j|$$
$$+ \lambda_{\text{th}} \sum_{k=Gn+1}^{\infty} \frac{e^{-\beta k}}{1 - e^{-\beta}}|k\rangle\langle k|. \tag{53}$$

From the above spectral decomposition of ρ, it is clear that only the first summation term contributes to the state derivatives. This contribution makes the effective dimension of the problem equal to the dimension of τ instead of that of ρ. Because of this reduction in the effective dimensionality of the problem, we can efficiently use
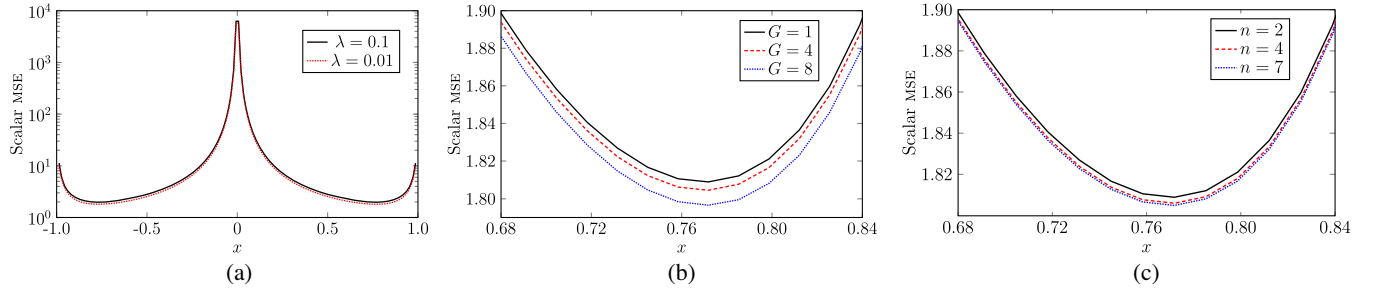
FIG. 8. Variation of HCRB with thermalization and binomial code parameters. For the plot in (a), we take $n = 2$, $\beta = 0.01$, and $G = 1$. In (b), we see the effect of noise by varying the number of correctable amplitude damping errors of the binomial code, using $n = 2$, $\lambda_{\text{th}} = 0.01$, and $\beta = 1$. In (c), we illustrate the effect of increasing the correctable phase errors of the binomial code on the HCRB using $\lambda_{\text{th}} = 0.01$, $\beta = 1$, and $G = 1$. (a) Thermalisation effect. (b) Amplitude damping error parameter, G. (c) Phase error parameter,n.

Algorithm 4 to evaluate upper and lower bounds to the HCRB for fixed u to benchmark parameter estimates for $\boldsymbol{\theta}$.

By optimizing over u, we find exact values of the HCRB. In Fig. 8, we illustrate how the HCRB changes with different values of the noise parameters λ_{th} and β and the code parameters G and n. Figure 8(a) shows that increasing the thermalization increases the scalar MSE. We also note that the value for x that minimizes the MSE is insensitive to the amount of thermalization in the state. Notice also that the effect of binomial codes in this application is limited at high thermalization, where the state has a lot of thermal noise. The binomial codes are able to protect against errors resulting on the state due to low temperatures. To see this protection, we observe the behavior of the HCRB in the region where it is minimized with respect to x. In Fig. 8(b), we illustrate the behavior of the HCRB within $x = [0.68, 0.84]$ for different values of G. Notice that increasing the number of correctable amplitude damping errors by increasing the value of G improves the precision of the simultaneous estimate for x and ϕ. In Fig. 8(c), we illustrate a similar improvement by increasing n, which is related to the number of correctable phase errors. This result demonstrates that the error-correcting codes can be used to improve simultaneous parameter estimates in the low error regime.

It is worth noting the performance of the simple bounds in Theorem 1 for this application. For the specific choice of parameters $u = 0$, $x = -0.8$, $\phi = 0.7$, $n = 5$, $G = 3$, $\lambda_{\text{th}} = 0.2$, and $\beta = 0.1$, we get the exact HCRB. In general, Theorem 1 returns nontight bounds to the HCRB for alternative values in the parameter space. This result illustrates that if the tightness of bounds is crucial, then one should apply Theorem 3.

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 noncompatible observables. Specifically, the optimal strategy to define the fundamental limits to precision estimates and their attainability is not known. Efforts to estimate multiple noncompatible observables have largely been focused on approaching the fundamental QCRB. This focus has led to efforts to devise nontrivial measurement schemes that approach the QCRB. An alternative approach is to focus on the tighter HCRB, which is physically attainable. However, the HCRB is difficult to evaluate, since it involves a difficult optimization over two observables. This difficulty 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 a larger number of parameters. In the two-parameter case, we reduce the complexity of the optimization 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 POVMs. Our results readily apply to a large range of physical applications. This range will provide deeper insight into the role of quantum measurements in quantum sensing and help continue the drive of realizing quantum technologies.

We illustrate an application of our results by considering the estimation of a magnetic field using noisy multiqubit probe states [71,84]. A recent numerical study by Albarelli, Friel, and Datta demonstrates the necessity of using the HCRB over the QCRB, based on a violation of the weak commutation condition [44]. 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.

A second application of our results explores how bosonic quantum error correction codes can improve noise resilience of parameter estimates. Bosonic codes are interesting because of their potential in reducing the number of physical systems required while having some robustness against errors. However, the infinite dimensionality of

bosonic systems renders a brute force numerical approach to determining the HCRB intractable. Instead, through our analytical approach, we reduce this problem to a finite-dimensional problem and, thereby, evaluate the corresponding precision bounds efficiently.

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 insight 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 extension would provide an immediate access to the tighter HCRB through experimental implementation.

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

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

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 continue to be an active area of research [85], and they have also been recently used in quantum information theory [86].

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

Lemma 4.—Let $Y, H \in L(\mathbb{C}^n)$. Then,

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

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

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

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

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 optimization problems is.

Lemma 5.—Let $A \in L(\mathbb{C}^n)$ and let ρ be a full-rank matrix in $\mathbb{H}_n$. Then,

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

$$(A6)$$

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

$$(A7)$$

with the minimum achieved by setting $Y = -A^\dagger \rho^{-1}$ and $Y = -\rho^{-1} A^\dagger$, respectively.

Proof.—We first prove Eqs. (A6) and (A7). The corresponding objective functions that are to be minimized are convex and differentiable, so it suffices to find when their Fréchet derivatives are equal to zero for any direction H. For this task, we use Lemma 4, 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 ρ 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 result back into the objective functions gives the result. ∎

APPENDIX B: SIMPLE TWO-PARAMETER BOUNDS TO THE HCRB

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 optimization problem

$$\underset{X_1,X_2}{\text{minimize}} \; \text{Tr}[W\text{Re}Z] + \|W\text{Im}Z\|_1,$$

$$\text{subject to} \quad \text{Tr}[\rho X_j] = 0,$$

$$\text{Tr}[\partial_j\rho X_k] = \delta_{jk}, \tag{B1}$$

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

$$Z = \begin{pmatrix} \text{Tr}[\rho X_1^2] & \text{Tr}[\rho X_1 X_2] \\ \text{Tr}[\rho X_2 X_1] & \text{Tr}[\rho X_2^2] \end{pmatrix}. \tag{B2}$$

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

1. Reformulation of the optimization problem

The optimization problem (10) can be solved analytically primarily from our ability to rewrite the objective function as a quadratic function in the optimization 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 optimization in finance [87]. A similar argument allows us to solve Eq. (10) 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 $X(\cdot)X^\dagger$ is a completely positive map. Second, the positivity of the diagonal entries of Z implies that

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

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

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

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

$$\text{ReTr}[Z] + iw = \text{Tr}[(X_1 + iX_2)\rho(X_1 + iX_2)^\dagger], \tag{B3}$$

$$\text{ReTr}[Z] - iw = \text{Tr}[(X_1 - iX_2)\rho(X_1 - iX_2)^\dagger]. \tag{B4}$$

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

$$\text{Tr}[\rho Y] = 0,$$

$$\text{Tr}[\partial_1\rho Y] = 1,$$

$$\text{Tr}[\partial_2\rho Y] = i. \tag{B5}$$

Hence, the optimization problem (10) can be written as

$$\underset{Y,t}{\text{minimize}} \; t,$$

$$\text{subject to} \; \text{Tr}[Y\rho Y^\dagger] \le t,$$

$$\text{Tr}[Y^\dagger \rho Y] \le t,$$

$$\text{Tr}[\rho Y] = 0,$$

$$\text{Tr}[\partial_1\rho Y] = 1,$$

$$\text{Tr}[\partial_2\rho Y] = i. \tag{B6}$$

Note that the optimization problem (B6) is a linear optimization problem with convex quadratic and linear constraints. When the equality constraints are satisfied, the quadratic terms in the inequality constraints are nonnegative, and, by setting t to be arbitrarily large, we can see that the inequality constraints in Eq. (B6) can always be strictly satisfied. Since Eq. (B6) is also a convex optimization problem because of its linear objective function and convex constraint functions, the Slater constraint qualification holds with respect to Eq. (B6). This result implies that the first-order KKT condition suffices to determine the optimality of Eq. (B6).

2. Analyzing the Lagrangian

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

$$\mathbf{z} = (z_1, z_2, z_3, z_4, z_5, z_6). \tag{B7}$$

The Lagrangian of Eq. (B6) is

$$
\begin{aligned}
\mathsf{L}(Y, t, u, v, \mathbf{z}) = {} & t + u\text{Tr}[Y\rho Y^\dagger] - ut + v\text{Tr}[Y^\dagger\rho Y] - vt \\
& + z_1\text{ReTr}[\rho Y] + z_4\text{ImTr}[\rho Y] \\
& + z_2(\text{ReTr}[\partial_1\rho Y] - 1) + z_5\text{ImTr}[\partial_1\rho Y] \\
& + z_3\text{ReTr}[\partial_2\rho Y] + z_6(\text{ImTr}[\partial_2\rho Y] - 1),
\end{aligned}
\tag{B8}
$$

where u and 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 [88] and their corresponding Lagrange multipliers is always zero. Third is the feasibility of the primal

variables, and fourth is feasibility of the dual variables. If these KKT optimality conditions hold, then we can obtain the optimal solution and value of the corresponding optimization problem.

Now we use the fact that

$$\text{ReTr}[\rho Y] = \frac{\text{Tr}[\rho Y] + \text{Tr}[\rho Y^\dagger]}{2}, \tag{B9}$$

$$\text{ImTr}[\rho Y] = \frac{\text{Tr}[\rho Y] - \text{Tr}[\rho Y^\dagger]}{2i}, \tag{B10}$$

$$\text{ReTr}[\partial_j \rho Y] = \frac{\text{Tr}[\partial_j \rho Y] + \text{Tr}[\partial_j \rho Y^\dagger]}{2}, \tag{B11}$$

$$\text{ImTr}[\partial_j \rho Y] = \frac{\text{Tr}[\partial_j \rho Y] - \text{Tr}[\partial_j \rho Y^\dagger]}{2i}. \tag{B12}$$

Using this fact, it follows that

$$\begin{aligned}\mathsf{L}(Y, t, u, v, \mathbf{z}) = {}& t(1 - u - v) - \mathbf{b}^\top \mathbf{z} + u\text{Tr}[Y\rho Y^\dagger] \\ & + v\text{Tr}[Y^\dagger \rho Y] + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger],\end{aligned} \tag{B13}$$

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

$$A = z_1 A_1 + z_2 A_2 + z_3 A_3 + z_4 A_4 + z_5 A_5 + z_6 A_6, \tag{B14}$$

where

$$A_1 = \frac{\rho}{2}, \qquad A_2 = \frac{\partial_1 \rho}{2}, \qquad A_3 = \frac{\partial_2 \rho}{2}, \tag{B15}$$

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 ρ. From the positive definiteness of ρ, t is equal to zero if and only if Y is 0, but this result would violate the feasibility constraints. Hence, t cannot be equal to zero.

(2) The stationarity KKT condition requires that the derivative of Lagrangian in Eq. (B8) be zero with respect to t. From this requirement, we observe that the optimal dual variables must satisfy $u + v = 1$.

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

$$\begin{aligned} u(\text{Tr}[Y\rho Y^\dagger] - t) &= 0, \\ v(\text{Tr}[Y^\dagger \rho Y] - t) &= 0. \end{aligned} \tag{B16}$$

If $\text{Tr}[Y\rho Y^\dagger] \neq \text{Tr}[Y^\dagger \rho Y]$, 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 situation corresponds to the scenario where the QCRB 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.

However, if $\text{Tr}[Y\rho Y^\dagger] = \text{Tr}[Y^\dagger \rho Y]$, then the *Ansätze* $(u, v) = (0, 1)$ and $(u, v) = (0, 1)$ will not yield tight bounds, because complementary slackness will not further constrain the optimal values of u and v.

3. Deriving the Lagrange dual functions

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

$$\mathsf{L}(Y, t, 1, 0, \mathbf{z}) = -\mathbf{b}^\top \mathbf{z} + \text{Tr}[Y\rho Y^\dagger] + \text{Tr}[AY] + \text{Tr}[A^\dagger Y^\dagger], \tag{B17}$$

where $\mathbf{b} = (0, 1, 0, 0, 0, 1)^\top$. Since ρ is full rank, ρ is invertible. Using Lemma 5, the above is minimized with respect to Y when $Y = -A^\dagger \rho^{-1}$ with optimal value $-\text{Tr}[A^\dagger \rho^{-1} A]$. In this scenario, the Lagrange dual function of Eq. (B6) evaluated with $(u, v) = (1, 0)$ is

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

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

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

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

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

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

$$\begin{aligned} Q_1 &= \sum_{j,k=1,\ldots,6} \text{Tr}[A_j^\dagger \rho^{-1} A_k] |j\rangle\langle k|, \\ Q_2 &= \sum_{j,k=1,\ldots,6} \text{Tr}[A_j \rho^{-1} A_k^\dagger] |j\rangle\langle k|. \end{aligned} \tag{B21}$$

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

$$g(1, 0, \mathbf{z}) = -\mathbf{z}^\top Q_1 \mathbf{z} - \mathbf{b}^\top \mathbf{z},$$
$$g(0, 1, \mathbf{z}) = -\mathbf{z}^\top Q_2 \mathbf{z} - \mathbf{b}^\top \mathbf{z}. \qquad (B22)$$

4. Upper and lower bounds

Using the fact that the Lagrange dual functions (B22) and the dual variables are real, lower bounds to the HCRB are given by

$$\max_{j=1,2} \max_{\mathbf{z} \in \mathbb{R}^6} [-\mathbf{z}^\top \mathrm{Re}(Q_j) \mathbf{z} - \mathbf{b}^\top \mathbf{z}], \qquad (B23)$$

from which it follows that

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

is the correct optimality condition to consider. Thus, when $\mathrm{Re}(Q_j)$ is full rank, the lower bounds to the HCRB can be written as

$$\max_{j=1,2} l_j, \quad \text{where } l_j = \frac{1}{4}\mathbf{b}^\top \mathrm{Re}(Q_j)^{-1}\mathbf{b}. \qquad (B24)$$

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

To obtain upper bounds to the HCRB, we appeal to the form of the primal problem is closely related to Eq. (B6), that has an objective function of

$$\max\{\mathrm{Tr}[Y\rho Y^\dagger], \mathrm{Tr}[Y^\dagger \rho Y]\}. \qquad (B25)$$

The upper bounds are expressed in terms of the dual variables that optimize Eq. (B23), which we can write as $\mathbf{z}_1 = (z_{1,1}, ..., z_{6,1})$ and $\mathbf{z}_2 = (z_{1,2}, ..., z_{6,2})$ where

$$z_{a,j} = -\frac{1}{2}\{[\mathrm{Re}(Q_j)^{-1}]_{a2} + [\mathrm{Re}(Q_j)^{-1}]_{a6}\} \qquad (B26)$$

and $[\mathrm{Re}(Q_j)^{-1}]_{ab}$ denotes the matrix element in the ath row and bth column of the inverse of $\mathrm{Re}(Q_j)$.

Recall that, when $(u, v) = (1, 0)$, the optimal solution to Y in minimizing the Lagrangian is $-A^\dagger \rho^{-1}$. By choosing $A = A_1 z_{1,1} + \cdots + A_6 z_{6,1}$, we find that the HCRB is thus upper bounded by

$$P_1 = \max\{\mathrm{Tr}[A^\dagger \rho^{-1}A], \mathrm{Tr}[\rho^{-2}A\rho A^\dagger]\}. \qquad (B27)$$

Then,

$$P_1 = \max\{l_1, m_1\}, \qquad (B28)$$

where

$$m_j = \sum_{a,b=1}^{6} \mathrm{Tr}[\rho^{-2}A_a \rho A_b^\dagger] z_{a,j} z_{b,j}. \qquad (B29)$$

When $(u, v) = (0, 1)$, the optimal solution to Y in minimizing the Lagrangian is $-\rho^{-1}A^\dagger$. In this case, the primal objective function is equal to

$$P_2 = \max\{l_2, m_2\}. \qquad (B30)$$

Hence, the HCRB is at most

$$\min_{j=1,2}\{\max\{l_j, m_j\}\}. \qquad (B31)$$

This concludes the proof of Theorem 1 for bounds on the two-parameter HCRB.

5. Full rankness of Q

The analytic solution to the HCRB requires $\mathrm{Re}(Q_j)$ to have full rank such that the solution can be determined. In this subsection, we demonstrate that the full rankness of the probe state ρ 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

$$Q_1 = \begin{pmatrix} H & -iH \\ iH & H \end{pmatrix}, \qquad Q_2 = \begin{pmatrix} H & iH \\ -iH & H \end{pmatrix}, \qquad (B32)$$

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

$$B_1 = \rho^{-1/2}A_1 = \rho^{-1/2}\rho/2 = \rho^{1/2}/2, \qquad (B33)$$

$$B_2 = \rho^{-1/2}A_2 = \rho^{-1/2}\delta_1\rho/2, \qquad (B34)$$

$$B_3 = \rho^{-1/2}A_3 = \rho^{-1/2}\delta_2\rho/2. \qquad (B35)$$

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

$$H_{i,j} = \langle B_i, B_j \rangle. \qquad (B36)$$

As a Gram matrix, it is positive semidefinite. 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 a one-parameter problem and not a two-parameter problem. Hence, $\{A_1, A_2, A_3\}$ are linearly independent. If we assume ρ 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 ρ entails full rankness of H.

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

$$\mathrm{Re}(Q_2) = \begin{pmatrix} \mathrm{Re}(H) & -\mathrm{Im}(H) \\ \mathrm{Im}(H) & \mathrm{Re}(H) \end{pmatrix}. \quad (B37)$$

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

$$\mathrm{Re}(Q_2) \to \begin{pmatrix} H & iH \\ \mathrm{Im}(H) & \mathrm{Re}(H) \end{pmatrix} \to \begin{pmatrix} H & 0 \\ \mathrm{Im}(H) & H^* \end{pmatrix}, \quad (B38)$$

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

APPENDIX C: LOWER BOUND IN THE MULTIPARAMETER SETTING

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

$$\underset{X_1,\dots,X_d}{\mathrm{minimize}} \ \mathrm{Tr}[\mathrm{Re}Z] + \|\mathrm{Im}Z\|_1,$$
$$\mathrm{subject\ to} \quad \mathrm{Tr}[\rho X_j] = 0$$
$$\mathrm{Tr}[\partial_j \rho X_k] = \delta_{jk}. \quad (C1)$$

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

$$w_{j,k} = \mathrm{Tr}[\rho X_j X_k] - \mathrm{Tr}[\rho X_k X_j]. \quad (C2)$$

1. Deriving a lower bound for the objective function

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

$$\mathrm{Tr}[\mathrm{Re}Z] = \sum_{j=1}^{d} \mathrm{Tr}[X_j \rho X_j^\dagger], \quad (C3)$$

$$\mathrm{Im}Z = \frac{1}{2i} \sum_{1 \le j < k \le d} w_{j,k}(|j\rangle\langle k| - |k\rangle\langle j|). \quad (C4)$$

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

$$\mathrm{Im}Z = \frac{1}{2i} \begin{pmatrix} 0 & w_{1,2} & w_{1,3} \\ -w_{1,2} & 0 & w_{2,3} \\ -w_{1,3} & -w_{2,3} & 0 \end{pmatrix}. \quad (C5)$$

Whenever $d \ge 3$, the trace norm of $\mathrm{Im}Z$ fails to be a quadratic form in the observables $X_1, \dots, X_d$. Hence, the objective function of the optimization problem (C1) fails to be quadratic for us to apply the techniques in Appendix B 1. We can, however, obtain lower bounds for the trace norm of $\mathrm{Im}Z$ that do have a quadratic structure, namely, by exploiting the following decomposition for the trace norm:

$$\|\mathrm{Im}Z\|_1 = \max\{\mathrm{Tr}[U\mathrm{Im}Z] : U \text{ is a unitary matrix}\}. \quad (C6)$$

Fortunately, it is possible to pick unitary matrices U such that $\mathrm{Tr}[U\mathrm{Im}Z]$ are quadratic in the observables $X_1, \dots, X_d$, which we prove in the following subsection. We achieve this possibility by constructing unitary matrices labeled by binary vectors $\alpha = (\alpha_1, \dots, \alpha_d)$ given by

$$U_\alpha = |d\rangle\langle 1|(-1)^{\alpha_d} + \sum_{j=1}^{d-1} |j\rangle\langle j+1|(-1)^{\alpha_j}. \quad (C7)$$

Using the unitary matrices in

$$\mathscr{U} = \{U_\alpha : \alpha \in \{0,1\}^d\}, \quad (C8)$$

we obtain the lower bound

$$\|\mathrm{Im}Z\|_1 \ge \max\{\mathrm{Tr}[U\mathrm{Im}Z] : U \in \mathscr{U}\}. \quad (C9)$$

2. Recasting the optimization problem

In this subsection, we prove the following lemma.

Lemma 6.—Let d be a positive integer, where $d \ge 2$. Now, given a binary vector $\alpha = (\alpha_1, \dots, \alpha_d)$, let U_α be as defined in Eq. (C7), and let us define

$$V_\alpha = \frac{1}{2}\mathrm{Tr}[[X_d + (-1)^{\alpha_1} iX_1]\rho[X_d + (-1)^{\alpha_1} iX_1]^\dagger]$$
$$+ \frac{1}{2}\sum_{j=1}^{d-1} \mathrm{Tr}[[X_j + (-1)^{\alpha_j} iX_{j+1}]\rho[X_j + (-1)^{\alpha_j} iX_{j+1}]^\dagger], \quad (C10)$$

where $X_1, \dots, X_d$ are Hermitian matrices and ρ is a density matrix. Let $Z = \sum_{j,k=1}^{d} \mathrm{Tr}[\rho X_j X_k]|j\rangle\langle k|$. Then,

$$\mathrm{Tr}[\mathrm{Re}Z] + \mathrm{Tr}[U_\alpha \mathrm{Im}Z] = V_\alpha. \quad (C11)$$

Proof.—Rewriting Eq. (C4), we get

$$\mathrm{Im}Z = \frac{1}{2i}\sum_{j,k=1}^{d} w_{j,k}|j\rangle\langle k|, \quad (C12)$$

where $w_{j,k}$ is as defined in Eq. (C2). Using Eqs. (C3) and (C12), we can find that

$$\text{Tr}[\text{Re}Z] + \text{Tr}[U_\alpha \text{Im}Z]$$
$$= \sum_{j=1}^{d} \text{Tr}[X_j \rho X_j^\dagger] + \frac{1}{2i} \sum_{j=1}^{d} \sum_{k=1}^{d} \text{Tr}[w_{j,k} U_\alpha |j\rangle\langle k|]. \quad (C13)$$

Now,

$$\text{Tr}[U_\alpha |j\rangle\langle k|]$$
$$= \text{Tr}[(|d\rangle\langle 1|)|j\rangle\langle k|](-1)^{\alpha_d}$$
$$+ \sum_{a=1}^{d-1} \text{Tr}[(|a\rangle\langle a+1|)|j\rangle\langle k|](-1)^{\alpha_a}$$
$$= \delta_{j,1}\delta_{k,d}(-1)^{\alpha_d} + \sum_{a=1}^{d-1} \delta_{j,a+1}\delta_{k,a}(-1)^{\alpha_a}. \quad (C14)$$

Hence,

$$\sum_{j,k=1}^{d} w_{j,k}\text{Tr}[U_\alpha |j\rangle\langle k|] = w_{1,d}(-1)^{\alpha_d} + \sum_{a=1}^{d-1} w_{a+1,a}(-1)^{\alpha_a}. \quad (C15)$$

Now, note that, for any $j,k = 1, ..., d$, we have

$$\frac{w_{j,k}}{2i} = \frac{-i}{2}(\text{Tr}[X_k \rho X_j] - \text{Tr}[X_j \rho X_k])$$
$$= \frac{1}{2}\{\text{Tr}[X_k \rho(iX_j)^\dagger] + \text{Tr}[(iX_j)\rho X_k^\dagger]\}. \quad (C16)$$

Hence, we get

$$\frac{1}{2i}\sum_{j=1}^{d}\sum_{k=1}^{d}\text{Tr}[w_{j,k} U_\alpha |j\rangle\langle k|]$$
$$= \frac{1}{2}\{\text{Tr}[X_d \rho(iX_1)^\dagger] + \text{Tr}[(iX_1)\rho X_d^\dagger]\}(-1)^{\alpha_d}$$
$$+ \sum_{k=1}^{d-1}\frac{1}{2}\{\text{Tr}[X_k \rho(iX_{k+1})^\dagger] + \text{Tr}[(iX_{k+1})^\dagger \rho X_k^\dagger]\}(-1)^{\alpha_k}. \quad (C17)$$

Now,

$$\frac{1}{2}\text{Tr}[[X_d + (-1)^{\alpha_d}iX_1]\rho[X_d + (-1)^{\alpha_d}iX_1]^\dagger]$$
$$= \frac{1}{2}\text{Tr}[X_d \rho X_d^\dagger] + \frac{1}{2}\text{Tr}[X_1 \rho X_1^\dagger]$$
$$+ \frac{1}{2}\text{Tr}[X_d \rho[(-1)^{\alpha_d}iX_1]^\dagger] + \frac{1}{2}\text{Tr}[[(-1)^{\alpha_d}iX_1]\rho X_d^\dagger], \quad (C18)$$

and, similarly, for all $a = 1, ..., d-1$, we have

$$\frac{1}{2}\text{Tr}[[X_a + (-1)^{\alpha_a}iX_{a+1}]\rho[X_a + (-1)^{\alpha_a}iX_{a+1}]^\dagger]$$
$$= \frac{1}{2}\text{Tr}[X_a \rho X_a^\dagger] + \frac{1}{2}\text{Tr}[X_{a+1}\rho X_{a+1}^\dagger]$$
$$+ \frac{1}{2}\text{Tr}[X_a \rho[(-1)^{\alpha_a}iX_{a+1}]^\dagger] + \frac{1}{2}\text{Tr}[[(-1)^{\alpha_a}iX_{a+1}]\rho X_a^\dagger]. \quad (C19)$$

From Eqs. (C13), (C17), (C18), and (C19), the lemma follows. ∎

To see how this lemma works explicitly for the three-parameter ($d = 3$) scenario, note that

$$\text{Tr}[\text{Re}Z] + \frac{i}{2}\sum_{(a,b)\in E_{d,1}} w_{a,b}$$
$$= \text{Tr}[X_1 \rho X_1^\dagger] + \text{Tr}[X_2 \rho X_2^\dagger] + \text{Tr}[X_3 \rho X_3^\dagger]$$
$$+ \frac{i}{2}(\text{Tr}[X_2 \rho X_1^\dagger] + \text{Tr}[X_3 \rho X_2^\dagger] + \text{Tr}[X_1 \rho X_3^\dagger])$$
$$- \frac{i}{2}(\text{Tr}[X_1 \rho X_2^\dagger] + \text{Tr}[X_2 \rho X_3^\dagger] + \text{Tr}[X_3 \rho X_1^\dagger])$$
$$= \frac{1}{2}\text{Tr}[(X_1 + iX_2)\rho(X_1 + iX_2)^\dagger]$$
$$+ \frac{1}{2}\text{Tr}[(X_2 + iX_3)\rho(X_2 + iX_3)^\dagger]$$
$$+ \frac{1}{2}\text{Tr}[(X_3 + iX_1)\rho(X_3 + iX_1)^\dagger]. \quad (C20)$$

Then, we can rewrite Eq. (C1) as an optimization over Hermitian matrices $X_1, ..., X_d$, where

$$\begin{aligned} &\underset{X_1,...,X_d}{\text{minimize}} \quad \underset{\alpha\in\{0,1\}^d}{\max} V_\alpha, \\ &\text{subject to} \quad \text{Tr}[\rho X_j] = 0 \\ &\qquad\qquad \text{Tr}[\partial_j \rho X_k] = \delta_{jk}. \end{aligned} \quad (C21)$$

We can rewrite with an introduction of an auxiliary variable $t \in \mathbb{R}$ so that Eq. (C21) is equivalent to

$$\begin{aligned} &\underset{X_1,...,X_d,t}{\text{minimize}} \quad t, \\ &\text{subject to} \quad \text{Tr}[\rho X_j] = 0 \\ &\qquad\qquad \text{Tr}[\partial_j \rho X_k] = \delta_{jk} \\ &\qquad\qquad V_\alpha \leq t \\ &\qquad\qquad \alpha \in \{0,1\}^d. \end{aligned} \quad (C22)$$

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

3. Diagonalizing the quadratic forms

Now, note that the lower bound V_α that we have for the objective function is a quadratic function of the optimization variables $X_1 + i(-1)^{\alpha_1} X_2, ..., X_{d-1} + i(-1)^{\alpha_{d-1}} X_d$, $X_d + i(-1)^{\alpha_d} X_1$, and these optimization variables depend on the binary vector α. We can alternatively write V_α in terms of optimization variables $Y_1, ..., Y_d$ that are independent of α. We can quantify the linear dependence of the variables $Y_1, ..., Y_d$ on the variables $X_1, ..., X_d$ using the following matrix equation:

$$\mathbf{Y} = \mathbf{SX}, \tag{C23}$$

where

$$\mathbf{S} = \begin{pmatrix} S_{1,1}\mathbf{1} & \cdots & S_{1,d}\mathbf{1} \\ \vdots & & \vdots \\ S_{d,1}\mathbf{1} & \cdots & S_{d,d}\mathbf{1} \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_d \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix}, \tag{C24}$$

and

$$S = \begin{cases} \sum_{j \in \mathbb{Z}_d} (|j\rangle\langle j| + i|j\rangle\langle j \oplus 1|) & d \neq 0 \pmod 4 \\ \sum_{j \in \mathbb{Z}_d} [|j\rangle\langle j| + (-1)^{\delta_{j,d}} i|j\rangle\langle j \oplus 1|] & \text{otherwise.} \end{cases} \tag{C25}$$

When d is not a multiple of 4, $S_{j,j} = 1$ and $S_{j,j+1} = i$ for all $j = 1, ..., d-1$, and $S_{d,d} = 1, S_{d,1} = i$, and all other matrix elements of S are zero. For instance, when $d = 3$, we have

$$S = \begin{pmatrix} 1 & i & 0 \\ 0 & 1 & i \\ i & 0 & 1 \end{pmatrix}, \tag{C26}$$

but when $d = 4$ we have

$$S = \begin{pmatrix} 1 & i & 0 & 0 \\ 0 & 1 & i & 0 \\ 0 & 0 & 1 & i \\ -i & 0 & 0 & 1 \end{pmatrix}. \tag{C27}$$

Note that the only difference when d is a multiple of 4 is that we flip the sign of the bottom-left matrix element.

Given such a set of Y_j variables, it then follows that

$$V_\alpha = \frac{1}{2}\sum_{j=1}^{d} (\delta_{0,\bar{\alpha}_j} \mathrm{Tr}[Y_j \rho Y_j^\dagger] + \delta_{1,\bar{\alpha}_j} \mathrm{Tr}[Y_j^\dagger \rho Y_j]), \tag{C28}$$

where $\bar{\alpha} = \alpha$ when d is not a multiple of 4, and when d is a multiple of 4, then $\bar{\alpha}$ differs from α by simply flipping the last bit.

In the following proposition, we determine when the matrix S is full rank.

Proposition 7.—Let d be a positive integer, and let S be a matrix as defined in Eq. (C25). Then S has full rank.

Proof.—First, we consider the case when d is not a multiple of 4. Since S is a circulant matrix, its eigenvectors are the Fourier modes $|\phi_k\rangle = \sum_{j=0}^{d-1} \omega^{jk}|j\rangle$, where $k = 0, 1, ..., d-1$ and $\omega = \exp(2\pi i/d)$ is a root of unity. The only way to get $S|\phi_k\rangle = \mathbf{0}$ is to have $\omega^k = i$ for some integer k, but this result is possible only if d divides 4. So if d does not divide 4, we cannot have $S|\phi_k\rangle = \mathbf{0}$, which implies that S does not have any zero eigenvalues. Hence, S is full rank.

Next, we consider the case when d is a multiple of 4. The first $d-1$ rows of S form an upper triangular matrix and are, therefore, linearly independent. We just need to show that the last row is linearly independent from the rest. Let us denote the jth row of S by s_j, where j goes from 1 to d. Consider an arbitrary sum of the first $d-1$ rows of the form

$$v = \sum_{j=1}^{d-1} c_j s_j. \tag{C29}$$

We wish to know if there exists a choice of constants c_j such that $v = s_d$ where our definition of S (when d is a multiple of 4) has

$$s_d = (-i, 0, ..., 0, 1). \tag{C30}$$

By setting the jth element of v equal to the jth element of s_d, we obtain an equation for each j:

$$c_1 = -i, \tag{C31}$$

$$ic_j + c_{j+1} = 0, \tag{C32}$$

$$ic_{d-1} = 1, \tag{C33}$$

with the middle line holding for all $1 \leq j < d-1$. It is simple to confirm that there does not exist a solution to this set of equations. In particular, we have the recursive equation $c_{j+1} = (-i)c_j$ with initial condition $c_1 = (-i)$, and this equation solves to $c_j = (-i)^j$. This equation entails $c_{d-1} = (-i)^{d-1} = i(-1)^d = i$ when d is a multiple of 4. However, this result contradicts $ic_{d-1} = 1$, and so no solution exists. Therefore, the last row S is linearly independent from the rest, and the matrix is full rank. ∎

From the above proposition, we see that, whenever d is not a multiple of 4, the matrix S is full rank, which implies that $T = S^{-1}$ exists. In this scenario, we can write

$$T = \begin{pmatrix} T_{1,1} & \cdots & T_{1,d} \\ \vdots & & \vdots \\ T_{d,1} & \cdots & T_{d,d} \end{pmatrix}, \qquad (C34)$$

and it follows that, for every $k = 1, \ldots, d$, we can express the Hermitian observables $X_1, \ldots, X_d$ as linear combinations of the matrices $Y_1, \ldots, Y_d$:

$$X_k = \sum_{\ell=1}^{d} T_{k,\ell} Y_\ell. \qquad (C35)$$

Recall that $\rho_j = \partial_j \rho$ for $j = 1, \ldots, d$. From Eq. (C35) and the Hermiticity of X_k, we recast the equality constraints in Eq. (C22) as

$$c_{0,k}(\mathbf{Y}) = \frac{1}{2} \sum_{\ell=1}^{d} (T_{k,\ell} \mathrm{Tr}[\rho Y_\ell] + T_{k,\ell}^* \mathrm{Tr}[\rho Y_\ell^\dagger]) = 0, \quad (C36)$$

$$c_{j,k}(\mathbf{Y}) = \frac{1}{2} \sum_{\ell=1}^{d} (T_{k,\ell} \mathrm{Tr}[\rho_j Y_\ell] + T_{k,\ell}^* \mathrm{Tr}[\rho_j Y_\ell^\dagger]) - \delta_{j,k} = 0. \qquad (C37)$$

Since the variables Y_ℓ 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 Eq. (C35) that

$$\sum_{\ell=1}^{d} (T_{k,\ell} Y_\ell - T_{k,\ell}^* Y_\ell^\dagger) = 0. \qquad (C38)$$

The left side of Eq. (C38) is, in general, an anti-Hermitian matrix, and, to make it Hermitian, we multiply both sides by i to get

$$H_k(\mathbf{Y}) = i \sum_{\ell=1}^{d} (T_{k,\ell} Y_\ell - T_{k,\ell}^* Y_\ell^\dagger) = 0. \qquad (C39)$$

With all these constraints, we recast the optimization problem (C22) as the following optimization problem:

$$\begin{aligned}
&\underset{Y_1, \ldots, Y_d, t}{\text{minimize}} \; t, \\
&\text{subject to } c_{0,k}(\mathbf{Y}) = 0 \\
&\qquad\qquad\quad c_{j,k}(\mathbf{Y}) = 0 \\
\frac{1}{2} \sum_{\ell=1}^{d} (\delta_{0,\alpha_\ell} \mathrm{Tr}[Y_\ell \rho Y_\ell^\dagger] + \delta_{1,\alpha_\ell} \mathrm{Tr}[Y_\ell^\dagger \rho Y_\ell]) &\leq t \\
&\qquad\qquad\quad H_k(\mathbf{Y}) = 0 \\
&\qquad\qquad\quad \boldsymbol{\alpha} \in \{0,1\}^d.
\end{aligned} \qquad (C40)$$

4. Analysis on the Lagrangian

Here, we consider the constraints in Eq. (C40) 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 Eq. (C40) can then be written as

$$\begin{aligned}
L_d = {}& t + \sum_{j=0}^{d} \sum_{k=1}^{d} z_{j,k} c_{j,k}(\mathbf{Y}) + \sum_{k=1}^{d} \mathrm{Tr}[\xi_k H_k(\mathbf{Y})] \\
&+ \frac{1}{2} \sum_{\boldsymbol{\alpha} \in \{0,1\}^d} v_\alpha \sum_{\ell=1}^{d} (\delta_{0,\tilde{\alpha}_\ell} \mathrm{Tr}[Y_\ell \rho Y_\ell^\dagger] + \delta_{1,\tilde{\alpha}_\ell} \mathrm{Tr}[Y_\ell^\dagger \rho Y_\ell]) \\
&- \sum_{\boldsymbol{\alpha} \in \{0,1\}^d} v_\alpha t. \qquad (C41)
\end{aligned}$$

Here, the Lagrange multipliers $z_{j,k}$ are real numbers, while the Lagrange multipliers v_α are non-negative numbers. The Lagrange multipliers ξ_k are Hermitian matrices in $\mathbb{H}_N$. Note that the multiparameter Lagrangian is a quadratic form in $\mathbf{Y}$ and, as such, can be minimized using Lemma 5. Before we do so, we consider the minimization of the Lagrangian with respect to the primal variable t.

If the Lagrangian multipliers v_α do not all sum to one, by picking t to approach either positive or negative infinity, the Lagrangian L_d becomes unbounded. Hence, the optimal multipliers v_α must sum to one. By picking a discrete set of values of v_α where v_α is equal to zero to all but one value of α and maximizing the Lagrange dual function for each of these cases, we can obtain our lower bound to the multiparameter HCRB.

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

$$\begin{aligned}
L_{d,\boldsymbol{\alpha}} = {}& \sum_{j=0}^{d} \sum_{k=1}^{d} z_{j,k} c_{j,k}(\mathbf{Y}) + \sum_{k=1}^{d} \mathrm{Tr}[\xi_k H_k(\mathbf{Y})] \\
&+ \frac{1}{2} \sum_{\ell=1}^{d} (\delta_{0,\tilde{\alpha}_\ell} \mathrm{Tr}[Y_\ell \rho Y_\ell^\dagger] + \delta_{1,\tilde{\alpha}_\ell} \mathrm{Tr}[Y_\ell^\dagger \rho Y_\ell]). \qquad (C42)
\end{aligned}$$

Now, define

$$\Gamma_\ell = \sum_{k=1}^{d} T_{k,\ell} \left(\sum_{j=0}^{d} z_{j,k} \rho_j + i\xi_k \right). \qquad (C43)$$

By rewriting the terms on the first line on the right side of Eq. (C42), the effective Lagrangian becomes

$$\begin{aligned}
L_{d,\boldsymbol{\alpha}} = {}& -\sum_{j=1}^{d} z_{j,j} + \frac{1}{2} \sum_{\ell=1}^{d} (\mathrm{Tr}[\Gamma_\ell Y_\ell] + \mathrm{Tr}[\Gamma_\ell^\dagger Y_\ell^\dagger] \\
&+ \delta_{0,\tilde{\alpha}_\ell} \mathrm{Tr}[Y_\ell \rho Y_\ell^\dagger] + \delta_{1,\tilde{\alpha}_\ell} \mathrm{Tr}[Y_\ell^\dagger \rho Y_\ell]). \qquad (C44)
\end{aligned}$$

Then, given that ρ is a Hermitian full-rank matrix, we can use Lemma 5 to get the corresponding Lagrange dual to be

$$g_\alpha = \min_{\mathbf{Y}} L_{d,\boldsymbol{\alpha}}$$

$$= -\sum_{j=1}^{d} z_{j,j} - \sum_{\ell=1}^{d} \frac{\delta_{0,\bar{\alpha}_\ell} \mathrm{Tr}[\Gamma_\ell \rho^{-1} \Gamma_\ell^\dagger] + \delta_{1,\bar{\alpha}_\ell} \mathrm{Tr}[\Gamma_\ell^\dagger \rho^{-1} \Gamma_\ell]}{2}.$$

$$(C45)$$

Our lower bound to the HCRB is, thus,

$$\max_{\boldsymbol{\alpha} \in \{0,1\}^d} \max\{g_\alpha : z_{j,k} \in \mathbb{R}, \xi_k \in \mathbb{H}_D\}, \qquad (C46)$$

where $j = \{0, \ldots, d\}, k = \{1, \ldots, d\}$. Any feasible value of g_α yields a lower bound to the HCRB.

APPENDIX D: MINIMIZING THE LAGRANGIAN

In this Appendix, we extend our formalism to account for arbitrary values of the Lagrange dual variable u. When $u + v = 1$, we minimize the Lagrangian, which we recall has the form

$$\mathsf{L}(Y, u, \mathbf{z}) = -\mathbf{b}^\top \mathbf{z} + u\mathrm{Tr}[Y\rho Y^\dagger] + (1-u)\mathrm{Tr}[Y^\dagger \rho Y]$$
$$+ \mathrm{Tr}[AY] + \mathrm{Tr}[A^\dagger Y^\dagger], \qquad (D1)$$

where

$$A = (z_1\rho + z_2\rho_1 + z_3\rho_2 - iz_4\rho - iz_5\rho_1 - iz_6\rho_2)/2. \quad (D2)$$

Recall the notation $\rho_j = \partial_j\rho, j \in \{1, 2\}$. The Fréchet derivative of the Lagrangian in the matrix direction H is given by

$$\nabla_Y(\mathsf{L}, H) = \lim_{h \to 0} \frac{\mathsf{L}(Y + hH, u, \mathbf{z}) - \mathsf{L}(Y, u, \mathbf{z})}{h}. \quad (D3)$$

Lemma 8.—Let L be the Lagrangian as defined in Eq. (D1) and A be as given in Eq. (D2). Suppose that

$$uY\rho + (1-u)\rho Y + A^\dagger = 0. \qquad (D4)$$

Then, $\nabla_Y(L, H) = 0$.
Proof.—Notice that

$$\nabla_Y(\mathsf{L}, H) = u(\mathrm{Tr}[H\rho Y^\dagger] + \mathrm{Tr}[Y\rho H^\dagger]) + (1-u)(\mathrm{Tr}[H^\dagger\rho Y]$$
$$+ \mathrm{Tr}[Y^\dagger\rho H]) + \mathrm{Tr}[AH] + \mathrm{Tr}[A^\dagger H^\dagger]. \qquad (D5)$$

Now, we use the cyclic property of the trace to write

$$\nabla_Y(\mathsf{L}, H) = u(\mathrm{Tr}[\rho Y^\dagger H] + \mathrm{Tr}[Y\rho H^\dagger]) + (1-u)(\mathrm{Tr}[\rho YH^\dagger]$$
$$+ \mathrm{Tr}[Y^\dagger\rho H]) + \mathrm{Tr}[AH] + \mathrm{Tr}[A^\dagger H^\dagger]$$
$$= \mathrm{Tr}[B^\dagger H] + \mathrm{Tr}[BH^\dagger], \qquad (D6)$$

where $B = uY\rho + (1-u)\rho Y + A^\dagger$ and $B^\dagger = u\rho Y^\dagger + (1-u)Y^\dagger\rho + A$. Since $B = 0$ by the assumption of our lemma, we must have $B^\dagger = 0$, and it follows that $\nabla_Y(\mathsf{L}, H) = 0$. ∎

Notice that Eq. (D4) is a Sylvester equation, and solving it is a standard procedure, where a variant of the Bartels-Stewart algorithm can apply. When ρ is a full-rank matrix, we can solve this equation analytically. Let ρ have the spectral decomposition $\rho = \sum_j p_j|e_j\rangle\langle e_j|$, where $|e_j\rangle$ are normalized eigenvectors of ρ. In this case, Eq. (D4) is equivalent to

$$u\sum_{j,k}\langle e_j|Y|e_k\rangle|e_j\rangle\langle e_k|p_k + (1-u)\sum_{j,k}p_j|e_j\rangle\langle e_k|\langle e_j|Y|e_k\rangle$$
$$+ \sum_{j,k}|e_j\rangle\langle e_k|\langle e_j|A^\dagger|e_k\rangle = 0. \qquad (D7)$$

Simplifying this equation, we get

$$\sum_{j,k}\langle e_j|Y|e_k\rangle[up_k + (1-u)p_j]|e_j\rangle\langle e_k|$$
$$= -\sum_{j,k}|e_j\rangle\langle e_k|\langle e_j|A^\dagger|e_k\rangle, \qquad (D8)$$

from which it follows that

$$\langle e_j|Y|e_k\rangle = -[up_k + (1-u)p_j]^{-1}\langle e_j|A^\dagger|e_k\rangle. \qquad (D9)$$

The following lemma then follows.
Lemma 9.—The Y that minimizes the Lagrangian is given by

$$Y = -\sum_{j,k}[up_k + (1-u)p_j]^{-1}\langle e_j|A^\dagger|e_k\rangle|e_j\rangle\langle e_k|. \qquad (D10)$$

Crucially, we can write the Y that minimizes the Lagrangian as a linear combination of the Lagrange multipliers $\mathbf{z}$.
Lemma 10.—The Y that minimizes the Lagrangian is given by

$$Y = z_1\gamma_1 + z_2\gamma_2 + z_3\gamma_3 + iz_4\gamma_1 + iz_5\gamma_2 + iz_6\gamma_3, \quad (D11)$$

where the complex matrices

$$\gamma_1 = -\mathbf{1}_d/2, \qquad (D12)$$

$$\gamma_2 = -\sum_{j,k}[up_k + (1-u)p_j]^{-1}\langle e_j|\rho_1/2|e_k\rangle|e_j\rangle\langle e_k|,$$

$$(D13)$$

$$\gamma_3 = -\sum_{j,k}[up_k + (1-u)p_j]^{-1}\langle e_j|\rho_2/2|e_k\rangle|e_j\rangle\langle e_k|.$$

$$(D14)$$

Proof.—Recall the definition of A in Eq. (D2). Then, through Lemma 9, we find

$$Y = -(z_1 + iz_4) \sum_{j,k} [up_k + (1-u)p_j]^{-1} \langle e_j | \rho/2 | e_k \rangle |e_j\rangle\langle e_k|$$

$$- (z_2 + iz_5) \sum_{j,k} [up_k + (1-u)p_j]^{-1} \langle e_j | \rho_1/2 | e_k \rangle |e_j\rangle\langle e_k|$$

$$- (z_3 + iz_6) \sum_{j,k} [up_k + (1-u)p_j]^{-1} \langle e_j | \rho_2/2 | e_k \rangle |e_j\rangle\langle e_k|. \tag{D15}$$

Now, we can make the simplification

$$\sum_{j,k} [up_k + (1-u)p_j]^{-1} \langle e_j | \rho/2 | e_k \rangle |e_j\rangle\langle e_k| = \sum_{j} [up_j + (1-u)p_j]^{-1} p_j/2 |e_j\rangle\langle e_j| = \mathbf{1}_d/2. \tag{D16}$$

Hence, the result follows. ∎

By substituting the optimal value of Y back into the Lagrangian, we find that the Lagrangian is a quadratic in $\mathbf{z}$. Namely, we have the following.

Lemma 11.—For fixed u such that $0 < u < 1$, and where $\mathbf{z} = (z_1, \ldots, z_6) \in \mathbb{R}^6$, the Lagrange dual of our Lagrangian is

$$g(u, \mathbf{z}) = -\mathbf{b}^\top \mathbf{z} + \mathbf{z}^\top Q \mathbf{z}, \tag{D17}$$

where

$$Q = u\mathrm{Re}(Q_1) + (1-u)\mathrm{Re}(Q_2) + \mathrm{Re}(Q_3), \tag{D18}$$

$$Q_1 = \begin{pmatrix} G_1 & -iG_1 \\ iG_1 & G_1 \end{pmatrix}, \qquad Q_2 = \begin{pmatrix} G_2 & iG_2 \\ -iG_2 & G_2 \end{pmatrix}, \qquad Q_3 = \frac{1}{2}\begin{pmatrix} G_3 + G_3^* & i(G_3 - G_3^*) \\ -i(G_3 - G_3^*) & G_3 + G_3^* \end{pmatrix}, \tag{D19}$$

and

$$G_1 = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & \mathrm{Tr}[\gamma_2\rho\gamma_2^\dagger] & \mathrm{Tr}[\gamma_2\rho\gamma_3^\dagger] \\ 0 & \mathrm{Tr}[\gamma_3\rho\gamma_2^\dagger] & \mathrm{Tr}[\gamma_3\rho\gamma_3^\dagger] \end{pmatrix}, \qquad G_2 = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & \mathrm{Tr}[\gamma_2^\dagger\rho\gamma_2] & \mathrm{Tr}[\gamma_2^\dagger\rho\gamma_3] \\ 0 & \mathrm{Tr}[\gamma_3^\dagger\rho\gamma_2] & \mathrm{Tr}[\gamma_3^\dagger\rho\gamma_3] \end{pmatrix}, \qquad G_3 = \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & \mathrm{Tr}[\rho_1\gamma_2] & \mathrm{Tr}[\rho_1\gamma_3] \\ 0 & \mathrm{Tr}[\rho_2\gamma_2] & \mathrm{Tr}[\rho_2\gamma_3] \end{pmatrix}. \tag{D20}$$

Proof.—The Lagrange dual is given by substituting the optimal solution for Y in the Lagrangian minimization. Recall the definition of the Lagrangian in Eq. (D1), and then the first term to evaluate is

$$\mathrm{Tr}[Y\rho Y^\dagger] = \mathrm{Tr}[(z_1\gamma_1 + z_2\gamma_2 + z_3\gamma_3 + iz_4\gamma_1 + iz_5\gamma_2 + iz_6\gamma_3)\rho(z_1\gamma_1^\dagger + z_2\gamma_2^\dagger + z_3\gamma_3^\dagger - iz_4\gamma_1^\dagger - iz_5\gamma_2^\dagger - iz_6\gamma_3^\dagger)], \tag{D21}$$

where we use the optimal solution for Y as given in Lemma 10. Writing this solution in matrix form, we have

$$\mathrm{Tr}[Y\rho Y^\dagger] = \mathbf{z}^\top \begin{pmatrix} G_1 & -iG_1 \\ iG_1 & G_1 \end{pmatrix} \mathbf{z} = \mathbf{z}^\top Q_1 \mathbf{z}, \tag{D22}$$

where $\mathbf{z}$ is the column vector of Lagrange multipliers and the block matrix

$$G_1 = \begin{pmatrix} \mathrm{Tr}[\gamma_1\rho\gamma_1^\dagger] & \mathrm{Tr}[\gamma_1\rho\gamma_2^\dagger] & \mathrm{Tr}[\gamma_1\rho\gamma_3^\dagger] \\ \mathrm{Tr}[\gamma_2\rho\gamma_1^\dagger] & \mathrm{Tr}[\gamma_2\rho\gamma_2^\dagger] & \mathrm{Tr}[\gamma_2\rho\gamma_3^\dagger] \\ \mathrm{Tr}[\gamma_3\rho\gamma_1^\dagger] & \mathrm{Tr}[\gamma_3\rho\gamma_2^\dagger] & \mathrm{Tr}[\gamma_3\rho\gamma_3^\dagger] \end{pmatrix} = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & \mathrm{Tr}[\gamma_2\rho\gamma_2^\dagger] & \mathrm{Tr}[\gamma_2\rho\gamma_3^\dagger] \\ 0 & \mathrm{Tr}[\gamma_3\rho\gamma_2^\dagger] & \mathrm{Tr}[\gamma_3\rho\gamma_3^\dagger] \end{pmatrix}. \tag{D23}$$

Here, we use the definitions for γ_j in Lemma 10 and the traceless property of the state derivatives ρ_1 and ρ_2:

$$\text{Tr}[\rho\gamma_s] = \text{Tr}[\rho\gamma_s^\dagger] = \sum_j p_j[(1-u)p_j + up_j]^{-1}\langle e_j|\rho_{s-1}/2|e_j\rangle = \sum_j \langle e_j|\rho_{s-1}|e_j\rangle/2 = 0, \tag{D24}$$

for $s \in \{2,3\}$. Similarly, we determine the second term in the Lagrangian:

$$\text{Tr}[Y^\dagger\rho Y] = \text{Tr}[(z_1\gamma_1^\dagger + z_2\gamma_2^\dagger + z_3\gamma_3^\dagger - iz_4\gamma_1^\dagger - iz_5\gamma_2^\dagger - iz_6\gamma_3^\dagger)\rho(z_1\gamma_1 + z_2\gamma_2 + z_3\gamma_3 + iz_4\gamma_1 + iz_5\gamma_2 + iz_6\gamma_3)]. \tag{D25}$$

In matrix form, this term can similarly be written as

$$\text{Tr}[Y^\dagger\rho Y] = \mathbf{z}^\top \begin{pmatrix} G_2 & iG_2 \\ -iG_2 & G_2 \end{pmatrix} \mathbf{z} = \mathbf{z}^\top Q_2 \mathbf{z}, \tag{D26}$$

where

$$G_2 = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & \text{Tr}[\gamma_2^\dagger\rho\gamma_2] & \text{Tr}[\gamma_2^\dagger\rho\gamma_3] \\ 0 & \text{Tr}[\gamma_3^\dagger\rho\gamma_2] & \text{Tr}[\gamma_3^\dagger\rho\gamma_3] \end{pmatrix}. \tag{D27}$$

Finally, we substitute the optimal solution for Y into the last two terms of the Lagrangian:

$$\text{Tr}[AY + A^\dagger Y^\dagger] = \frac{1}{2}\text{Tr}[(z_1\rho + z_2\rho_1 + z_3\rho_2 - iz_4\rho - iz_5\rho_1 - iz_6\rho_2)(z_1\gamma_1 + z_2\gamma_2 + z_3\gamma_3 + iz_4\gamma_1 + iz_5\gamma_2 + iz_6\gamma_3)$$

$$+ (z_1\rho + z_2\rho_1 + z_3\rho_2 + iz_4\rho + iz_5\rho_1 + iz_6\rho_2)(z_1\gamma_1^\dagger + z_2\gamma_2^\dagger + z_3\gamma_3^\dagger - iz_4\gamma_1^\dagger - iz_5\gamma_2^\dagger - iz_6\gamma_3^\dagger)], \tag{D28}$$

where we use the Hermiticity of the state derivatives. Hence,

$$\text{Tr}[AY + A^\dagger Y^\dagger] = \frac{1}{2}\mathbf{z}^\top \begin{pmatrix} G_3 & iG_3 \\ -iG_3 & G_3 \end{pmatrix} \mathbf{z} + \frac{1}{2}\mathbf{z}^\top \begin{pmatrix} G_4 & -iG_4 \\ iG_4 & G_4 \end{pmatrix} \mathbf{z} = \mathbf{z}^\top Q_3 \mathbf{z}, \tag{D29}$$

where

$$G_3 = \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & \text{Tr}[\rho_1\gamma_2] & \text{Tr}[\rho_1\gamma_3] \\ 0 & \text{Tr}[\rho_2\gamma_2] & \text{Tr}[\rho_2\gamma_3] \end{pmatrix}, \quad G_4 = \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & \text{Tr}[\rho_1\gamma_2^\dagger] & \text{Tr}[\rho_1\gamma_3^\dagger] \\ 0 & \text{Tr}[\rho_2\gamma_2^\dagger] & \text{Tr}[\rho_2\gamma_3^\dagger] \end{pmatrix}. \tag{D30}$$

Notice that $G_4 = G_3^*$. Since the Lagrange dual must be real, and the dual variables u and $\mathbf{z}$ must be real, we can take the real part of the matrices Q_1, Q_2 and Q_3 to complete the proof. ∎

Because our optimization problem is convex, we are promised that $\mathsf{L}(Y, u, \mathbf{z})$ will be a concave function in both u and z. This promise implies that, for fixed u, $\mathsf{L}(Y, u, \mathbf{z})$ is concave in $\mathbf{z}$, which implies that Q is negative definite. From the stationary point of $g(u, \mathbf{z})$ with respect to $\mathbf{z}$, we find that the optimal solution to $\mathbf{z}$ is given by the solution to the linear equation

$$2Q\mathbf{z} = \mathbf{b}. \tag{D31}$$

If Q is full rank, then we reach the optimal values for the Lagrange multipliers:

$$\mathbf{z} = \frac{1}{2}Q^{-1}\mathbf{b}. \tag{D32}$$

Substituting this result into the solution for Y in Eq. (D11) provides an *Ansatz* that can saturate the HCRB, which is upper and lower bounded by

$$\mathcal{U}_u = \max\{\mathbf{z}^\top Q_1\mathbf{z}, \mathbf{z}^\top Q_2\mathbf{z}\},$$
$$\mathcal{L}_u = -\mathbf{b}^\top\mathbf{z} + \mathbf{z}^\top Q\mathbf{z}, \tag{D33}$$

respectively. For fixed u, our *Ansatz* for Y gives a tight bound when these two bounds are equivalent. If such a solution exists, we can optimize over the dual variable u to find the optimal value. This value can be determined numerically for any application in $\mathcal{O}[\text{polylog}(1/\epsilon)]$ time, where ϵ is the duality gap. Alternatively, we can find the

optimal u by looking solely at the lower bound to the HCRB:

$$\mathscr{L}_u = -\frac{1}{2}\mathbf{b}^\top \mathcal{Q}^{-1}\mathbf{b} + \frac{1}{4}\mathbf{b}^\top (\mathcal{Q}^{-1})^\top \mathcal{Q}\mathcal{Q}^{-1}\mathbf{b}$$
$$= -\frac{1}{4}\mathbf{b}^\top \mathcal{Q}^{-1}\mathbf{b}, \qquad \text{(D34)}$$

where we use Eq. (D32) and the fact that $\mathcal{Q}$ must be a symmetric matrix. The function l_u is continuous and differentiable with respect to u. Also, duality theory of convex optimization promises that $\mathscr{L}_u$ is concave in u. Hence, in the scenario where the optimal u is not attained for the values $u = \{0, 1\}$, we have that $\mathscr{L}_u$ is optimized at its stationary point $d\mathscr{L}_u/du = 0$. The optimality condition of our HCRB is, hence, reduced to finding the roots of the stationary points of $\mathscr{L}_u$. This concludes our proof of Theorem 3 in the main body of the text for any two-parameter estimation problem.

APPENDIX E: COMPLEXITY ANALYSIS FOR LAGRANGE MINIMIZATION

For fixed u, our method to bound the HCRB requires minimizing the Lagrangian. We find that this method amounts to solving the Sylvester equation $A^\dagger = -[uY\rho + (1 - u)\rho Y]$ for Y. For large dimensional systems, evaluating the HCRB with concrete analytical results becomes increasingly cumbersome. In this scenario, a numerical approach can be used to handle the state diagonalization to evaluate the HCRB. Here, we bound the complexity of our formalism to evaluating the HCRB numerically.

For large dimensional systems, the Sylvester equation is more efficiently solved by first vectorizing the equation to

$$\text{vec}(A^\dagger) = -[\mathbb{1}_D \otimes (1 - u)\rho + u\rho^\top \otimes \mathbb{1}_D]\text{vec}(Y) \qquad \text{(E1)}$$

and then solving using any system of linear equations solver. The Bartels-Stewart algorithm is an efficient and robust numerical solver for the Sylvester matrix for large D [64,89], which outperforms well-known primitive implementations of Gaussian elimination. The complexity of the Bartels-Stewart algorithm scales as $\mathcal{O}(D^3)$.

To circumvent any time complexity involved, we must have access to the basis that diagonalizes the state. In this case, we solve the Sylvester equation analytically. For large D, there are efficient numerical methods that can attain the spectral decomposition of the state in subcubic time. For example, with $\rho(\boldsymbol{\theta}) \in \mathbb{H}_D$, the matrix inversion operation can be practically achieved using the Coppersmith-Winograd algorithm, which scales as $\mathcal{O}(D^{2.376})$ [90].

As a point of comparison, semidefinite programming (SDP) provides an alternative method to optimization tasks. SDP programs can be applied to general problems and admit polynomial-time solvers, which highlight the power of this approach. The HCRB is recast as an SDP program in Ref. [44]. For a consistent complexity comparison with our method, we consider the nontrivial case where the state is full rank. Hence, by observation of Eq. (11) in Ref. [44], the variable X that is optimized has order D^2 terms. Furthermore, notice that, for each iteration of the SDP algorithm, the first constraint requires knowledge of the spectral decomposition of a matrix parameterized in terms of X. Therefore, it is easy to observe that this brute force SDP approach has a time complexity greater than $\mathcal{O}(D^{2\times 2.376})$. This result indicates at least a quadratic speedup, which amounts to a significant improvement with increasing D.

[1] S. L. Braunstein and C. M. Caves, *Statistical Distance and the Geometry of Quantum States*, Phys. Rev. Lett. **72**, 3439 (1994).

[2] V. Giovannetti, S. Lloyd, and L. Maccone, *Advances in Quantum Metrology*, Nat. Photonics **5**, 222 (2011).

[3] G. Tóth and I. Apellaniz, *Quantum Metrology from a Quantum Information Science Perspective*, J. Phys. A **47**, 424006 (2014).

[4] C. Helstrom, *The Minimum Variance of Estimates in Quantum Signal Detection*, IEEE Trans. Inf. Theory **14**, 234 (1968).

[5] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

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

[7] K. Matsumoto, *A New Approach to the Cramér-Rao-Type Bound of the Pure-State Model*, J. Phys. A **35**, 3111 (2002).

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

[9] R. D. Gill and M. Guta, *On Asymptotic Quantum Statistical Inference*, arXiv:1112.2078.

[10] D. Burgarth and A. Ajoy, *Evolution-Free Hamiltonian Parameter Estimation through Zeeman Markers*, Phys. Rev. Lett. **119**, 030402 (2017).

[11] M. Tsang, H. M. Wiseman, and C. M. Caves, *Fundamental Quantum Limit to Waveform Estimation*, Phys. Rev. Lett. **106**, 090401 (2011).

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

[13] T. Baumgratz and A. Datta, *Quantum Enhanced Estimation of a Multidimensional Field*, Phys. Rev. Lett. **116**, 030801 (2016).

[14] M. Tsang, *Quantum Imaging beyond the Diffraction Limit by Optical Centroid Measurements*, Phys. Rev. Lett. **102**, 253601 (2009).

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

[16] C. N. Gagatsos, D. Branford, and A. Datta, *Gaussian Systems for Quantum-Enhanced Multiple Phase Estimation*, Phys. Rev. A **94**, 042342 (2016).

[17] L. Pezzè, M. A. Ciampini, N. Spagnolo, P. C. Humphreys, A. Datta, I. A. Walmsley, M. Barbieri, F. Sciarrino, and

A. Smerzi, *Optimal Measurements for Simultaneous Quantum Estimation of Multiple Phases*, Phys. Rev. Lett. **119**, 130504 (2017).

[18] J. Rubio and J. Dunningham, *Bayesian Multiparameter Quantum Metrology with Limited Data*, Phys. Rev. A **101**, 032114 (2020).

[19] T. J. Proctor, P. A. Knott, and J. A. Dunningham, *Networked Quantum Sensing*, arXiv:1702.04271.

[20] J. S. Sidhu and P. Kok, *Quantum Metrology of Spatial Deformation Using Arrays of Classical and Quantum Light Emitters*, Phys. Rev. A **95**, 063829 (2017).

[21] J. S. Sidhu and P. Kok, *Quantum Fisher Information for General Spatial Deformations of Quantum Emitters*, arXiv:1802.01601.

[22] J. Rubio, P. A. Knott, T. J. Proctor, and J. A. Dunningham, *Quantum Sensing Networks for the Estimation of Linear Functions*, J. Phys. A **53**, 344001 (2020).

[23] M. G. A. Paris, *Quantum Estimation for Quantum Technologies*, Int. J. Quantum. Inform. **07**, 125 (2009).

[24] D. A. S. Fraser, *On Local Unbiased Estimation*, J. R. Stat. Soc. Ser. B **26**, 46 (1964).

[25] M. J. W. Hall and H. M. Wiseman, *Does Nonlinear Metrology Offer Improved Resolution? Answers from Quantum Information Theory*, Phys. Rev. X **2**, 041006 (2012).

[26] R. Demkowicz-Dobrzanski, M. Jarzyna, and J. Kołodyński, *Quantum Limits in Optical Interferometry*, Prog. Opt. **60**, 345 (2015).

[27] S. Kay, *Estimation Theory*, Fundamentals of Statistical Signal Processing Vol. I, 1 ed. (Prentice-Hall, Englewood Cliffs, NJ, 1993).

[28] L. Pezzé and A. Smerzi, *Quantum Theory of Phase Estimation*, in *Proceedings of the International School of Physics Enrico Fermi*, Course 188, Vol. 691 (IOS, Amsterdam, 2014).

[29] C. L. Degen, F. Reinhard, and P. Cappellaro, *Quantum Sensing*, Rev. Mod. Phys. **89**, 035002 (2017).

[30] J. S. Sidhu and P. Kok, *Geometric Perspective on Quantum Parameter Estimation*, AVS Quantum Sci. **2**, 014701 (2020).

[31] H. Zhu, *Information Complementarity: A New Paradigm for Decoding Quantum Incompatibility*, Sci. Rep. **5**, 14317 (2015).

[32] T. Heinosaari, T. Miyadera, and M. Ziman, *An Invitation to Quantum Incompatibility*, J. Phys. A **49**, 123001 (2016).

[33] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, *Compatibility in Multiparameter Quantum Metrology*, Phys. Rev. A **94**, 052108 (2016).

[34] M. Guţă and J. Kahn, *Local Asymptotic Normality for Qubit States*, Phys. Rev. A **73**, 052108 (2006).

[35] M. Hayashi and K. Matsumoto, *Asymptotic Performance of Optimal State Estimation in Qubit System*, J. Math. Phys. (N.Y.) **49**, 102101 (2008).

[36] K. Yamagata, A. Fujiwara, and R. D. Gill, *Quantum Local Asymptotic Normality Based on a New Quantum Likelihood Ratio*, Ann. Statist. **41**, 2197 (2013).

[37] Y. Yang, G. Chiribella, and M. Hayashi, *Attaining the Ultimate Precision Limit in Quantum State Estimation*, Commun. Math. Phys. **368**, 223 (2019).

[38] J. Suzuki, *Explicit Formula for the Holevo Bound for Two-Parameter Qubit-State Estimation Problem*, J. Math. Phys. (N.Y.) **57**, 042201 (2016).

[39] M. Bradshaw, S. M. Assad, and P. K. Lam, *A Tight Cramér-Rao Bound for Joint Parameter Estimation with a Pure Two-Mode Squeezed Probe*, Phys. Lett. A **381**, 2598 (2017).

[40] P. A. Knott, T. J. Proctor, A. J. Hayes, J. F. Ralph, P. Kok, and J. A. Dunningham, *Local versus Global Strategies in Multiparameter Estimation*, Phys. Rev. A **94**, 062312 (2016).

[41] J. Řeháček, Z. Hradil, B. Stoklasa, M. Paur, J. Grover, A. Krzic, and L. L. Sánchez-Soto, *Multiparameter Quantum Metrology of Incoherent Point Sources: Towards Realistic Superresolution*, Phys. Rev. A **96**, 062107 (2017).

[42] L. Zhang and K. W. C. Chan, *Quantum Multiparameter Estimation with Generalized Balanced Multimode NOON-like States*, Phys. Rev. A **95**, 032321 (2017).

[43] T. J. Proctor, P. A. Knott, and J. A. Dunningham, *Multiparameter Estimation in Networked Quantum Sensors*, Phys. Rev. Lett. **120**, 080501 (2018).

[44] F. Albarelli, J. F. Friel, and A. Datta, *Evaluating the Holevo Cramér-Rao Bound for Multiparameter Quantum Metrology*, Phys. Rev. Lett. **123**, 200503 (2019).

[45] F. Fröwis, R. Schmied, and N. Gisin, *Tighter Quantum Uncertainty Relations Following from a General Probabilistic Bound*, Phys. Rev. A **92**, 012102 (2015).

[46] R. D. Gill and S. Massar, *State Estimation for Large Ensembles*, Phys. Rev. A **61**, 042312 (2000).

[47] O. E. Barndorff-Nielsen and R. D. Gill, *Fisher Information in Quantum Statistics*, J. Phys. A **33**, 4481 (2000).

[48] E. Bagan, M. Baig, R. Muñoz-Tapia, and A. Rodriguez, *Collective versus Local Measurements in a Qubit Mixed-State Estimation*, Phys. Rev. A **69**, 010304(R) (2004).

[49] E. Bagan, M. A. Ballester, R. D. Gill, A. Monras, and R. Muñoz-Tapia, *Optimal Full Estimation of Qubit Mixed States*, Phys. Rev. A **73**, 032301 (2006).

[50] J. Kahn and M. Guta, *Local Asymptotic Normality for Finite Dimensional Quantum Systems*, Commun. Math. Phys. **289**, 597 (2009).

[51] C. W. Helstrom, *Minimum Mean-Squared Error of Estimates in Quantum Statistics*, Phys. Lett. A **25**, 101 (1967).

[52] H. P. Yuen and M. Lax, *Multiple-Parameter Quantum Estimation and Measurement of Nonselfadjoint Observables*, IEEE Trans. Inf. Theory **19**, 740 (1973).

[53] V. P. Belavkin, *Generalized Uncertainty Relations and Efficient Measurements in Quantum Systems*, Theor. Math. Phys. **26**, 213 (1976).

[54] One can also show this result by first proving that $V \geq M$ implies that $\mathrm{Tr}[V] \geq \mathrm{Tr}[\mathrm{Re}[M]] + \mathrm{Tr}|\mathrm{Im}[M]|$ and subsequently replacing V and M by $\sqrt{W} V \sqrt{W}$ and $\sqrt{W} M \sqrt{W}$, respectively. The multiple on both sides by $\sqrt{W}$ is to ensure that $\sqrt{W} V \sqrt{W}$ is symmetric and $\sqrt{W} M \sqrt{W}$ is Hermitian.

[55] H. Nagaoka, *A New Approach to Cramér-Rao Bounds for Quantum State Estimation*, IT 89-42, IEICE technical report, 1989, reprinted in Ref. [8], Chap. 8, pp. 100–112.

[56] M. G. Genoni, M. G. A. Paris, G. Adesso, H. Nha, P. L. Knight, and M. S. Kim, *Optimal Estimation of Joint Parameters in Phase Space*, Phys. Rev. A **87**, 012107 (2013).

[57] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).

[58] A. Carollo, B. Spagnolo, A. A. Dubkov, and D. Valenti, *On Quantumness in Multi-parameter Quantum Estimation*, J. Stat. Mech. (2019) 094010.

[59] M. Tsang, *The Holevo Cramér-Rao Bound Is at Most Thrice the Helstrom Version*, arXiv:1911.08359.

[60] F. Albarelli, M. Tsang, and A. Datta, *Quantum Semiparametric Estimation*, Phys. Rev. X **10,** 031023 (2020).

[61] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, *Compatibility in Multiparameter Quantum Metrology*, Phys. Rev. A **94,** 052108 (2016).

[62] J. Suzuki, *Information Geometrical Characterization of Quantum Statistical Models in Quantum Estimation Theory*, Entropy **21,** 703 (2019).

[63] J. Nocedal and S. Wright, *Numerical Optimization* (Springer Science & Business Media, New York, 2006).

[64] R. H. Bartels and G. W. Stewart, *Algorithm 432: Solution of the Matrix Equation* ax + xb = c, Commun. ACM **15,** 820 (1972).

[65] J. Gorski, F. Pfeuffer, and K. Klamroth, *Biconvex Sets and Optimization with Biconvex Functions: A Survey and Extensions*, Math. Methods Oper. Res. **66,** 373 (2007), http://www2.math.uni-wuppertal.de/~klamroth/publications/gopfkl07.pdf.

[66] L. Razzoli, L. Ghirardi, I. Siloi, P. Bordone, and M. G. A. Paris, *Lattice Quantum Magnetometry*, Phys. Rev. A **99,** 062330 (2019).

[67] Y.-X. Liu, A. Ajoy, and P. Cappellaro, *Nanoscale Vector dc Magnetometry via Ancilla-Assisted Frequency Up-Conversion*, Phys. Rev. Lett. **122,** 100501 (2019).

[68] S. Pang and T. A. Brun, *Quantum Metrology for a General Hamiltonian Parameter*, Phys. Rev. A **90,** 022117 (2014).

[69] J. A. Jones, S. D. Karlen, J. Fitzsimons, A. Ardavan, S. C. Benjamin, G. A. D. Briggs, and J. J. L. Morton, *Magnetic Field Sensing beyond the Standard Quantum Limit Using 10-Spin NOON States*, Science **324,** 1166 (2009).

[70] Y. Ouyang, *Permutation-Invariant Quantum Codes*, Phys. Rev. A **90,** 062317 (2014).

[71] Y. Ouyang, N. Shettell, and D. Markham, *Robust Quantum Metrology with Explicit Symmetric States*, arXiv:1908.02378.

[72] D. Gottesman, A. Kitaev, and J. Preskill, *Encoding a Qubit in an Oscillator*, Phys. Rev. A **64,** 012310 (2001).

[73] K. Noh, S. M. Girvin, and L. Jiang, *Encoding an Oscillator into Many Oscillators* Phys. Rev. Lett. **125,** 080503 (2020).

[74] I. L. Chuang, D. W. Leung, and Y. Yamamoto, *Bosonic Quantum Codes for Amplitude Damping*, Phys. Rev. A **56,** 1114 (1997).

[75] W. Wasilewski and K. Banaszek, *Protecting an Optical Qubit against Photon Loss*, Phys. Rev. A **75,** 042316 (2007).

[76] M. Bergmann and P. van Loock, *Quantum Error Correction against Photon Loss Using NOON States*, Phys. Rev. A **94,** 012311 (2016).

[77] Y. Ouyang and R. Chao, *Permutation-Invariant Constant-Excitation Quantum Codes for Amplitude Damping*, IEEE Trans. Inf. Theory **66,** 2921 (2019).

[78] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, *New Class of Quantum Error-Correcting Codes for a Bosonic Mode*, Phys. Rev. X **6,** 031006 (2016).

[79] A. L. Grimsmo, J. Combes, and B. Q. Baragiola, *Quantum Computing with Rotation-Symmetric Bosonic Codes*, Phys. Rev. X **10,** 011058 (2020).

[80] B. M. Terhal, *Quantum Error Correction for Quantum Memories*, Rev. Mod. Phys. **87,** 307 (2015).

[81] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, *New Class of Quantum Error-Correcting Codes for a Bosonic Mode*, Phys. Rev. X **6,** 031006 (2016).

[82] B. M. Terhal, J. Conrad, and C. Vuillot, *Towards Scalable Bosonic Quantum Error Correction*, Quantum Sci. Technol. **5,** 043001 (2020).

[83] The parameters G and n here are reminiscent of GNU states in permutation-invariant quantum codes [70].

[84] Y. Ouyang and N. Rengaswamy, *Weight Distribution of Classical Codes Influences Robust Quantum Metrology*, arXiv:2007.02859.

[85] E. Deadman and S. D. Relton, *Taylor's Theorem for Matrix Functions with Applications to Condition Number Estimation*, Linear Algebra Appl. **504,** 354 (2016).

[86] Y. Ouyang, *Quantum Storage in Quantum Ferromagnets*, arXiv:1904.01458.

[87] M. J. Best, *Portfolio Optimization*, 1st ed. (Chapman and Hall/CRC, London, 2010).

[88] A constraint function is $f(x)$ written in the canonical form $f(x) = 0$ or $f(x) \le 0$.

[89] S. N. G. Golub and C. V. Loan, *A Hessenberg-Schur Method for the Problem* ax + xb = c, IEEE Trans. Autom. Control **24,** 909 (1979).

[90] D. Coppersmith and S. Winograd, *Matrix Multiplication via Arithmetic Progressions*, J. Symb. Comput. **9,** 251 (1990).