Tight Cramér-Rao type bounds for multiparameter quantum metrology through conic programming

Masahito Hayashi[1,2,3] and Yingkai Ouyang[4]

[1]School of Data Science, The Chinese University of Hong Kong, Shenzhen, Longgang District, Shenzhen, 518172, China

[2]International Quantum Academy (SIQA), Futian District, Shenzhen 518048, China

[3]Graduate School of Mathematics, Nagoya University, Nagoya, 464-8602, Japan

[4]Department of Physics & Astronomy, University of Sheffield, Sheffield, S3 7RH, United Kingdom

In the quest to unlock the maximum potential of quantum sensors, it is of paramount importance to have practical measurement strategies that can estimate incompatible parameters with best precisions possible. However, it is still not known how to find practical measurements with optimal precisions, even for uncorrelated measurements over probe states. Here, we give a concrete way to find uncorrelated measurement strategies with optimal precisions. We solve this fundamental problem by introducing a framework of conic programming that unifies the theory of precision bounds for multiparameter estimates for uncorrelated and correlated measurement strategies under a common umbrella. Namely, we give precision bounds that arise from linear programs on various cones defined on a tensor product space of matrices, including a particular cone of separable matrices. Subsequently, our theory allows us to develop an efficient algorithm that calculates both upper and lower bounds for the ultimate precision bound for uncorrelated measurement strategies, where these bounds can be tight. In particular, the uncorrelated measurement strategy that arises from our theory saturates the upper bound to the ultimate precision bound. Also, we show numerically that there is a strict gap between the previous efficiently computable bounds and the ultimate precision bound.

1 Introduction

Quantum sensors, by employing quantum resources, promise to estimate physical parameters with unprecedented precision beyond what is possible using classical resources. Quantum metrology is a research field that studies quantum sensors. Quantum metrology schemes require the ability to both prepare parameter-dependent quantum probe states and perform quantum measurements on these states. Armed with the statistics of the measurement outcomes, one can thereafter estimate the underlying parameters. A central question in quantum metrology is to find measurement strategies with the ultimate precision for these multiparameter estimates. In the simplest scenario of estimating a single-parameter, the ultimate precision, given by the quantum (CR) Cramér-Rao bound

Masahito Hayashi: hmasahito@cuhk.edu.cn, masahito@math.nagoya-u.ac.jp

Yingkai Ouyang: y.ouyang@sheffield.ac.uk

[1, 2, 3, 4, 5], along with the corresponding optimal measurement strategy, are efficient to compute with knowledge of both the probe state and its dependence on the single parameter.

The theory of multiparameter quantum metrology[1] is considerably richer than the single parameter setting. For example, parameters can be fundamentally incompatible, as is often the case in quantum systems. If we want to design the best quantum sensors that can simultaneously estimate incompatible parameters, we must find optimal practical measurement strategies for multiparameter quantum metrology. However, even after decades of research, the question of how to determine these optimal measurement strategies remains unanswered and unknown.

In lieu of determining these optimal measurement strategies, the field has focused on determining bounds for the ultimate precision of quantum sensors that estimate incompatible parameters simultaneously. Since most optimizations for precision bounds [3, 5, 8, 9, 10, 11, 12] are not based directly on measurement strategies, even if we can calculate the best precision bounds from these optimizations, we still will not know what the optimal measurement strategies are.

Regarding the theory of precision bounds, the theory of single parameter estimation differs substantially from the multiparameter case. Namely, while the CR bound is tight for both correlated and uncorrelated measurement strategies in single-parameter estimation, this is not the case for multiple parameters. This is because the SLD Cramér Rao (SLD) bound does not give the tight bound in the multiple-parameter case nevertheless it gives the tight bound in the single parameter case. That is, for multiparameter quantum metrology, the Holevo-Nagaoka (HN) bound [3, 4, 5, 8, 9] is efficient to compute and always tight for correlated measurement strategies across multiple probe states. Although it is often called the Holevo Cramér Rao bound, it is called Holevo-Nagaoka bound in this paper and its reason is explained later. Such correlated measurement strategies however require a large quantum device across multiple probe states, this bound is not practical. To accomplish state estimation in a practical way, we need to design uncorrelated measurement strategies across multiple probe states, using only measurement devices that access individual probe states. (See Fig. 1.) The Nagaoka-Hayashi (NH) [4, 10, 11, 12] bound is not only efficient to compute, but also addresses such uncorrelated measurement strategies. However, the NH bound might not be always tight for uncorrelated measurement strategies. Hence, an efficient way to determine the ultimate precision bound for uncorrelated strategies remains unknown.

The ultimate precision bound for uncorrelated strategies in the multiparameter setting, or simply the tight bound, was formulated more than two decades ago as the optimal value of an infinite-dimensional optimization program with linear objective and constraint functions on a topological vector space [13]. The advantage of this formulation of tight bound is that it is a direct optimization over measurement strategies in contrast to all other multiparameter precision bounds.

There however remain many outstanding questions pertaining to the tight bound.

(1) How to efficiently determine the tight bound?

(2) How to determine the optimal uncorrelated measurement strategy for multiparameter quantum metrology that saturates the tight bound?

(3) What is the relationship between the SLD bound, the HN bound, the NH bound, and the tight bound?
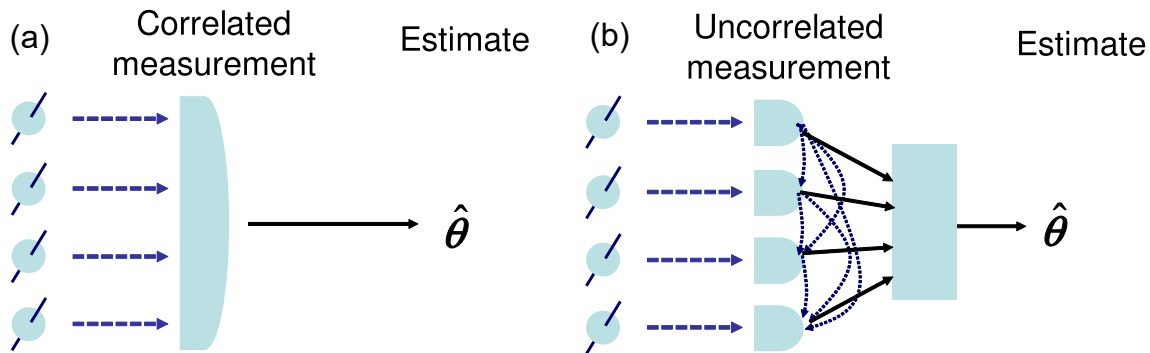
---

[1]See [6] and [7, Section V] for a recent review.

Figure 1: Measurement strategies in parameter estimation using quantum probe states can be either correlated (a), or uncorrelated (b) across identical copies of the probe states. In (a), the measurement device collectively measures the input states. That is, the measurement device needs to access a large quantum system. In (b) the measurement device individually measures the input states, but classical feedback is allowed to improve the measurement. This strategy is composed of measurement device to access a single system.

(4) Is there a gap between the tight bound and the NH bound?

With regards to question (3), we unify the theory of the SLD, HN, NH and tight bounds under a common umbrella. Remarkably, these bounds can alternatively be formulated as conic programs with the same linear objective and constraint functions. The only difference between these programs is the different choices of their cones. Namely, the cone for the tight bound is a strict subset of the cone for the NH bound, and the cone for the NH bound is a strict subset of the cone for the HN and SLD bounds. This solves the open problem (3).

Regarding question (2), using our reformulation of the tight bound, we also construct an efficient algorithm that calculates the tight bound. From the dual program of our reformulation of the tight bound, we construct an associated semidefinite program (SDP), and show how to use its optimal solution to approximately solve the tight bound. Namely, we provide an efficient algorithm that calculates both upper and lower bounds to the tight bound, thereby solving open problem (4). Regarding question (4), using our algorithm, we numerically demonstrate that the tight bound can be strictly tighter than the NH bound.

We also solve questions (1) and (2), where we show concretely how we can efficiently compute optimal uncorrelated measurement strategies that saturate the tight bound along with the tight bound.

Our numerically tight estimates to the tight bounds are useful beyond multiparameter estimation theory. More abstractly, we develop an approach to optimize over the separable cone for bipartite systems. Hence we expect our theory can pave the way towards gaining new insights into optimizations over separable bipartite states, which could be useful for entanglement theory [14, 15] and more general quantum resource theories [16, 17, 18].

Now we sketch the organization of our paper. We also give the structure of our paper visually in Fig. 2. In Table 2, we list the important notations we used in our paper along with their meanings. In Section 2 we review various CR-type bounds in multiparameter quantum metrology and define the tight bound. In Section 3, we answer question (3), where we unify various CR-type bounds via conic linear programming with various cones on a space that is the tensor product of real symmetric matrices and complex Hermitian matrices. Here, we reformulate the tight bound, (i.e.Theorem 2), the HN bound (Theorem 4) and the NH bound (Theorem 4). In Section 4, we develop our theory of how to calculate upper and lower bounds on the tight bound by applying an SDP with constraints labeled by
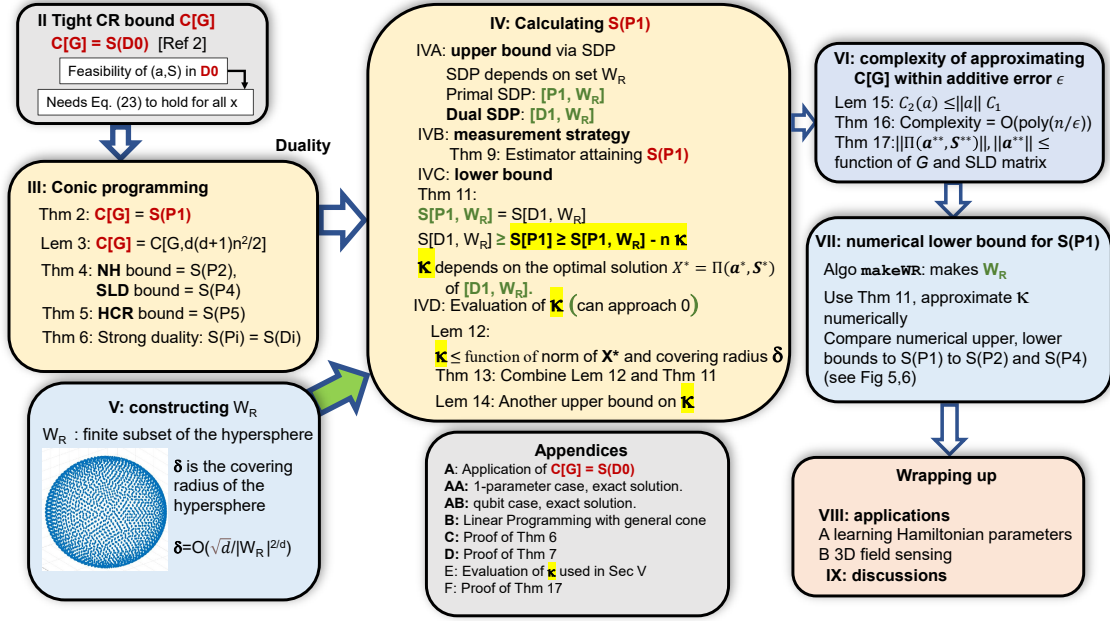
Figure 2: Structure of our paper.

unit vectors from a real vector space. We also derive an SDP with optimal value equal to the upper bound, and which is directly optimized over uncorrelated measurement strategies. From the solution for this SDP, we derive a concrete uncorrelated measurement strategy that has its precision given by our upper bound to the tight bound. Since our two-sided bounds to the bound can be tight, we therefore have derived a near-optimal measurement strategy We thereby can compute optimal uncorrelated measurement strategies in the asymptotic limit, and this answers question (2). In Section 5, we construct a strategic subset of the above unit vectors from design theory. In Section 6, based on the results of Section 5, we give the calculation complexity of the tight bound within an additive error of ϵ. This thereby answers question (1). In Section 7, we answer question (4), where we describe how we obtain our numerical lower bounds for the tight bound and illustrate its results. In Section 8, we explain how quantum multiparameter estimation theory is applicable in learning parameters of Hamiltonian models, and also in 3D-field sensing. In Section 9, we discuss our results and its implications in more detail.

2 Formulation and review of existing results

2.1 Various lower bounds for tight CR bound

We describe the formulation of quantum state estimation, and briefly review existing results on quantum parameter estimation. We refer readers to references [2, 3, 19, 20, 21, 22] for more details.

A *quantum system* is represented by a Hilbert space $\mathcal{H}$. In the following, when $\mathcal{H}$ is finite-dimensional, we can ignore the word "bounded" and "trace class", and can replace "self-adjoint operator" by Hermitian matrix. Let $\mathcal{B}_{sa}(\mathcal{H})$ be the set of bounded self-adjoint operators on $\mathcal{H}$, which is a real vector space. Let $\mathcal{T}_{sa}(\mathcal{H})$ be the real vector space composed of set of trace class self-adjoint operators on $\mathcal{H}$. A *quantum state* ρ is a positive semi-definite matrix on $\mathcal{H}$ with unit trace. The set of all quantum states on $\mathcal{H}$ is denoted by $\mathcal{S}(\mathcal{H}) := \{\rho \,|\, \rho \geq 0, \mathrm{Tr}\rho = 1\} \subset \mathcal{T}_{sa}(\mathcal{H})$.

Symbol	Meaning
d	the number of parameters to be estimated
$\theta = (\theta^1, \ldots, \theta^d)$	the parameters' true value
$\Theta \subseteq \mathbb{R}^d$	set of all possible parameter vectors
$\mathcal{M}$	model $\{\rho_\theta : \theta \in \Theta\}$
$\mathcal{H}$	the Hilbert space for the quantum probe state
n	the number of dimensions $\mathcal{H}$ has
$\mathcal{T}_{sa}(\mathcal{H})$	trace-class self-adjoint operators on $\mathcal{H}$
$\mathcal{B}_{sa}(\mathcal{H})$	bounded self-adjoint operators on $\mathcal{H}$
$\mathcal{S}(\mathcal{H})$	set of density operators on $\mathcal{H}$
$\hat{\theta} = (\hat{\theta}^1, \ldots, \hat{\theta}^d)$	an estimator of θ
ρ_θ, ρ	a probe state that is parametrized by θ
Π	a POVM, a measurement
Π_x	positive semidefinite operators in Π
$\hat{\Pi} = (\Pi, \hat{\theta})$	an estimator based on Π and $\hat{\theta}$
$V_\theta[\hat{\Pi}], V[\hat{\Pi}]$	mean-square error (MSE) matrix of $\hat{\Pi}$
l.u.atθ	locally unbiased at θ
$C_\theta[G], C[G]$	fundamental precision bound: $\min_{\hat{\Pi} \text{ l.u.at } \theta} \mathrm{Tr} G V_\theta[\hat{\Pi}]$
a	size d real square matrix
S	Hermitian operator on $\mathcal{H}$
$D0$	Maximization of $\mathrm{Tr} a + \mathrm{Tr} S$ subject to (23) for all $x \in \mathbb{R}^d$
(a^{**}, S^{**})	the optimal solution of $D0$
$S(D0) = \mathrm{Tr} a^{**} + \mathrm{Tr} S^{**}$	the optimal value of $D0$, (equal to $C[G]$)
$J_\theta(\Pi)$	Fisher information matrix
$\mathcal{X}$	set of labels x corresponding to Π_x
$D_j = \frac{\partial}{\partial \theta^j} \rho_\theta$	jth partial derivative of ρ_θ
$C[G, m]$	denotes $C[G]$ when $|\mathcal{X}| = m$ and $m < \infty$
L_i	SLD corresponding to D_i
$C^S[G]$	SLD CR bound
$C^{HN}[G]$	Holevo-Nagaoka (HN) bound
$C_\theta^N[G], C^N[G]$	Nagaoka bound
$C^{NH}[G]$	Nagaoka-Hayashi (NH) bound

Table 1: Notations for Section 2. We define the quantum parameter estimation problem in terms of a quantum model $\mathcal{M}$ using the notation θ, Θ, $\mathcal{H}$, d and n. The vector θ encapsulates the true value of the d parameters, and we estimate θ using an estimator $\hat{\theta}$ that is obtained from measurements Π. The estimator $\hat{\Pi}$ encapsulates information about both Π and how to construct $\hat{\theta}$ from the measurements. For quantum parameter estimation, we like to minimize the MSE $V_\theta[\hat{\Pi}]$. For uncorrelated measurement strategies the minimum MSE for locally unbiased $\hat{\Pi}$ is given by the optimal value $S(D0)$ of a optimization program $D0$. In fact $S(D0)$ is the tight bound $C[G]$, where G denotes the weight matrix for the d-parameter estimation problem. Evaluating $C[G]$ is non-trivial. There is a plethora of upper bounds to the tight bound $C[G]$, given by the SLD CR bound, $C^S[G]$, the Holevo-Nagaoka (HN) bound $C^{HN}[G]$, the Nagaoka bound $C^N[G]$ and the Nagaoka-Hayashi (NH) bound $C^{HN}[G]$.

A *measurement* Π can be represented mathematically as a positive operator-valued measurement (POVM) which is a set of positive semidefinite matrices $\Pi = \{\Pi_x\}_{x \in \mathcal{X}}$ that satisfies a completeness condition. When the set of measurement outcomes is finite, $\mathcal{X}$ is a finite set and the completeness condition is $\sum_{x \in \mathcal{X}} \Pi_x = I$. When the set of measurement outcomes is a continuum, such as when $\mathcal{X} = \mathbb{R}$, the completeness condition is $\int_{\mathcal{X}} \Pi_x dx = I$. For notational simplicity, for any POVM Π, we employ the notation for discrete-valued POVMs, and we always use $\sum_{x \in \mathcal{X}} \Pi_x = I$ to represent the completeness condition.

When one performs a POVM Π on ρ, the *Born rule*

$$p_\rho(x|\Pi) = \mathrm{Tr}\Big[\rho \Pi_x\Big] \tag{1}$$

gives the probability of getting an outcome $x \in \mathcal{X}$. We are interested in a model of quantum parameter estimation given by a parametric family of quantum states on $\mathcal{H}$:

$$\mathcal{M} := \{\rho_\theta \,|\, \theta \in \Theta\} \subset \mathcal{S}(\mathcal{H}), \tag{2}$$

where $\Theta \subset \mathbb{R}^d$ denotes the set of parameters. Here, d denotes the number of parameters that we want to simultaneously estimate. To avoid mathematical subtleties, we impose regularity conditions; we require ρ_θ to be differentiable sufficiently many times, assume $\partial \rho_\theta / \partial \theta_i$ to be linearly independent, and for the sake of clarity only consider full-rank states here [2].

Now given a measurement Π and an estimator $\hat{\boldsymbol{\theta}}$, we denote $\hat{\Pi} = (\Pi, \hat{\boldsymbol{\theta}})$ as an *estimator*. We define the mean-square error (MSE) matrix for the estimator $\hat{\Pi}$ as

$$
\begin{aligned}
V_{\boldsymbol{\theta}}[\hat{\Pi}] &= \left[\sum_{x \in \mathcal{X}} \mathrm{Tr}\Big[\rho_{\boldsymbol{\theta}} \Pi_x\Big] (\hat{\theta}^i(x) - \theta^i)(\hat{\theta}^j(x) - \theta^j)\right] \\
&= \left[E_{\boldsymbol{\theta}}\Big[(\hat{\theta}^i(x) - \theta^i)(\hat{\theta}^j(x) - \theta^j)|\Pi\Big]\right].
\end{aligned}
\tag{3}
$$

where $E_{\boldsymbol{\theta}}[f(X)|\Pi]$ denotes the expectation of a random variable $f(X)$ with respect to the probability distribution $p_{\rho_{\boldsymbol{\theta}}}(x|\Pi) = \mathrm{Tr}\rho_{\boldsymbol{\theta}} \Pi_x$ obtained from the Born rule. In multiparameter quantum metrology, the objective is to find an optimal estimator $\hat{\Pi} = (\Pi, \hat{\boldsymbol{\theta}})$ that in some sense minimizes the MSE matrix.

Since the minimization of an MSE matrix is not properly defined, we seek precision bounds where we minimize $\mathrm{Tr}[G V_{\boldsymbol{\theta}}[\hat{\Pi}]]$, which is the weighted trace of the MSE matrix according to a given positive matrix G. Here, G is a *weight matrix* and quantifies the trade-off between estimating different vector components of the parameter $\boldsymbol{\theta}$. For instance, when G is the size d identity matrix I_d, minimizing $\mathrm{Tr} G V_{\boldsymbol{\theta}}[\hat{\Pi}]$ corresponds to minimizing the average variance of estimators.

A problem fundamental to quantum metrology is that of finding the ultimate precision bound under reasonable assumptions on the estimators we use. We say that an estimator $\hat{\Pi}$ is *unbiased* if for all $\boldsymbol{\theta} = (\theta^1, \ldots, \theta^d) \in \Theta$, we have

$$E_{\boldsymbol{\theta}}\Big[\hat{\theta}^i(X)|\Pi\Big] = \sum_{x \in \mathcal{X}} \hat{\theta}^i(x) \mathrm{Tr}\Big[\rho_{\boldsymbol{\theta}} \Pi_x\Big] = \theta^i \quad (\forall i = 1, 2, \ldots, d).$$

However, such an unbiased estimator invariably does not exist. Hence one often relaxes this unbiasedness condition to a *locally unbiased* condition in the neighborhood of a chosen

---

[2]We refer the reader to [23, 24] for problems in the pure-state model.

point $\boldsymbol{\theta}$. To have the locally unbiased condition to hold, we require that for all parameter indices $i, j \in \{1, 2, \ldots, d\}$, we have the equations

$$E_{\boldsymbol{\theta}}\left[\hat{\theta}^i(X)|\Pi\right] = \sum_{x \in \mathcal{X}} \hat{\theta}^i(x)\mathrm{Tr}\left[\rho_{\boldsymbol{\theta}}\Pi_x\right] = \theta^i, \tag{4}$$

$$\frac{\partial}{\partial \theta^j} E_{\boldsymbol{\theta}}\left[\hat{\theta}^i(X)|\Pi\right] = \sum_{x \in \mathcal{X}} \hat{\theta}^i(x)\mathrm{Tr}\left[\frac{\partial}{\partial \theta^j}\rho_{\boldsymbol{\theta}}\Pi_x\right] = \delta_i^j. \tag{5}$$

Note that we can derive this condition by applying the Taylor expansion to the usual unbiasedness condition at a point $\boldsymbol{\theta}$ to first order.

Then, we introduce the fundamental precision limit by

$$C_{\boldsymbol{\theta}}[G] := \min_{\hat{\Pi}\,:\,\mathrm{l.u.at}\,\boldsymbol{\theta}} \mathrm{Tr}\left[GV_{\boldsymbol{\theta}}[\hat{\Pi}]\right], \tag{6}$$

where the minimization is carried out for all possible estimators under the locally unbiasedness condition, which is indicated by l.u. at $\boldsymbol{\theta}$. In this paper, any lower bound for the weighted trace of the MSE matrix $V_{\boldsymbol{\theta}}[\hat{\Pi}]$ is referred to as the *CR type bound*. When a CR type bound equals to the fundamental precision limit $C_{\boldsymbol{\theta}}[G]$ as in (6), it is called the *tight CR bound* in our discussion. That is, $C_{\boldsymbol{\theta}}[G]$ is called the tight CR bound. In the following, we discuss some CR type and tight CR bounds.

In fact, the set of MSE matrix $V_{\boldsymbol{\theta}}[\hat{\Pi}]$ under the local unbiasedness condition is characterized as follows.

$$\{V_{\boldsymbol{\theta}}[\hat{\Pi}]|\hat{\Pi} \text{ is l.u.at } \boldsymbol{\theta}\} = \{J_{\boldsymbol{\theta}}(\Pi)^{-1}|\Pi \text{ is a POVM }\}, \tag{7}$$

where $J_{\boldsymbol{\theta}}(\Pi)$ is the Fisher information matrix of the distribution family $\{P_{\boldsymbol{\theta},\Pi}\}_{\boldsymbol{\theta}}$ and $P_{\boldsymbol{\theta},\Pi}(x) := \mathrm{Tr}[\rho_{\boldsymbol{\theta}}\Pi_x]$ [25, Exercise 6.44].

When multiple copies of the unknown state are prepared, only individual measurements for each copy are allowed, and the error is measured by weighted sum of mean square error with the weight matrix G, it is impossible to realize estimation precisions exceeding $C_{\boldsymbol{\theta}}[G]$. Although the measurement to achieve this bound depends on the true parameter $\boldsymbol{\theta}$, when classical adaptive improvement for the choice of measurement is allowed, it is possible to achieve the bound $C_{\boldsymbol{\theta}}[G]$ [26, 27, 28, 29]. Therefore, we can consider that the tight CR bound $C_{\boldsymbol{\theta}}[G]$ expresses the ultimate precision of the optimal estimator in the asymptotic limit of infinitely many probe states when only adaptive individual measurements for each copy are allowed. This setting corresponds to the strategy A2 in Section 3.2 of [22]. Furthermore, when n copies of the unknown state are given, even when any separable measurement over the n-fold system is allowed, it is impossible to overcome the precision error $C_{\boldsymbol{\theta}}[G]$ although there exists a separable measurement that requires quantum correlation over the n-fold system [25, Exercise 6.42].

In the latter discussion, we focus on this problem only at one point $\boldsymbol{\theta} \in \Theta$. Hence, we omit the subscript θ later, using $C[G]$ to denote $C_{\boldsymbol{\theta}}[G]$. Furthermore, we simplify ρ_θ, $\frac{\partial}{\partial \theta^j}\rho_{\boldsymbol{\theta}}$, $V_{\boldsymbol{\theta}}[\hat{\Pi}]$ to ρ, D_j, and $V[\hat{\Pi}]$, respectively. Further, when $\mathcal{H}$ is infinite-dimensional, ρ and D_j are assumed to be trace-class operators. Note that, even when we remove (4), the minimum value of $\mathrm{Tr}[GV[\hat{\Pi}]]$ is not changed due to the following reason. Given an estimator $\hat{\Pi} = (\Pi, \hat{\boldsymbol{\theta}})$ satisfying (5), the new estimator $\hat{\Pi}' = (\Pi, \hat{\boldsymbol{\theta}} - E_{\boldsymbol{\theta}}\left[\hat{\theta}(X)|\Pi\right] + \boldsymbol{\theta})$ satisfies the relation $V[\hat{\Pi}] = V[\hat{\Pi}'] + (E_{\boldsymbol{\theta}}\left[\hat{\theta}(X)|\Pi\right] - \boldsymbol{\theta})(E_{\boldsymbol{\theta}}\left[\hat{\theta}(X)|\Pi\right] - \boldsymbol{\theta})^T \geq V[\hat{\Pi}']$. Hence, we ignore the condition (4).

We denote the minimum of (6) when our measurement is limited to measurement with discrete value and the number of elements in $\mathcal{X}$ is m by $C[G, m]$. Clearly, $C[G, m] \geq$

$C[G, m']$ for $m' \geq m$. In addition, since our interest is the minimization (6), without loss of generality, we can assume that $\boldsymbol{\theta}$ is zero.

To get a CR type bound, we often focus on the SLD L_i, which is defined as any Hermitian matrix that satisfies

$$D_i = \frac{1}{2}\Big(L_i\rho + \rho L_i\Big). \tag{8}$$

The SLD Fisher information matrix $J = (J_{i,j})$ is defined as

$$J_{i,j} := \frac{1}{2}\mathrm{Tr}\Big[L_i\Big(L_j\rho + \rho L_j\Big)\Big]. \tag{9}$$

Here, when ρ is strictly positive, the choice of Hermitian matrix L_i is unique. Otherwise, it is not unique. However, the definition of the SLD Fisher information matrix J in (9) does not depend on the choice of Hermitian matrix L_i under the condition (8). Under the locally unbiasedness condition at $\boldsymbol{\theta}$, we have SLD CR inequality [2]

$$V[\hat{\Pi}] \geq J^{-1}. \tag{10}$$

For the proof, see [2, 3], [25, Section 6.6], [22, Appendix B] for more details. When we can choose SLDs L_i for $i = 1, \ldots, d$ such that these SLDs L_i all commute, the equality in (10) can be achieved by a local unbiased estimator constructed by their simultaneous spectral decomposition. In the choice of SLDs L_i, extending the Hilbert space is allowed. However, when ρ is a strictly positive density matrix, it is sufficient to check for the commutativity of SLDs L_i without extending the Hilbert space. In general, there is a possibility that the equality in (10) be achieved only with an extending Hilbert space.

Taking a weighted trace in (10), we obtain the following bound.

- The SLD CR bound, which is the tight CR bound for any one-parameter model [2]:

$$C^{\mathrm{S}}[G] := \mathrm{Tr}[GJ^{-1}], \tag{11}$$

where J denotes the SLD Fisher information matrix about the model $\mathcal{M}$.

To get a tighter bound than SLD bound, we focus on the vector of self-adjoint operators $\vec{Z} = (Z^1, \ldots, Z^d)$ that satisfies the condition

$$\mathrm{Tr}[D_j Z^i] = \delta_i^j \text{ for } i, j = 1, \ldots, d. \tag{12}$$

Then, we define the Hermitian matrix $\mathcal{Z}(\vec{Z})$ whose (i, j) component is $\mathrm{Tr}\rho Z^i Z^j$. When an estimator $\hat{\Pi} = (\Pi, \hat{\boldsymbol{\theta}})$ satisfies the condition:

$$\sum_{x \in \mathcal{X}} \hat{\theta}^i(x)\Pi_x = Z^i, \tag{13}$$

we have [3, (6.7.73)]

$$V[\hat{\Pi}] \geq \mathcal{Z}(\vec{Z}). \tag{14}$$

Using this relation, we have [4]

$$\mathrm{Tr}GV_{\boldsymbol{\theta}}[\hat{\Pi}] \geq \mathrm{Tr}G\mathrm{Re}\,\mathcal{Z}(\vec{Z}) + \mathrm{Tr}|G^{\frac{1}{2}}\mathrm{Im}\,\mathcal{Z}(\vec{Z})G^{\frac{1}{2}}|, \tag{15}$$

where the operator $|X|$ is defined as $\sqrt{X^\dagger X}$.

Therefore, we obtain the following bound;

- The Holevo-Nagaoka (HN) bound [3, 4] (Nagaoka [4] derived this bound by using several formulas obtained in Holevo [3], and called this bound the Holevo bound. Recently, the reference [30] called this bound the Holevo Cramer Rao bound, and cited Nagaoka [4] as its formulation. However, this bound does not appear in Holevo [3]. It should be called Holevo-Nagaoka CR bound. For its detailed history, see the latest review [31].)

$$C^{HN}[G] := \min_{\vec{Z}=(Z^1,\dots,Z^d)} \text{Tr}[G\,\text{Re}\,\mathcal{Z}(\vec{Z})] + \text{Tr}[|G^{\frac{1}{2}}\text{Im}\,\mathcal{Z}(\vec{Z})G^{\frac{1}{2}}|], \tag{16}$$

where the minimization takes the vector of self-adjoint operators $\vec{Z} = (Z^1, \dots, Z^d)$ to satisfy the condition (12).

Notice that the minimum (16) is achieved when the vector of Hermitian matrices $\vec{Z}$ satisfies the condition $\text{Tr}\rho Z^i = 0$ for $i = 1, \dots, d$. When the model is composed of pure states, the equality in inequality $C[G] \geq C^{HN}[G]$ holds [32].

The HN bound $C^{HN}[G]$ improves the SLD CR bound $C^S[G]$, and gives the asymptotic limit of precision of the minimum estimation error when any quantum correlation is allowed in measurement apparatus and multiple copies of unknown states are prepared [33, 5, 34, 35, 29].

In addition, the HN bound $C^{HN}[G]$ satisfies the additivity condition, i.e., this value with the m-copy setting equals the value with the one-copy setting divided by m [5]. The reference [8, Eq. (11)] derived a calculation formula based on SDP for the bound $C^{HN}(G)$.

When $d = 2$, the lower bound in (15) is written as

$$\begin{aligned}
&\text{Tr}[G\,\text{Re}\,\mathcal{Z}(\vec{Z})] + \text{Tr}[|G^{\frac{1}{2}}\text{Im}\,\mathcal{Z}(\vec{Z})G^{\frac{1}{2}}|]\\
=&G_{1,1}\text{Tr}[Z^1\rho Z^1] + G_{2,2}\text{Tr}[Z^2\rho Z^2]\\
&+ G_{1,2}\text{Tr}[\rho(Z^1Z^2 + Z^2Z^1)]\\
&+ 2\sqrt{\det G}|\text{Tr}[\rho[Z^1, Z^2]]|.
\end{aligned} \tag{17}$$

To improve the HN bound, Nagaoka [4, 10] derived the following inequality with $d = 2$:

$$\begin{aligned}
\text{Tr}[GV_{\boldsymbol{\theta}}[\hat{\Pi}]] \geq &G_{1,1}\text{Tr}[Z^1\rho Z^1] + G_{2,2}\text{Tr}[Z^2\rho Z^2]\\
&+ G_{1,2}\text{Tr}[\rho(Z^1Z^2 + Z^2Z^1)]\\
&+ 2\sqrt{\det G}\text{Tr}[|\rho^{1/2}[Z^1, Z^2]\rho^{1/2}|]
\end{aligned} \tag{18}$$

when an estimator $\hat{\Pi} = (\Pi, \hat{\boldsymbol{\theta}})$ satisfies the condition (13). Since $\text{Tr}[|\rho^{1/2}[Z^1, Z^2]\rho^{1/2}|] \geq |\text{Tr}[\rho[Z^1, Z^2]]|$, the inequality (17) implies (15).

Using (17), we obtain the following bound.

- The Nagaoka bound [4, 10], which is given only in the case with $d = 2$, and is tighter than the HN bound.

$$\begin{aligned}
C_{\boldsymbol{\theta}}^N[G] := \min_{\vec{Z}=(Z^1,Z^2)} &G_{1,1}\text{Tr}[Z^1\rho Z^1] + G_{2,2}\text{Tr}[Z^2\rho Z^2]\\
&+ G_{1,2}\text{Tr}[\rho(Z^1Z^2 + Z^2Z^1)]\\
&+ 2\sqrt{\det G}\text{Tr}[|\rho^{1/2}[Z^1, Z^2]\rho^{1/2}|],
\end{aligned} \tag{19}$$

where the minimization takes the vector of self-adjoint operators $\vec{Z} = (Z^1, Z^2)$ under the condition (12).

In the qubit case, the Nagaoka bound $C_{\boldsymbol{\theta}}^N[G]$ equals the tight CR bound $C[G]$ for a two-parameter model ($d = 2$).

To generalize the above improvement, we consider the tensor product between the Hilbert space $\mathcal{H}$ and the real vector space $\mathcal{R}'$ spanned by $\{|1\rangle, \ldots, |d\rangle\}$. Then, we define $\mathcal{B}'$ as

$$\mathcal{B}' := \left\{ \sum_{j=1}^{d} \sum_{k=1}^{d} \Big| |k\rangle\langle j| \otimes X^{k,j} | X^{k,j} \in \mathcal{B}_{sa}(\mathcal{H}), X^{k,j} = X^{j,k} \right\}. \tag{20}$$

To generalize (17), for the vector of self-adjoint operators $\vec{Z} = (Z^1, \ldots, Z^d)$, the paper [11] focuses on $\Pi(\vec{Z}) \in \mathcal{B}'$ with components $\Pi(\vec{Z})^{i,j} = Z^i Z^j$. When an estimator $\hat{\Pi} = (\Pi, \hat{\boldsymbol{\theta}})$ satisfies the condition (13), the paper [11] derived the inequality

$$\mathrm{Tr}[GV_{\boldsymbol{\theta}}[\hat{\Pi}]] \geq \min_{X' \in \mathcal{B}'} \{\mathrm{Tr}[(G \otimes \rho)X'] | X' \geq \Pi(\vec{Z})\}. \tag{21}$$

Using (17), the paper [12] introduced the following bound.

• The Nagaoka-Hayashi (NH) bound [12], which is tighter than the HN bound.

$$C^{NH}[G] := \min_{\vec{Z}} \min_{X' \in \mathcal{B}'} \{\mathrm{Tr}[(G \otimes \rho)X'] | X' \geq \Pi(\vec{Z})\}, \tag{22}$$

where the minimization takes the vector of self-adjoint operators $\vec{Z} = (Z^1, \ldots, Z^d)$ under the condition (12).

The reference [12, (22)] derived a calculation formula based on SDP for the bound $C^{NH}(G)$. Overall, the HN and NH bounds focused on the relation between the difficulty of joint measurement and the multiparameter estimation. This kind of relation was also discussed in the recent paper [36].

2.2 Another expression of tight CR bound by [37, 13]

To get another form of $C[G]$, the papers [37, 13] treated the minimization $C[G]$ as a minimization with respect to a POVM $\hat{\Pi}$ over $\mathbb{R}^d$. In this case, a POVM $\hat{\Pi}$ is considered as an element of convex cone. Then, it focused on the following maximization problem.

For a $d \times d$ real matrix a and a self-adjoint operator S on the Hilbert space $\mathcal{H}$, we consider the condition:

$$(x^T G x)\rho - \sum_{i,j} a_i^j x^i D_j - S \geq 0 \tag{23}$$

for $x = (x^1, \ldots, x^d) \in \mathbb{R}^d$. We consider the maximization:

$$S(D0) := \max_{a,S} \sum_i a_i^i + \mathrm{Tr}S, \tag{24}$$

where the maximization takes the pair (a, S) to satisfy the condition (23). Here, $D0$ shows the maximization problem presented in (24), and the maximum value is denoted by $S(D0)$. Then, the paper [13] showed the following proposition:

Proposition 1 ([13, Theorem 6]).

$$C[G] = S(D0). \tag{25}$$

In fact, the maximization (24) can be regarded as the dual problem when the minimization $C[G]$ is considered as a conic linear programming with respect to a POVM $\hat{\Pi}$ over $\mathbb{R}^d$. Hence, we call the maximization (24) the dual problem $D0$.

To see this, we focus on the relation

$$\int_{\mathbb{R}^d} \mathrm{Tr}\Big((x^T G x)\rho - \sum_{i,j} a_i^j x^i D_j - S)\Pi(dx) \Big) \geq 0, \tag{26}$$

where $x = (x^1, \ldots, x^d)$ takes values in $\mathbb{R}^d$. Using the above relation, we have

$$\int_{\mathbb{R}^d} (x^T G x)\mathrm{Tr}[\rho\Pi(dx)]$$
$$\geq \int_{\mathbb{R}^d} \sum_{i,j} a_i^j x^i \mathrm{Tr}[D_j \Pi(dx)] + \int_{\mathbb{R}^d} \mathrm{Tr}[S\Pi(dx)]$$
$$= \sum_{i,j} a_i^j \int_{\mathbb{R}^d} x^i \mathrm{Tr}[D_j \Pi(dx)] + \mathrm{Tr}[S\Pi(\mathbb{R}^d)]$$
$$= \sum_{i,j} a_i^j \delta_j^i + \mathrm{Tr}[SI] = \sum_i a_i^i + \mathrm{Tr}S. \tag{27}$$

Hence, we can easily see the inequality $\geq$ in (25).

To show the opposite inequality, one needs to show the non-existence of the duality gap, i.e., the minimum of the primal problem equals the maximum of the dual problem in the framework of a conic linear programming, as explained in Appendix B. Although the non-existence of the duality gap holds for a conic linear program in a finite-dimensional vector space, the set of POVMs $\hat{\Pi}$ over $\mathbb{R}^d$ is a subset of an infinite-dimensional space, because the number of elements in $\mathbb{R}^d$ is not finite, even when $\mathcal{H}$ is finite-dimensional. Hence, the paper [13] showed the non-existence of the duality gap in this problem setting by discussing a complicated issue related to topological vector space.

In fact, the formula (25) has various merits. Appendix A summarizes its two applications. For example, use of the relation (25) enables us to derive the minimum MSE under the locally unbiasedness condition under the one-parameter case. As another example, using the relation (25), the papers [37, 13] derived the tight CR bound $C[G]$ for a three-parameter model ($d = 3$) in the qubit system.

3 Conic programming with various cones on tensor product

When $\mathcal{H}$ is finite-dimensional, Proposition 1 guarantees that the tight CR bound $C[G]$ is given as the maximum value $S(D0)$ with the variables (a, S) of finite-dimension. However, even with a finite-dimensional space $\mathcal{H}$, the calculation of the maximum value $S(D0)$ is not so easy,

To get a more computable form for $C[G]$, we introduce the real vector space $\mathcal{R}$ spanned by $\{|0\rangle, |1\rangle, \ldots, |d\rangle\}$. Let $\mathcal{M}_{rs}(\mathcal{R})$ be the set of real symmetric matrices on $\mathcal{R}$. The weight matrix G can be considered as an element of $\mathcal{M}_{rs,+}(\mathcal{R}) \subset \mathcal{M}_{rs}(\mathcal{R})$, where $\mathcal{M}_{rs,+}(\mathcal{R})$ denotes the set of positive semidefinite matrices in $\mathcal{M}_{rs}(\mathcal{R})$. Here, we denote $G = \sum_{1 \leq j, i \leq d} G_{i,j} |i\rangle\langle j|$.

This section aims to derive various CR bounds as conic linear programming problems on the vector spaces $\mathcal{B} := \mathcal{M}_{rs}(\mathcal{R}) \otimes \mathcal{B}_{sa}(\mathcal{H})$ and $\mathcal{T} := \mathcal{M}_{rs}(\mathcal{R}) \otimes \mathcal{T}_{sa}(\mathcal{H})$. That is, $\mathcal{B}$ is defined as

$$\mathcal{B} := \Big\{ \sum_{j=0}^d \sum_{k=0}^d |k\rangle\langle j| \otimes X^{k,j} \Big| X^{k,j} \in \mathcal{B}_{sa}(\mathcal{H}), X^{k,j} = X^{j,k} \Big\}. \tag{28}$$

Symbol	Meaning
$\mathcal{R}$	real vector space spanned by $|0\rangle, |1\rangle, \ldots, |d\rangle$
$\mathcal{R}_C$	complex vector space spanned by $|0\rangle, \ldots, |d\rangle$
$\mathcal{R}'$	real vector space spanned by $|1\rangle, \ldots, |d\rangle$
$\mathcal{R}'_C$	complex vector space spanned by $|1\rangle, \ldots, |d\rangle$
$\mathcal{M}_{rs}(\mathcal{R})$	set of real symmetric matrices on $\mathcal{R}$
$\mathcal{M}_{rs,+}(\mathcal{R})$	positive semidefinite matrices in $\mathcal{M}_{rs}(\mathcal{R})$
$\mathcal{B}$	Equal to $\mathcal{M}_{rs}(\mathcal{R}) \otimes \mathcal{B}_{sa}(\mathcal{H})$
$\mathcal{T}$	Equal to $\mathcal{M}_{rs}(\mathcal{R}) \otimes \mathcal{T}_{sa}(\mathcal{H})$
$\mathcal{B}', \mathcal{B}''$	both are extensions of $\mathcal{B}$
$\mathcal{B}''_\rho$	ρ-dependent subspace of $\mathcal{B}''$
$\mathcal{S}_P, \mathcal{S}_{SEP}$	cones in $\mathcal{B}$
$X = \Pi(a, S)$	an operator in $\mathcal{T}$ given by (57)
$\mathcal{S}_P^*, \mathcal{S}_{SEP}^*$	dual cones
$P1, P2$	conic programs with cones $\mathcal{S}_{SEP}, \mathcal{S}_P$
$\mathcal{S}_3 = \mathcal{S}(\mathcal{R}_C \otimes H)_{PPT}$	PPT cone
$\mathcal{S}_4 = \mathcal{S}(\mathcal{R}_C \otimes H)_P$	cone $S(\mathcal{R}_C \otimes H)_P$
$Pi, \quad i = 3, 4$	conic program with variable in $\mathcal{B}''$ and $\mathcal{S}_i$
$P5$	conic program with variable in $\mathcal{B}''_\rho$ and $\mathcal{S}_4$
$S(Pi)$	optimal value of Pi, for $i = 1, 2, 3, 4, 5$
Di	dual program of Pi for $i = 1, 2, 3, 4, 5$
Objective function of Di	$\mathrm{Tr}\,a + \mathrm{Tr}\,S$
Constraints of Di	$\Pi(a, S)$ lies in appropriate dual cone
$S(Di)$	optimal value of Di, for $i = 0, 1, 2, 3, 4, 5$
$\mathcal{W}_R$	a finite set of norm 1 vectors in $\mathcal{R}$
$\mathcal{W}_H$	a finite set of norm 1 vectors in $\mathcal{H}$
$[P1, \mathcal{W}_R]$	Primal SDP
$[D1, \mathcal{W}_R]$	Dual SDP of $[P1, \mathcal{W}_R]$
$S[P1, \mathcal{W}_R]$	optimal value of $[P1, \mathcal{W}_R]$
$S[D1, \mathcal{W}_R]$	optimal value of $[D1, \mathcal{W}_R]$
$C_2(a)$	real number depending on a, D_j and ρ (see (80))
C_1	real number depending on D_j and ρ (see (110))
$\delta(\mathcal{W}_R)$	covering radius of $\mathcal{W}_R$
κ	related to the covering radius of $\mathcal{W}_R$

Table 2: Notations from Section 3 onwards. For our conic programming framework, we consider primal conic programs $P1, P2, P3, P4, P5$ and their dual conic programs $D1, D2, D3, D4, D5$. Primal programs are minimizations over cones, while dual programs are maximizations over the dual cones (also cones). There is no duality gap between the primal programs and the dual programs, so Pi has the same optimal value as Di, so that $S(Pi) = S(Di)$. The optimizations for the primal programs are over variables of the form X, which are matrices in the tensor product space $\mathcal{R}_C \otimes \mathcal{H}$. The primal conic programs all have objective functions $\mathrm{Tr}(G \otimes \rho)X$ and the dual conic programs all have objective functions $\mathrm{Tr}\,a + \mathrm{Tr}\,S$, where a is a size d real matrix, and S is a Hermitian matrix. Now $S(P1)$ and $D(P1)$ are equal to the tight bound $C[G]$, and can be approximated using the semidefinite program (SDP) $[P1, \mathcal{W}_R]$, where we can take $\mathcal{W}_R$ to be a collection of points on a norm-1 hypersphere in a real vector space. Namely $C[G] = S(P1) = D(P1) \approx S[P1, \mathcal{W}_R]$. The smaller the covering radius $\delta(\mathcal{W}_R)$, the better $S[P1, \mathcal{W}_R]$ approximates $C[G]$. To approximate $C[G]$, for a chosen set $\mathcal{W}_R$, we solve the primal SDP $[P1, \mathcal{W}_R]$ and the dual SDP $[P1, \mathcal{W}_R]$, denoting the optimal solutions as X^* and (a^*, S^*) respectively. Using (a^*, S^*) we can calculate $C_2(a^*)$. Then we find that (in Theorem 13) $S[D1, \mathcal{W}_R] - n\|X^*\|(1 + C_2(a^*)^2)\delta(\mathcal{W}_R) \leq C[G] \leq S[D1, \mathcal{W}_R]$.

$\mathcal{T}$ is similarly defined.

Since any real symmetric matrix can be regarded as a Hermitian matrix on the complexified space $\mathcal{R}_C$ of $\mathcal{R}$, any element of $\mathcal{B}$ can be considered as a self-adjoint operator on $\mathcal{R}_C \otimes \mathcal{H}$. Then, we define several cones in the space $\mathcal{B}$ as follows. Let $\mathcal{B}_{sa,+}(\mathcal{H})$ be the set of bounded positive semi-definite operators on $\mathcal{H}$. Then, we define the cone $\mathcal{S}_{SEP}$ in $\mathcal{B}$ as $\mathcal{S}_{SEP} := conv(\mathcal{M}_{rs,+}(\mathcal{R}) \otimes \mathcal{B}_{sa,+}(\mathcal{H}))$, where $conv$ is the convex hull. Then, we define the cone $\mathcal{S}_{SEP}^*$ in $\mathcal{B}$ as the dual cone of $\mathcal{S}_{SEP}$. That is,

$$\mathcal{S}_{SEP}^* := \{X \in \mathcal{T} | \mathrm{Tr} XY \geq 0 \text{ for } Y \in \mathcal{S}_{SEP}\}. \tag{29}$$

In the context of entanglement theory, $\mathcal{S}_{SEP}^*$ is often called entanglement witnesses [38, 39, 40].

Also, we define the cone $\mathcal{S}_P$ as

$$\mathcal{S}_P := \{X \in \mathcal{B} | \langle v|X|v\rangle \geq 0 \text{ for } v \in \mathcal{R}_C \otimes \mathcal{H}\}. \tag{30}$$

Then, we have

$$\begin{aligned} \mathcal{S}_P^* &:= \{X \in \mathcal{T} | \mathrm{Tr} XY \geq 0 \text{ for } Y \in \mathcal{S}_P\} \\ &\supset \{X \in \mathcal{T} | \mathrm{Tr}[X|v\rangle\langle v|] \geq 0 \text{ for } v \in \mathcal{R}_C \otimes \mathcal{H}\} \\ &= \{X \in \mathcal{T} | \langle v|X|v\rangle \geq 0 \text{ for } v \in \mathcal{R}_C \otimes \mathcal{H}\}. \end{aligned} \tag{31}$$

Hence, when $\mathcal{H}$ is finite-dimensional, we have

$$\mathcal{S}_P \subset \mathcal{S}_P^*. \tag{32}$$

Relaxing the condition of $\mathcal{B}$ we extend the space $\mathcal{B}$ as

$$\mathcal{B}'' := \left\{ \sum_{j=0}^d \sum_{k=0}^d |k\rangle\langle j| \otimes X^{k,j} \Big| X^{k,0} \in \mathcal{B}_{sa}(\mathcal{H}), X^{k,j} = (X^{j,k})^\dagger \right\}. \tag{33}$$

We define the set $\mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P$ ($\mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_{PPT}$) of positive semi-definite (positive partial transpose) self-adjoint operators on $\mathcal{R}_C \otimes \mathcal{H}$. Since any component of an element X of $\mathcal{S}_P$ is a self-adjoint operator, the element X satisfies the conditions of $\mathcal{B}''$. Also, the operator transposed on the system $\mathcal{R}_C$ of X is positive. Hence, the element X satisfies the conditions of $\mathcal{B}''$ and $\mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_{PPT}$. Therefore, since the set $\mathcal{S}_{SEP}$ is generated by separable pure states in $\mathcal{R}_C \otimes \mathcal{H}$, we find the relations

$$\mathcal{S}_{SEP} \subset \mathcal{S}_P \subset \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_{PPT} \cap \mathcal{B}'' \subset \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}''. \tag{34}$$

To derive various CR bounds as conic linear programming problems on the vector spaces, we discuss the relation between our estimator and $\mathcal{S}_{SEP}$. For any estimator $(\Pi, \hat{\theta}) := (\{\Pi_x\}_{x \in \mathcal{X}}, \hat{\theta}(x))$, we define the operator $X(\Pi, \hat{\theta}) \in \mathcal{S}_{SEP}$ as

$$\begin{aligned} &X(\Pi, \hat{\theta}) \\ &:= \sum_{x \in \mathcal{X}} \left(|0\rangle + \sum_{i=1}^d \hat{\theta}^i(x)|i\rangle \right) \left(\langle 0| + \sum_{i=1}^d \langle i|\hat{\theta}^i(x) \right) \otimes \Pi_x. \end{aligned} \tag{35}$$

Although the references [41, (2011)] considered a matrix with the separable form from a POVM, our matrix $X(\Pi, \hat{\theta})$ with the separable form is different from their matrix with the separable form in the following point. Their matrix with the separable form is composed

of the tensor product of the POVM element and the density matrix of their guess. We consider the tensor product of the POVM element and the one-dimensional operator given by the superposition of $|0\rangle$ and the parameter of our estimate. The component $|0\rangle\langle0|$ enables us to check the condition for the POVM.

Since Π satisfies the condition of POVM, $X(\Pi, \hat{\theta})$ satisfies the following condition:

$$\text{Tr}_{\mathcal{R}}\Big[(|0\rangle\langle0| \otimes I_{\mathcal{H}})X(\Pi, \hat{\theta})\Big] = I_{\mathcal{H}}. \tag{36}$$

The condition (5) for a locally unbiased estimator guarantees the following condition:

$$\text{Tr}\Big[(\frac{1}{2}(|0\rangle\langle i| + |i\rangle\langle0|) \otimes D_j)X(\Pi, \hat{\theta})\Big] = \delta_{i,j}. \tag{37}$$

Under the above condition, the objective function $\text{Tr}GV_{\boldsymbol{\theta}}[\hat{\Pi}]$ is rewritten with $X(\Pi, \hat{\theta})$ as

$$\text{Tr}\Big[(G \otimes \rho)X(\Pi, \hat{\theta})\Big]. \tag{38}$$

In this notation, the components $G_{0,0}, G_{0,j}$, and $G_{j,0}$ are zero. Therefore, we consider the minimization:

$$S(P1) := \min_{X \in \mathcal{S}_{SEP}} \Big\{\text{Tr}\Big[(G \otimes \rho)X\Big]\Big|(36),(37) \text{ hold.}\Big\}, \tag{39}$$

and we denote the above primal conic linear programming problem by $P1$, and the minimum value is denoted by $S(P1)$. Although the above discussion shows the inequality $C[G] \geq S(P1)$, we have the following theorem.

Theorem 2. *We have*

$$C[G] = S(P1). \tag{40}$$

The above theorem shows the tight CR bound can be calculated by solving the primal conic linear programming $P1$. Since the tight CR bound equals the tight bound, the primal conic linear programming $P1$ characterizes the tight bound.

Proof. To show the opposite inequality $C[G] \leq S(P1)$, we assume that an element $X \in \mathcal{S}_{SEP}$ satisfies the conditions (36) and (37). X is written as

$$X = \sum_{x \in \mathcal{X}} T(x) \otimes M_x = \sum_{x \in \mathcal{X}_1} T(x) \otimes M_x + \sum_{x \in \mathcal{X}_0} T(x) \otimes M_x, \tag{41}$$

where $\mathcal{X}_0 := \{x \in \mathcal{X}|\langle0|T(x)|0\rangle = 0\}$ and $\mathcal{X}_1 := \mathcal{X} \setminus \mathcal{X}_0$. Then, for $x \in \mathcal{X}_0$, $T(x)|0\rangle = 0$ and $\langle0|T(x) = 0$. Hence, the second term in (41) does not contribute the conditions (36) and (37). Thus, the first term in (41) satisfies the conditions (36) and (37). Since

$$\text{Tr}\Big[(G \otimes \rho) \sum_{x \in \mathcal{X}_1} T(x) \otimes M_x\Big]$$
$$\leq \text{Tr}\Big[(G \otimes \rho)\Big(\sum_{x \in \mathcal{X}_1} T(x) \otimes M_x + \sum_{x \in \mathcal{X}_0} T(x) \otimes M_x\Big)\Big]$$
$$= \text{Tr}\Big[(G \otimes \rho)X\Big], \tag{42}$$

we have (40). $\qquad\square$

When $\mathcal{H}$ is finite-dimensional and $\dim \mathcal{H} = n$, the dimension of $\mathcal{B}$ is $\frac{d(d+1)}{2} \cdot n^2$. Due to Carathéodory's theorem, any element of $\mathcal{S}_{SEP}$ can be written as a convex sum of $\frac{d(d+1)n^2}{2} + 1$ extremal elements of $\mathcal{S}_{SEP}$. Then, we have the following lemma.

Lemma 3. *When* $\dim \mathcal{H} = n$, *we have* $C[G, \frac{d(d+1)n^2}{2} + 1] = C[G]$.

When we replace the cone $\mathcal{S}_{SEP}$ by another cone in the primal conic linear program $P1$, we obtain another primal conic linear program. The relation (34) lists various alternative cones. For instance, when the condition $X \in \mathcal{S}_{SEP}$ is replaced by $X \in \mathcal{S}_P$, we denote the primal conic linear programming problem as $P2$. In fact, since the cone $\mathcal{S}_P$ is given as the set of positive semi-definite matrices in $\mathcal{B}$, $P2$ can be simply solved by semi-definite programming.

When the condition $X \in \mathcal{S}_{SEP}$ is replaced by $X \in \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_{PPT} \cap \mathcal{B}''$, $(X \in \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}'')$, we denote this primal conic linear programming problem as $P3$ $(P4)$. When we denote the minimum of the primal conic linear programming problem Pl by $S(Pl)$ for $l = 2, 3, 4$. The relation (34) implies

$$S(P1) \geq S(P2) \geq S(P3) \geq S(P4). \tag{43}$$

Then, the NH bound $C^{NH}[G]$ and the SLD bound $C^S[G]$ can be calculated by solving $S(P2)$ and $S(P4)$, respectively, as follows.

Theorem 4. *We have*

$$C^{NH}[G] = S(P2) \tag{44}$$
$$C^S[G] = S(P4). \tag{45}$$

The above theorem shows that the NH and SLD bounds are the optimal values of the primal conic linear programs $P2$ and $P4$. Their difference lies in the choice of their cones. The cone of $P2$ is composed of positive semi-definite matrices in $\mathcal{B} = \mathcal{M}_{rs}(\mathcal{R}) \otimes \mathcal{B}_{sa}(\mathcal{H})$, and the cone of $P4$ is composed of all positive semi-definite matrices on $\mathcal{R}_C \otimes \mathcal{H}$ that are also in the cone $\mathcal{B}''$.

Although the minimizations $P2, P3, P4$ are defined in cones different from the cone $\mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P$, these minimizations $P2, P3, P4$ can be considered as semi-definite programming (SDP) in the cone $\mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P$ because their cones are given as linear constraints in the cone $\mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P$. Hence, the NH and SLD bounds $C^{NH}[G]$ and $C^S[G]$ can be solved by semi-definite programming while the reference [12, (22)] already gave the same SDP form as $S(P2)$ for $C^{NH}(G)$.

Proof. First, we show (44). For $X' \in \mathcal{B}'$ and $\vec{Z} = (Z^1, \ldots, Z^d)$ satisfying $\operatorname{Tr} Z^i D_j = \delta^i_j$, we define $X(X', \vec{Z}) \in \mathcal{B}$ as follows.

$$X(X', \vec{Z}) := X' + \sum_{j=1}^{d} (|0\rangle\langle j| + |j\rangle\langle 0|) \otimes Z^j + |0\rangle\langle 0| \otimes I. \tag{46}$$

Since any element in $\mathcal{B}$ can be written in the above form due to the definition of $\mathcal{B}$, it is sufficient to discuss $S(P2)$ to consider the matrix with the form $X(X', \vec{Z})$.
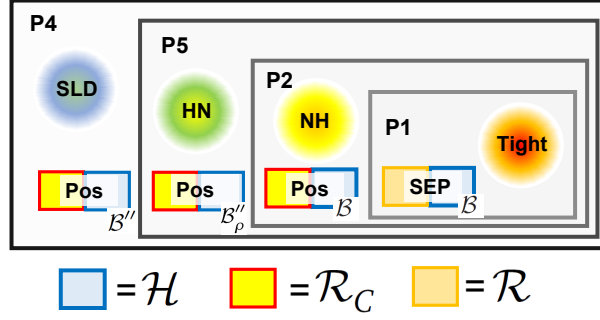
Figure 3: $P1, P2, P4$ and $P5$ are all conic linear programs on the space of operators $\mathcal{B}$, $\mathcal{B}''_\rho$, or $\mathcal{B}''$. The objective functions of all of these conic programs is the same, and the only difference between them lies in the constraints that their operator variables must satisfy. The operator spaces for each of these conic programs are different; P1 is optimized over a bipartite separable space on the tensor product of $\mathcal{R}$ and $\mathcal{H}$. All the other spaces use the complexification of $\mathcal{R}$. The possible choices of X in the programs $P1, P2, P4$ and $P5$ are contained within one another, which we depict visually as nested rectangles. Because of this, we have $S(P1) \geq S(P2) \geq S(P5) \geq S(P4)$.

Any element of $\mathcal{R}_C \otimes \mathcal{H}$ can be written as the form $|(y,z)\rangle := |y\rangle \oplus (|0\rangle \otimes |z\rangle)$ with $|y\rangle(:= \sum_{j=1}^d |j\rangle \otimes |y_j\rangle) \in \mathcal{R}'_C \otimes \mathcal{H}$ and $|z\rangle \in \mathcal{H}$, where $|y_j\rangle \in \mathcal{H}$. Then, we have

$$\langle (y,z)|X(X', \vec{Z})|(y,z)\rangle$$

$$= \langle y|X'|y\rangle + \|z\|^2 + \langle z|\sum_{j=1}^d Z^j y_j\rangle + \langle \sum_{j=1}^d Z^j y_j|z\rangle$$

$$= \langle y|X'|y\rangle + \left\| z + \sum_{j=1}^d Z^j y_j \right\|^2 - \left\| \sum_{j=1}^d Z^j y_j \right\|^2$$

$$= \langle y|X' - \Pi(\vec{Z})|y\rangle + \left\| z + \sum_{j=1}^d Z^j y_j \right\|^2. \tag{47}$$

Considering the case with $z = -\sum_{j=1}^d Z^j y_j$, as pointed by [12, (22)], we find the equivalence between the following two conditions for X' and $\vec{Z}$.

(i) The inequality $\langle (y,z)|X(X', \vec{Z})|(y,z)\rangle \geq 0$ holds for any $|(y,z)\rangle \in \mathcal{R}_C \otimes \mathcal{H}$.

(ii) The inequality $\langle y|X' - \Pi(\vec{Z})|y\rangle \geq 0$ holds for any $|y\rangle \in \mathcal{R}'_C \otimes \mathcal{H}$.

Since the condition (i) is the conditions for $P2$ and the condition (ii) is the conditions for the NH bound $C^{NH}[G]$, the desired statement is obtained.

Next, we proceed to the proof of (45). We choose X' as an element of $\mathcal{S}(\mathcal{R}'_c \otimes \mathcal{H})_P \cap \mathcal{B}''$. When $\vec{Z}$ is fixed, the optimum X' under the condition $X(X', \vec{Z}) \in \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P$ is the operator $\Pi(\vec{Z})$. In this case, the value of the objective function $\mathrm{Tr}[(G \otimes \rho)X]$ is $\mathrm{Tr}[G\mathcal{Z}(\vec{Z})]$. Under the condition (12), we can show $\mathcal{Z}(\vec{Z}) \geq J^{-1}$ in the same way as the SLD Cramér-Rao inequality. When we choose a suitable $\vec{Z}$, the above inequality becomes an equality. Thus, the minimum of $\mathrm{Tr}[G\mathcal{Z}(\vec{Z})]$ under the condition (12) is $\mathrm{Tr}[GJ^{-1}] = C^S[G]$. Therefore, we obtain (45). $\qquad \square$

To get a relation with HN bound, $C^{HN}(G)$, we introduce a linear constraint to the operator $X \in \mathcal{B}''$ as

$$\mathrm{Tr}\left[X((|j\rangle\langle i| - |i\rangle\langle j|) \otimes \rho)\right] = 0 \tag{48}$$

for $i, j = 1, 2, \ldots, d$. Using this linear constraint, we define the subspace $\mathcal{B}''_\rho$ of $\mathcal{B}''$ as

$$\mathcal{B}''_\rho := \{X \in \mathcal{B}'' | (48) \text{ holds.}\}. \tag{49}$$

We consider the minimization:

$$S(P5) := \min_{X \in \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}''_\rho} \left\{ \mathrm{Tr}\Big[(G \otimes \rho)X\Big] \Big| (36), (37) \text{ hold.} \right\}$$

$$= \min_{X \in \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}''} \left\{ \mathrm{Tr}\Big[(G \otimes \rho)X\Big] \Big| (36), (37), (48) \text{ hold.} \right\}. \tag{50}$$

Since we have the relation

$$\mathcal{S}_P \subset \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}''_\rho \subset \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}'', \tag{51}$$

we have the following relations

$$S(P4) \leq S(P5) \leq S(P2). \tag{52}$$

Then, we have the following theorem.

Theorem 5. *We have*

$$C^{HN}[G] = S(P5). \tag{53}$$

The minimization $P5$ can be considered as an SDP in the cone $\mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P$ because the cone is given as linear constraints in the cone $\mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P$ in the same way as the minimizations $P2, P4$. Therefore, the tight, NH, SLD, and HN bounds are summarized as Fig. 3. Although the reference [8, Eq. (11)] derived an SDP form of the HN bound, their SDP is different from $P5$. They assumed the finite-dimensional system for $\mathcal{H}$ in the derivation [8, Eq. (11)]. In contrast, we do not require this assumption to derive the equation (53).

Proof. We show (53) by using the second expression of $S(P5)$ in (50). Since $X(X', \vec{Z}) \in \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}''$, the components of $\vec{Z}$ are self-adjoint. First, we fix $\vec{Z}$. When the condition (48) holds, the condition $X(X', \vec{Z}) \in \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}''$ is rewritten as follows.

$$X'_* := X' - \left(\Pi(\vec{Z}) - \Big(\sum_{1 \leq i,j \leq d} \frac{\mathrm{Tr}\Big[\rho[X^i, X^j]\Big]}{2} |i\rangle\langle j| \Big) \otimes I \right)$$

$$\geq \Big(\sum_{1 \leq i,j \leq d} \frac{\mathrm{Tr}\Big[\rho[X^i, X^j]\Big]}{2} |i\rangle\langle j| \Big) \otimes I. \tag{54}$$

Since $\Big(\Pi(\vec{Z}) - \sum_{1 \leq i,j \leq d} \frac{\mathrm{Tr}[\rho[X^i, X^j]]}{2} |i\rangle\langle j| \otimes I \Big)$ satisfies the condition (48), when X'_* satisfies the condition (48), when X' satisfies the condition (48). The pair of the condition (48) and the condition $X(X', \vec{Z}) \in \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}''$ is rewritten as the pair of the condition (48) for X'_* and the condition (54).

When X'_* satisfies (54), we have $\mathrm{Tr}_{\mathcal{H}} \rho X'_* \geq \sum_{1 \leq i,j \leq d} \frac{\mathrm{Tr}[\rho[X^i, X^j]]}{2} |i\rangle\langle j|$. Conversely, when a $d \times d$ symmetric matrix V' satisfies the condition $V' \geq \sum_{1 \leq i,j \leq d} \frac{\mathrm{Tr}[\rho[X^i, X^j]]}{2} |i\rangle\langle j|$, $X'_* = V' \otimes I$ satisfies (54).

Since $\mathrm{Tr}\Big[(G\otimes\rho)X'_*\Big]=\mathrm{Tr}\Big[G(\mathrm{Tr}_{\mathcal{H}}\rho X'_*)\Big]$, our minimization with fixed $\vec{Z}$ is rewritten as

$$\min_{V':\mathrm{symmetric}}\left\{\mathrm{Tr}GV'\Big|V'\geq\sum_{1\leq i,j\leq d}\frac{\mathrm{Tr}\Big[\rho[X^i,X^j]\Big]}{2}|i\rangle\langle j|\right\}$$
$$=\min_{V':\mathrm{symmetric}}\left\{\mathrm{Tr}GV'\Big|V'\geq\sqrt{-1}\mathrm{Im}\,\mathcal{Z}(\vec{Z})\right\},\tag{55}$$

which equals $\mathrm{Tr}\Big[|G^{\frac{1}{2}}\mathrm{Im}\,\mathcal{Z}(\vec{Z})G^{\frac{1}{2}}|\Big]$.

Since

$$\mathrm{Tr}\Big[(G\otimes\rho)\Big(\Pi(\vec{Z})-\Big(\sum_{1\leq i,j\leq d}\frac{\mathrm{Tr}\Big[\rho[X^i,X^j]\Big]}{2}|i\rangle\langle j|\Big)\otimes I\Big)\Big]$$
$$=\mathrm{Tr}\Big[(G\otimes\rho)\Pi(\vec{Z})\Big]=\mathrm{Tr}\Big[G\mathrm{Re}\,\mathcal{Z}(\vec{Z})\Big],\tag{56}$$

the minimum value of the objective function with fixed $\vec{Z}$ is $\mathrm{Tr}[G\mathrm{Re}\,\mathcal{Z}(\vec{Z})]+\mathrm{Tr}[|G^{\frac{1}{2}}\mathrm{Im}\,\mathcal{Z}(\vec{Z})G^{\frac{1}{2}}|]$. Since the condition for $\vec{Z}$ is the same as the HN bound, we obtain (53). $\qquad\square$

Now, for a $d\times d$ real matrix a and a self-adjoint operator S on the Hilbert space $\mathcal{H}$, let us define the operator as an element of $\mathcal{T}$:

$$\Pi(a,S):=G\otimes\rho-\frac{1}{2}\Big(\sum_{1\leq i,j\leq d}a_i^j(|0\rangle\langle i|+|i\rangle\langle 0|)\otimes D_j\Big)$$
$$-|0\rangle\langle 0|\otimes S.\tag{57}$$

The dual problem $D1$ of $P1$ is given as the following maximization:

$$S(D1):=\max_{(a,S)\in\mathbb{R}^{d\times d}\times\mathcal{T}_{sa}(\mathcal{H})}\left\{\sum_i a_i^i+\mathrm{Tr}S\Big|\Pi(a,S)\in\mathcal{S}^*_{SEP}\right\}.\tag{58}$$

Also, the dual problem $D2$ of $P2$ is given as the following maximization:

$$S(D2):=\max_{(a,S)\in\mathbb{R}^{d\times d}\times\mathcal{T}_{sa}(\mathcal{H})}\left\{\sum_i a_i^i+\mathrm{Tr}S\Big|\Pi(a,S)\in\mathcal{S}^*_P\right\}.\tag{59}$$

In the same way, the dual problems $D3$ and $D4$ of $P3$ and $P4$ are given as the following maximizations:

$$S(D3)$$
$$:=\max_{(a,S)\in\mathbb{R}^{d\times d}\times\mathcal{T}_{sa}(\mathcal{H})}\left\{\sum_i a_i^i+\mathrm{Tr}S\Big|\right.$$
$$\left.\Pi(a,S)\in(\mathcal{S}(\mathcal{R}_C\otimes\mathcal{H})_{PPT}\cap\mathcal{B}'')^*\right\}.\tag{60}$$

$$S(D4)$$
$$:=\max_{(a,S)\in\mathbb{R}^{d\times d}\times\mathcal{T}_{sa}(\mathcal{H})}\left\{\sum_i a_i^i+\mathrm{Tr}S\Big|\right.$$
$$\left.\Pi(a,S)\in(\mathcal{S}(\mathcal{R}_C\otimes\mathcal{H})_P\cap\mathcal{B}'')^*\right\}.\tag{61}$$

$$S(D5)$$
$$:=\max_{(a,S)\in\mathbb{R}^{d\times d}\times\mathcal{T}_{sa}(\mathcal{H})}\left\{\sum_i a_i^i+\mathrm{Tr}S\Big|\right.$$
$$\left.\Pi(a,S)\in(\mathcal{S}(\mathcal{R}_C\otimes\mathcal{H})_P\cap\mathcal{B}''_\rho)^*\right\}.\tag{62}$$

Since the minimizations $P1, \dots, P5$ are conic linear programmings, we have the following theorem.

Theorem 6. *We have*

$$S(Pl) = S(Dl) \tag{63}$$

for $l = 1, 2, 3, 4, 5$.

For its proof, see Appendix C. When $\mathcal{H}$ is a finite-dimensional space, the conic linear programing problems appearing in Theorem 6 are conic linear programing problems on finite-dimensional vector spaces.

Also, as shown in Appendix D, the following theorem holds.

Theorem 7. *We have*

$$S(D0) = S(D1). \tag{64}$$

The combination of Theorems 2, 6, and 7 implies Proposition 1. The combination of Theorems 2, 6, and 7 implies another proof of Proposition 1.

4 Calculation of $S(P1)$

4.1 Upper bound

Although the minimization problem $P2$, i.e., the NH bound $C^{NH}[G]$ can be solved by semi-definite programming, the cone in the primal conic linear programming $P1$ is the separable cone $\mathcal{S}_{SEP}$, which is different from the set of positive semi-definite matrices. Hence, the primal conic linear programming $P1$ is still not so easy even with a finite-dimensional space $\mathcal{H}$. In fact, this type of problem appears in the membership problem of the separable cone $\mathcal{S}_{SEP}$, which appears in the decision problem of the existence of entanglement [14, 15] and in generalized robustness of entanglement [17]. Also, the reference [42, Proposition 2] showed that the communication value can be calculated as a conic linear programming with the separable cone $\mathcal{S}_{SEP}$.

In this section, we consider an efficient algorithm to solve $S(P1)$. Although we present our algorithm for the conic linear programming $S(P1)$, this algorithm can be applied to a general conic linear programming with the separable cone $\mathcal{S}_{SEP}$ including the problem presented in [42, Proposition 2]. To solve this problem, we choose a finite set $\mathcal{W}_R := \{|w_s\rangle\}_{s=1}^m$ composed of several normalized vectors in $\mathbb{R}^{d+1}$. Then, we choose a subset $\mathcal{S}(\mathcal{W}_R) := \{\sum_{s=1}^m |w_s\rangle\langle w_s| \otimes X_s | X_s \in \mathcal{T}_{sa,+}(\mathcal{H})\} \subset \mathcal{S}_{SEP}$. Replacing $\mathcal{S}_{SEP}$ by $\mathcal{S}(\mathcal{W}_R)$, we consider the minimization:

$$S[P1, \mathcal{W}_R] := \min_{X \in \mathcal{S}(\mathcal{W}_R)} \left\{ \text{Tr}\Big[(G \otimes \rho)X\Big] \Big| (36), (37) \text{ hold.} \right\}. \tag{65}$$

The inclusion relation $\mathcal{S}(\mathcal{W}_R) \subset \mathcal{S}_{SEP}$ implies the relation

$$S[P1, \mathcal{W}_R] \geq S(P1). \tag{66}$$

We illustrate the relationship between $P1$ and the SDP $[P1, \mathcal{W}_R]$ in Fig. 4. To calculate $S[P1, \mathcal{W}_R]$, we define the set of block diagonal matrices $\mathcal{T}(\mathcal{W}_R) := \{\sum_{s=1}^m |s\rangle\langle s| \otimes X_s | X_s \in \mathcal{T}_{sa,+}(\mathcal{H})\}$ on the vector space $\mathbb{C}^m \otimes \mathcal{H}$. Identifying an element $\sum_{s=1}^m |w_s\rangle\langle w_s| \otimes X_s \in$
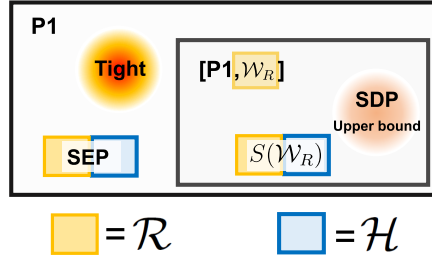
Figure 4: Relationship between $P1$ and $[P1, \mathcal{W}_R]$. To convert the problem $P1$ to an SDP, we have introduced a subset $\mathcal{S}(\mathcal{W}_R)$ of $\mathcal{S}_{SEP}$. The inclusion relation $\mathcal{S}(\mathcal{W}_R) \subset \mathcal{S}_{SEP}$ implies the relation $S[P1, \mathcal{W}_R] \geq S(P1)$. Also, due to the construction of $\mathcal{S}(\mathcal{W}_R)$, $S[P1, \mathcal{W}_R]$ can be solved by an SDP.

$\mathcal{S}(\mathcal{W}_R)$, $X_s \in \mathcal{T}_{sa,+}(\mathcal{H})$ with a block diagonal matrix $\sum_{s=1}^{m} |s\rangle\langle s| \otimes X_s \in \mathcal{T}(\mathcal{W}_R)$, we have another form of $S[P1, \mathcal{W}_R]$:

$$S[P1, \mathcal{W}_R]$$
$$= \min_{X=(X_s)\in\mathcal{T}(\mathcal{W}_R)} \left\{ \sum_{s=1}^{m} \langle w_s|G|w_s\rangle \mathrm{Tr} X_s \rho \,\middle|\, C[P1, \mathcal{W}_R] \text{ holds.} \right\}, \tag{67}$$

where the condition $C[P1, \mathcal{W}_R]$ is defined as

$$\sum_{s=1}^{m} \langle w_s|0\rangle\langle 0|w_s\rangle X_s = I, \tag{68}$$

$$\frac{1}{2}\sum_{s=1}^{m} \langle w_s|(|0\rangle\langle i| + |i\rangle\langle 0|)|w_s\rangle \mathrm{Tr}[X_s D_j] = \delta_i^j. \tag{69}$$

We call the above problem $[P1, \mathcal{W}_R]$. This problem can be considered as an SDP with $d^2 + n^2$ constraints with block diagonal matrices, which is a typical case of sparse SDP.

We define the number $\delta(\mathcal{W}_R)$ as

$$\delta(\mathcal{W}_R) := \max_{x\in\mathbb{R}^{d+1}:\|x\|=1} \min_{w\in\mathcal{W}_R} \||x\rangle\langle x| - |w\rangle\langle w|\|_1. \tag{70}$$

We consider a sequence $\mathcal{W}_{R,n}$ such that $\delta(\mathcal{W}_{R,n}) \to 0$. As shown in Theorem 13, we have

$$\lim_{n\to\infty} S[P1, \mathcal{W}_{R,n}] = S(P1). \tag{71}$$

Hence, we can use SDP for this problem.

Many existing algorithms for SDP work well for sparse positive semi-definite matrices. The paper [43] studied the calculation complexity for generic primal-dual interior-point method for SDP. When we apply their analysis for the calculation complexity to the case with block diagonal matrices, as explained in Remark 8, we find that the calculation complexity of $S[P1, \mathcal{W}_R]$ is

$$O(m((d^2 + n^2)n^3 + (d^2 + n^2)^2 n^2))$$
$$= O(m(n^6 + d^2 n^3 + d^4 n^2)). \tag{72}$$

Remark 8. *To get (72), we explain how to apply the analysis by the reference [43]. In the reference [43], m is the number of constraint, and n is the dimension of the vector space to*

define the positive semi-definite matrix. In their analysis for generic primal-dual interior-point method for SDP, the dominant part of the calculation complexity is given as the calculation of $\mathbf{B}'$ and $\mathbf{r}'$ in Eq. (8') of [43]. When the matrices are given as block diagonal matrices, the number of blocks is n_1 and size of each block is n_2, the calculation complexity of $\mathbf{B}'$ is $O(mn_2^3 n_1 + m^2 n_2^2 n_1)$ and the calculation complexity of $\mathbf{r}'$ is $O(n_2^3 n_1 + mn_2^2 n_1)$. Therefore, the total calculation complexity is given as $O(mn_2^3 n_1 + m^2 n_2^2 n_1)$. Applying this estimation to our setting, we obtain (72).

4.2 Estimator attaining upper bound

Here, we explain how to construct an estimator to attain the upper bound $S[P1, \mathcal{W}_R]$ with $\mathcal{W}_R = \{|w_s\rangle\}_{s=1}^m$, where $|w_s\rangle = \sum_{j=0}^d w_s^j |j\rangle$. For the optimal solution $X_1^*, \ldots, X_m^*$ to the SDP $[P1, \overline{\mathcal{W}}_R]$, we can define the positive semi-definite matrix

$$M_s := |w_s^0|^2 X_s^*. \tag{73}$$

The first condition (68) implies

$$\sum_{s=1}^m M_s = \sum_{s=1}^m |w_s^0|^2 X_s^* = \sum_{s=1}^m \langle w_s|0\rangle\langle 0|w_s\rangle X_s^* = I, \tag{74}$$

which shows that $\{M_s\}$ satisfies the condition of POVM. For each outcome s, we define $\hat{\theta}(s) \in \mathbb{R}^d$ as

$$\hat{\theta}^i(s) := \frac{w_s^i}{w_s^0} + \theta^i \tag{75}$$

so that the pair $(\{M_s\}, \hat{\theta})$ forms an estimator. The second condition (69) implies

$$\sum_{s=1}^m \hat{\theta}^i(s) \mathrm{Tr} M_s D_j = \sum_{s=1}^m \frac{w_s^i}{w_s^0} \mathrm{Tr} |w_s^0|^2 X_s^* D_j + \sum_{s=1}^m \theta^i \mathrm{Tr} |w_s^0|^2 X_s^* D_j$$

$$= \frac{1}{2} \sum_{s=1}^m \langle w_s|(|0\rangle\langle i| + |i\rangle\langle 0|)|w_s\rangle \mathrm{Tr} X_s^* D_j + \theta^i \mathrm{Tr} I D_j = \delta_i^j + 0. \tag{76}$$

Therefore, the estimator $(\{M_s\}, \hat{\theta})$ satisfies the locally unbiasedness condition. Its weighted mean square error is calculated as

$$\sum_{i,j} G_{i,j} \sum_{s=1}^m (\hat{\theta}^j(s) - \theta^j)(\hat{\theta}^i(s) - \theta^i) \mathrm{Tr} M_s D_j \tag{77}$$

$$= \sum_{i,j} G_{i,j} \sum_{s=1}^m \frac{w_s^j}{w_s^0} \frac{w_s^i}{w_s^0} \mathrm{Tr} |w_s^0|^2 X_s^* D_j = \sum_{s=1}^m \langle w_s|G|w_s\rangle \mathrm{Tr} X_s^* \rho. \tag{78}$$

Therefore we have the following result.

Theorem 9. *Let us use the optimal solution of $[P1, \mathcal{W}_R]$ to construct M_s according to* (73) *and $\hat{\theta}$ according to* (75). *Then the estimator $(\{M_s\}, \hat{\theta})$ attains the upper bound $S[P1, \mathcal{W}_R]$.*

4.3 Lower bound

Since the solution $S[P1, \mathcal{W}_R]$ of the SDP gives an upper bound of $S(P1)$, we need to derive a lower bound for $S(P1)$. However, it is not easy to derive an lower bound of $S(P1)$ from the solution $S[P1, \mathcal{W}_R]$. In the following, we present a lower bound of $S(P1)$ under the algorithm presented in the above section. Although the algorithm and the upper bound presented in the above section work with a general conic linear programming with the separable cone $\mathcal{S}_{SEP}$, the following lower bound does not necessarily work with a general conic linear programming with the separable cone $\mathcal{S}_{SEP}$ because the lower bound depends on the form of the conic linear programming $S(P1)$ with the separable cone $\mathcal{S}_{SEP}$.

To derive its lower bound, we denote the optimum a and S of the dual problem $[D1, \mathcal{W}_R]$ of $[P1, \mathcal{W}_R]$ by $a^* = \sum_{i,j} a^*_{i,j} |i\rangle\langle j|$ and S^*.

Namely, we define the dual problem $[D1, \mathcal{W}_R]$ as

$$\underset{(a, S) \in \mathbb{R}^{d \times d} \times \mathcal{T}_{sa}(\mathcal{H})}{\text{maximize}} \quad \sum_i a^i_i + \text{Tr} S \tag{79a}$$

$$\text{subject to} \quad \sum_{s=1}^m |s\rangle\langle s| \otimes \langle w_s | \Pi(a, S) | w_s \rangle \geq 0 \tag{79b}$$

with optimal value $S[D1, \mathcal{W}_R] := \sum_i (a^*)^i_i + \text{Tr} S^*$.

Then, we define the real number

$$C_2(a) := \frac{1}{2} \left\| \sum_j \left(\rho^{-1/2} (\sum_{j'} a^{j'}_j D_{j'}) \rho^{-1/2} \right)^2 \right\|^{1/2}. \tag{80}$$

for a $d \times d$ matrix a. Using this value, we define

$$X^* := \Pi(a^*, S^*) \tag{81}$$

$$\kappa := - \min_{v \in \mathcal{H}: \|v\| = 1} \min_{y: \|y\| \leq C_2(a^*)} \langle (1, y), v | X^* | (1, y), v \rangle. \tag{82}$$

Then, we have the following lemma.

Lemma 10. *A pair (a, S) satisfies the condition (23) if and only if the relation*

$$\|y\|^2 \rho + \sum_j (ay)^j D_j - S \geq 0 \tag{83}$$

holds for $y \in \mathbb{R}^d$ with $\|y\| \leq C_2(a)$.

Proof. In this proof, $(c_j)_j$ expresses the vector whose components are $c_1, \ldots, c_d$. A pair (a, S) satisfies the condition (23) if and only if any $y \in \mathbb{R}^d$ and $|\phi\rangle \in \mathcal{H}$ satisfy

$$\|y\|^2 \langle \phi | \rho | \phi \rangle + \sum_{j,j'} a^j_{j'} y^{j'} \langle \phi | D_j | \phi \rangle - \langle \phi | S | \phi \rangle \geq 0. \tag{84}$$

The LHS is rewritten as

$$\langle \phi | \rho | \phi \rangle \left\| y - \left(\frac{1}{2\langle \phi | \rho | \phi \rangle} \sum_{j'} a^j_{j'} \langle \phi | D_{j'} | \phi \rangle \right)_j \right\|^2$$

$$- \langle \phi | S | \phi \rangle - \frac{1}{4\langle \phi | \rho | \phi \rangle} \left\| \left(\sum_{j'} a^j_{j'} \langle \phi | D_{j'} | \phi \rangle \right)_j \right\|^2. \tag{85}$$

That is, when $|\phi\rangle$ is fixed, the minimum of the LHS for y is realized when $y = \left(\frac{1}{2\langle\phi|\rho|\phi\rangle} \sum_{j'} a_j^{j'} \langle\phi|D_{j'}|\phi\rangle \right)_j$.

Therefore, the range of $y \in \mathbb{R}^d$ in the condition (84) can be limited to the set $\left\{ y \in \mathbb{R}^d \middle| \|y\| \leq \right.$

$\left. \max_{|\phi\rangle} \left\| \left(\frac{1}{2\langle\phi|\rho|\phi\rangle} \sum_{j'} a_j^{j'} \langle\phi|D_{j'}|\phi\rangle \right)_j \right\| \right\}$.

Since we have

$$\left\| \left(\frac{1}{2\langle\phi|\rho|\phi\rangle} \sum_{j'} a_j^{j'} \langle\phi|D_{j'}|\phi\rangle \right)_j \right\|^2$$

$$= \sum_j \left(\frac{1}{2\langle\phi|\rho|\phi\rangle} \sum_{j'} a_j^{j'} \langle\phi|D_{j'}|\phi\rangle \right)^2$$

$$= \sum_j \left(\frac{1}{2\langle\phi|\rho|\phi\rangle} \langle\phi| \sum_{j'} a_j^{j'} D_{j'} |\phi\rangle \right)^2, \tag{86}$$

using $|\phi'\rangle := \rho^{1/2}|\phi\rangle$, we have

$$\max_{|\phi\rangle} \left\| \left(\frac{1}{2\langle\phi|\rho|\phi\rangle} \sum_{j'} a_j^{j'} \langle\phi|D_{j'}|\phi\rangle \right)_j \right\|^2$$

$$= \max_{|\phi'\rangle} \sum_j \left(\frac{1}{2\langle\phi'|\phi'\rangle} \langle\phi'|\rho^{-1/2}(\sum_{j'} a_j^{j'} D_{j'})\rho^{-1/2}|\phi'\rangle \right)^2$$

$$= \max_{|\phi'\rangle} \sum_j \frac{1}{4\langle\phi'|\phi'\rangle^2} \left(\langle\phi'|\rho^{-1/2}(\sum_{j'} a_j^{j'} D_{j'})\rho^{-1/2}|\phi'\rangle \right)^2$$

$$= \max_{|\phi'\rangle:\|\phi'\|=1} \sum_j \frac{1}{4} \left(\langle\phi'|\rho^{-1/2}(\sum_{j'} a_j^{j'} D_{j'})\rho^{-1/2}|\phi'\rangle \right)^2$$

$$\leq \max_{|\phi'\rangle:\|\phi'\|=1} \sum_j \frac{1}{4} \langle\phi'| \left(\rho^{-1/2}(\sum_{j'} a_j^{j'} D_{j'})\rho^{-1/2} \right)^2 |\phi'\rangle$$

$$= \max_{|\phi'\rangle:\|\phi'\|=1} \frac{1}{4} \langle\phi'| \sum_j \left(\rho^{-1/2}(\sum_{j'} a_j^{j'} D_{j'})\rho^{-1/2} \right)^2 |\phi'\rangle$$

$$= \frac{1}{4} \left\| \sum_j \left(\rho^{-1/2}(\sum_{j'} a_j^{j'} D_{j'})\rho^{-1/2} \right)^2 \right\| = C_2(a)^2. \tag{87}$$

Hence, the range of $y \in \mathbb{R}^d$ in the condition (84) can be limited to the set $\left\{ y \in \mathbb{R}^d \middle| \|y\| \leq \right.$
$\left. C_2(a) \right\}$. Thus, we obtain the desired statement. $\qquad\square$

Then, for the lower bound of $S(P1)$, we have the following theorem.

Theorem 11. *We have the following relation.*

$$S[D1, \mathcal{W}_R] = S[P1, \mathcal{W}_R] \geq S(P1) = S(D1)$$
$$\geq \underline{S}[D1, \mathcal{W}_R] := S[D1, \mathcal{W}_R] - n\kappa, \tag{88}$$

where n is the dimension of Hilbert space $\mathcal{H}$. In particular, (94) holds with equality in the limit $\delta \to 0$.

When we calculate a lower bound of $S(P1)$ from $S[P1, \mathcal{W}_R]$, we have knowledge for a^* and X^* in a real numerical calculation. Hence, it is possible to calculate κ by applying Theorem 11.

Proof. It is sufficient to show that the pair $(a^*, S^* - \kappa I)$ satisfies the condition (23) because in this case, the corresponding objective function of $[D1, \mathcal{W}_R]$ is $\mathrm{Tr}\, a^* + \mathrm{Tr}(S^* - \kappa I)$. Due to Lemma 10, it is sufficient to show that the pair $(a^*, S^* - \kappa I)$ satisfies

$$\|y\|^2 \rho + \sum_j (a^* y)^j D_j - S^* + \kappa I \geq 0 \tag{89}$$

for $y \in \mathbb{R}^d$ with $\|y\| \leq C_2(a^*)$.

For $y \in \mathbb{R}^d$ with $\|y\| \leq C_2(a^*)$ and $v \in \mathcal{H}$ with $\|v\| = 1$, we have

$$\langle v | \left(\|y\|^2 \rho + \sum_j (a^* y)^j D_j - S^* + \kappa I \right) | v \rangle$$
$$= \langle (1, y), v | X^* | (1, y), v \rangle + \langle v | \kappa I | v \rangle$$
$$= \langle (1, y), v | X^* | (1, y), v \rangle + \kappa \geq 0, \tag{90}$$

which implies (88). $\qquad \square$

4.4 Theoretical Evaluation of κ

When we derive a lower bound of $S(P1)$ in our numerical calculation, we can calculate κ numerically, and can directly use Theorem 11. However, when we estimate the calculation complexity to get the tight CR bound within additive error ϵ, we need to evaluate κ theoretically. For this aim, we define

$$\delta(\mathcal{W}_R) := \max_{x \in \mathbb{R}^{d+1} : \|x\| = 1} \min_{|w\rangle \in \mathcal{W}_R} \||x\rangle\langle x| - |w\rangle\langle w|\|_1. \tag{91}$$

Then, the following lemma gives an upper bound of κ.

Lemma 12. *The quantity $\delta(\mathcal{W}_R)$ gives an upper bound of κ.*

$$\kappa \leq \|X^*\|(1 + C_2(a^*)^2)\delta(\mathcal{W}_R). \tag{92}$$

Proof. For $y \in \mathbb{R}^d$ with $\|y\| \leq C_2(a^*)$ and $\delta = \delta(\mathcal{W}_R)$, we have

$$\langle (1, y) | X^* | (1, y) \rangle$$
$$\geq (1 + \|y\|^2)\langle w | X^* | w \rangle - \|X^*\| \||(1, y)\rangle\langle(1, y)|$$
$$\quad - (1 + \|y\|^2)|w\rangle\langle w|\|_1 I$$
$$\geq (1 + \|y\|^2)\langle w | X^* | w \rangle - \|X^*\|(1 + \|y\|^2)\delta I$$
$$\geq (1 + \|y\|^2)\langle w | X^* | w \rangle - \|X^*\|(1 + C_2(a^*)^2)\delta I$$
$$\geq -\|X^*\|(1 + C_2(a^*)^2)\delta I, \tag{93}$$

which implies (92). $\qquad \square$

Combining Theorem 11 and Lemma 12, we have the following theorem.

Theorem 13.

$$S[D1, \mathcal{W}_R] = S[P1, \mathcal{W}_R] \geq S(P1) = S(D1)$$
$$\geq S[D1, \mathcal{W}_R] - n\|X^*\|(1 + C_2(a^*)^2)\delta(\mathcal{W}_R). \tag{94}$$

Considering the structure of κ, we have another evaluation of κ. For this aim, we define

$$\delta(s, \mathcal{W}_R) := \max_{x=(\cos\theta,(\sin\theta)z), z\in S^{d-1}} \min_{|w\rangle\in\mathcal{W}_R} \||x\rangle\langle x| - |w\rangle\langle w|\|_1, \quad (95)$$

with $\theta = \tan^{-1} s$. Then, the following lemma gives another upper bound of κ.

Lemma 14. *Using the quantity $\delta(s, \mathcal{W}_R)$, we have the following upper bound of κ.*

$$\kappa \le \|X^*\| \max_{s\in[0,C_2(a^*)]} (1+s^2)\delta(s, \mathcal{W}_R). \quad (96)$$

Proof. For $y \in \mathbb{R}^d$ with $\|y\| \le C_2(a^*)$, we have

$$\begin{aligned}
&\langle(1,y)|X^*|(1,y)\rangle \\
\ge& (1+\|y\|^2)\langle w|X^*|w\rangle \\
& - \|X^*\| \||(1,y)\rangle\langle(1,y)| - (1+\|y\|^2)|w\rangle\langle w|\|_1 I \\
\ge& (1+\|y\|^2)\langle w|X^*|w\rangle - \|X^*\|(1+\|y\|^2)\delta(\|y\|, \mathcal{W}_R)I \\
\ge& (1+\|y\|^2)\langle w|X^*|w\rangle - \|X^*\| \max_{s\in[0,C_2(a^*)]} (1+s^2)\delta(s, \mathcal{W}_R)I \\
\ge& - \|X^*\| \max_{s\in[0,C_2(a^*)]} (1+s^2)\delta(s, \mathcal{W}_R)I.
\end{aligned} \quad (97)$$

$\square$

5 Construction of $\mathcal{W}_R$

5.1 Quantum t-design

To implement the proposed method, we need to construct the subset $\mathcal{W}_R \subset \mathbb{R}^{d+1}$. For this aim, we like to pack m points uniformly on a d-sphere of radius 1, such that any point on the sphere is as close to some point as possible. As one possible choice, we randomly generate pure states $|w_1\rangle\langle w_1|, \ldots, |w_l\rangle\langle w_l|$ on the system $\mathcal{H}$ subject to the Haar measure. Another choice for $\mathcal{W}_R$ is a quantum t-design.

A subset $\mathcal{W}_H$ of normalized vectors on the finite-dimensional Hilbert space $\mathcal{H}$ is called a quantum t-design on $\mathcal{H}$ when

$$\sum_{w\in\mathcal{W}_H} \frac{1}{|\mathcal{W}_H|} |w\rangle\langle w|^{\otimes t} = \int |x\rangle\langle x|^{\otimes t} \mu_{\mathcal{H}}(dx), \quad (98)$$

where μ is the Haar measure on the set of pure states on $\mathcal{H}$ [44] [45, Section 4.5]. A subset $\mathcal{W}_R$ of normalized vectors on $\mathcal{R}$ is called a quantum t-design on $\mathcal{R}$ when

$$\sum_{w\in\mathcal{W}_R} \frac{1}{|\mathcal{W}_R|} |w\rangle\langle w|^{\otimes t} = \int |x\rangle\langle x|^{\otimes t} \mu_{\mathcal{R}}(dx), \quad (99)$$

where μ is the Haar measure on the set of pure states on $\mathcal{R}$.

For a good choice of $\mathcal{W}_R$ and $\mathcal{W}_H$, we can consider a quantum t-design because a quantum t-design has a symmetric property. For example, the paper [46] discusses approximate construction of quantum t-designs.

5.2 Spherical t-design

A quantum t design has not been sufficiently studied, but a spherical t-design has been well studied [47]. The references [48, 49, 50] study a spherical t-design. Fortunately, a quantum t-design can be constructed from a spherical t-design [44].

A subset $\mathcal{V}$ of S^{d-1} is called a spherical t-design when the relation

$$\sum_{x \in S^{d-1}} f(x) = \int f(x) \mu_{S^{d-1}}(dx) \tag{100}$$

holds any polynomial f with degree t, where $\mu_{S^{d-1}}$ is the Haar measure on S^{d-1}.

Given a spherical $2t$-design $\mathcal{V} \subset S^{2d-1}$, we define a subset $\mathcal{V}_H$ of pure states on $\mathcal{H} = \mathbb{C}^d$ as follows. For $x \in S^{2d-1}$, we define the normalized vector $w(x) \in \mathcal{H}$ as $w(x)_j := x_{2j-1} + x_{2j}i$. As shown below, the set $\mathcal{V}_H := \{w(x)\}_{x \in \mathcal{V}}$ is a quantum t-design. For any sequences $e_1, \ldots, e_t$ and $f_1, \ldots, f_t$, we have

$$\sum_{x \in \mathcal{V}} \frac{1}{|\mathcal{V}|} \mathrm{Tr}\Big[|w(x)\rangle\langle w(x)|^{\otimes t}|e_1, \ldots, e_t\rangle\langle f_1, \ldots, f_t|\Big]$$

$$= \sum_{x \in \mathcal{V}} \frac{1}{|\mathcal{V}|} \prod_{j=1}^{t} \prod_{j'=1}^{t} w(x)_{e_j} \overline{w(x)}_{f_{j'}}$$

$$= \int \prod_{j=1}^{t} \prod_{j'=1}^{t} w(x)_{e_j} \overline{w(x)}_{f_{j'}} \mu_{S^{2d-1}}(dx)$$

$$= \int \mathrm{Tr}\Big[|w(x)\rangle\langle w(x)|^{\otimes t}|e_1, \ldots, e_t\rangle\langle f_1, \ldots, f_t|\Big] \mu_{S^{2d-1}}(dx)$$

$$= \int \mathrm{Tr}\Big[|x\rangle\langle x|^{\otimes t}|e_1, \ldots, e_t\rangle\langle f_1, \ldots, f_t|\Big] \mu_{\mathcal{H}}(dx). \tag{101}$$

Hence, we have

$$\sum_{x \in \mathcal{V}} \frac{1}{|\mathcal{V}|} |w(x)\rangle\langle w(x)|^{\otimes t} = \int |x\rangle\langle x|^{\otimes t} \mu_{\mathcal{H}}(dx). \tag{102}$$

which shows that the set $\mathcal{V}_H := \{w(x)\}_{x \in \mathcal{V}}$ is a quantum t-design. In the same way, given a spherical t-design $\mathcal{V} \subset S^{d-1}$, the set $\mathcal{V}$ is a quantum t-design on $\mathcal{R} = \mathbb{R}^d$.

5.3 Construction reflecting structure of κ

Since the lower bound of $S(D1)$ depends on the value κ, we may construct $\mathcal{W}$ by reflecting the structure of κ as follows. We choose a subset $\mathcal{S} \subset S^{d-1}$, which can be a spherical t-design. Then, we define a subset $\mathcal{S}(\phi) := \{(\cos\phi, \sin\phi y)\}_{y \in \mathcal{S}}$. We choose $\phi_0 > 0$ as $\tan\phi_0 = C_2(a^*)$. Using k, we choose $\mathcal{W}_R$ as $\cup_{j=0}^{k} \mathcal{S}(\frac{\phi_0 j}{k})$.

We can modify this construction as follows. That is, to construct $\mathcal{S}(\frac{\phi_0 j}{k})$, depending on j, we choose a subset $\mathcal{S}_j \subset S^{d-1}$, which also may be a spherical t-design. We define a subset $\mathcal{S}(\frac{\phi_0 j}{k}) := \{(\cos\frac{\phi_0 j}{k}, \sin\frac{\phi_0 j}{k} y)\}_{y \in \mathcal{S}_j}$. We choose $\phi_0 > 0$ as $\tan\phi_0 = C_2(a^*)$. Using k, we choose $\mathcal{W}_R$ as $\cup_{j=0}^{k} \mathcal{S}(\frac{\phi_0 j}{k})$.

Indeed, when $d = 2$, the choice of $\mathcal{S}$ is very easy because we can choose $\mathcal{S}$ as $\{(\cos\frac{2\pi l}{N}, \sin\frac{2\pi l}{N})\}_{l=1}^{N}$. In this case, we find that $\delta(\mathcal{S}) = \|(1, 0) - (\cos 2\pi/N, \sin 2\pi/N)\| = \sqrt{2 - 2\cos 2\pi/N}$.

Appendix E discusses the theoretical evaluation of κ in these choices while it is better to directly calculate κ from the obtained data a^* and X^* in a real numerical calculation.

5.4 Another construction

First, we construct the real discrete subset $\mathcal{D}_{n,d} \subset \mathcal{S}_d$, where $\mathcal{S}_d$ is the d-dimensional sphere in $\mathbb{R}^{d+1}$. We define $\mathcal{D}_{n,1}$ as

$$\mathcal{D}_{n,1} := \left\{ \left(\cos \frac{2\pi j}{n}, \sin \frac{2\pi j}{n} \right) \right\}_{j=0}^{n-1} \tag{103}$$

Then, we inductively define $\mathcal{D}_{n,d}$ as

$$\mathcal{D}_{n,d} := \left\{ \left(\cos \frac{2\pi j}{n} v, \sin \frac{2\pi j}{n} \right) \middle| v \in \mathcal{D}_{n,d-1,R}, j = 0, \ldots, n-1 \right\}. \tag{104}$$

Then, $|\mathcal{D}_{n,d}| = n^d$. We have

$$F(\mathcal{D}_{n,1}) := \min_{\psi_1 \in \mathcal{S}_2} \max_{\psi_2 \in \mathcal{D}_{n,2,R}} |\langle \psi_1 | \psi_2 \rangle| = \cos \frac{\pi}{n}. \tag{105}$$

In general, we have

$$F(\mathcal{D}_{n,d}) = \min_{\psi_1 \in \mathcal{S}_d} \max_{\psi_2 \in \mathcal{D}_{n,d}} |\langle \psi_1 | \psi_2 \rangle| = \cos^d \frac{\pi}{n}. \tag{106}$$

This can be inductively shown as follows.

$$F(\mathcal{D}_{n,d}) = \min_{\psi_1 \in \mathcal{S}_d} \max_{\psi_2 \in \mathcal{D}_{n,d}} |\langle \psi_1 | \psi_2 \rangle|$$

$$= \min_{\theta} \min_{\psi_1 \in \mathcal{S}_{d-1}} \max_{j=0,\ldots,n-1} \max_{\psi_2 \in \mathcal{D}_{n,d-1}} \left| \langle \psi_1 | \psi_2 \rangle \cos \theta \cos \frac{2\pi j}{n} + \sin \theta \sin \frac{2\pi j}{n} \right|$$

$$= \min_{\theta} \min_{\psi_1 \in \mathcal{S}_{d-1}} \max_{j=0,\ldots,n-1} \max_{\psi_2 \in \mathcal{D}_{n,d-1}} \left| \langle \psi_1 | \psi_2 \rangle \left(\cos \theta \cos \frac{2\pi j}{n} + \frac{1}{\langle \psi_1 | \psi_2 \rangle} \sin \theta \sin \frac{2\pi j}{n} \right) \right|$$

$$= \min_{\theta} \min_{\psi_1 \in \mathcal{S}_{d-1}} \max_{\psi_2 \in \mathcal{D}_{n,d-1}} |\langle \psi_1 | \psi_2 \rangle| \left(\max_{j=0,\ldots,n-1} \left| \cos \theta \cos \frac{2\pi j}{n} + \frac{1}{\langle \psi_1 | \psi_2 \rangle} \sin \theta \sin \frac{2\pi j}{n} \right| \right)$$

$$\geq \min_{\theta} \min_{\psi_1 \in \mathcal{S}_{d-1}} \max_{\psi_2 \in \mathcal{D}_{n,d-1}} |\langle \psi_1 | \psi_2 \rangle| \left(\max_{j=0,\ldots,n-1} \max \left(\left| \cos \theta \cos \frac{2\pi j}{n} + \sin \theta \sin \frac{2\pi j}{n} \right|, \right. \right.$$

$$\left. \left. \left| \cos \theta \cos \frac{2\pi j}{n} - \sin \theta \sin \frac{2\pi j}{n} \right| \right) \right)$$

$$= \min_{\theta} \min_{\psi_1 \in \mathcal{S}_{d-1}} \max_{\psi_2 \in \mathcal{D}_{n,d-1}} |\langle \psi_1 | \psi_2 \rangle| \left(\max_{j=0,\ldots,n-1} \max \left(\left| \sin(\frac{2\pi j}{n} + \theta) \right|, \left| \sin(\frac{2\pi j}{n} - \theta) \right| \right) \right)$$

$$= \left(\min_{\theta} \max_{j=0,\ldots,n-1} \max \left(\left| \sin(\frac{2\pi j}{n} + \theta) \right|, \left| \sin(\frac{2\pi j}{n} - \theta) \right| \right) \right) \left(\min_{\psi_1 \in \mathcal{S}_{d-1}} \max_{\psi_2 \in \mathcal{D}_{n,d-1}} |\langle \psi_1 | \psi_2 \rangle| \right)$$

$$\geq \cos \frac{\pi}{n} F(\mathcal{D}_{n,d-1}) = \cos^d \frac{\pi}{n}. \tag{107}$$

Therefore,

$$\delta(\mathcal{D}_{n,d}) = \max_{|y\rangle \in \mathcal{R}: \|x\|=1} \min_{|x\rangle \in \mathcal{D}_{n,d}} \||y\rangle\langle y| - |x\rangle\langle x|\|$$

$$= 2\sqrt{1 - F(\mathcal{D}_{n,d})^2} = 2\sqrt{1 - \cos^{2d} \frac{\pi}{n}}. \tag{108}$$

Thus, when $\delta(\mathcal{D}_{n,d}) = \delta$, we have

$$1 - \frac{\delta^2}{4} = \cos^{2d} \frac{\pi}{n} = (1 - \sin^2 \frac{\pi}{n})^d \geq (1 - \frac{\pi^2}{n^2})^d \geq 1 - \frac{d\pi^2}{n^2}. \tag{109}$$

Hence, $|\mathcal{D}_{n,d}| = n^d \leq \frac{(2\pi)^d d^{d/2}}{\delta^d}$.

6 Calculation complexity of tight CR bound within additive error ϵ

Now, we evaluate the calculation complexity of tight CR bound within additive error ϵ under the choice of $\mathcal{W}_R$ given in Section 5.4.

For further evaluation, we introduce the quantity:

$$C_1 := \frac{1}{2}\Big\|\sum_j (\rho^{-1/2}(D_j\rho)\rho^{-1/2})^2\Big\|^{1/2}, \tag{110}$$

Then, we have the following lemma;

Lemma 15.

$$C_2(a) \leq \|a\|C_1. \tag{111}$$

Proof. We define $A_j = \rho^{-1/2}(D_j\rho)\rho^{-1/2}$. Then, since $\|a\|^2 I \geq a^\dagger a$, we have

$$\|a\|^2 \sum_{j=1}^d A_j^2 = \|a\|^2 (A_1, \ldots, A_d) I \begin{pmatrix} A_1 \\ \vdots \\ A_d \end{pmatrix}$$

$$\geq (A_1, \ldots, A_d) a^\dagger a \begin{pmatrix} A_1 \\ \vdots \\ A_d \end{pmatrix} = \sum_{j=1}^d (\sum_{j'=1}^d a_j^{j'} A_{j'})^2. \tag{112}$$

Then, we have

$$\|a\|^2 C_1^2 = \|a\|^2 \Big\| \sum_{j=1}^d A_j^2 \Big\| \geq \Big\| \sum_{j=1}^d (\sum_{j'=1}^d a_j^{j'} A_{j'})^2 \Big\| = C_2(a)^2. \tag{113}$$

$\square$

When the pair a^*, S^* is the optimal solution of $S[D1, \mathcal{W}_R]$ with a weight matrix W, we define $\xi := n\|\Pi(a^*, S^*)\|(1 + \|a^*\|^2 C_1^2)$. Then, Theorem 13 and Lemma 15 guarantee that the error ϵ is upper bounded by $\xi\delta(W_R)$. Using this fact, we have the following theorem

Theorem 16. *Suppose that we choose $\mathcal{W}_R$ as $\mathcal{D}_{n,d}$ with $\epsilon\xi^{-1} = \delta(\mathcal{D}_{n,d})$. Then, the calculation complexity of the tight CR bound for probe states of size n within additive error ϵ is*

$$O\Big(\frac{\xi^d d^{d/2}}{\epsilon^d}(n^6 + d^2 n^3 + d^4 n^2)\Big). \tag{114}$$

Proof. We consider how $|\mathcal{W}_R|$ relates to the additive error of the CR bound. From Theorem 13, our lower bound to $O(D0)$ depends on the covering radius of $|\mathcal{W}_R|$. When the relation $\epsilon \geq \xi\delta$ holds, Lemma 15 and (94) guarantee that the error of CR bound is upper bounded by ϵ.

Since $\epsilon\xi^{-1} = \delta$, we have $m = |\mathcal{W}_R| \leq \frac{(2\pi)^d d^{d/2}}{\delta^d} = \frac{(2\pi\xi)^d d^{d/2}}{\epsilon^d}$. Hence, substituting the above value to m in (72), the overall complexity of the algorithm is

$$O\Big(\frac{\xi^d d^{d/2}}{\epsilon^d}(n^6 + d^2 n^3 + d^4 n^2)\Big). \tag{115}$$

$\square$

Since Theorem 16 evaluates the calculation complexity of the tight CR bound within additive error ϵ by using the value x, we need to evaluate the quantity $\|\Pi(a^{**}, S^{**})\|$, where (a^{**}, S^{**}) denotes the optimal solution of $D0$. The following theorem gives its upper bound.

Theorem 17. *We denote the SLD Fisher information matrix by J. We have*

$$\|\Pi(a^{**}, S^{**})\|$$
$$\leq \|G\| + \frac{d}{2\sqrt{\mathrm{Tr}G^{1/2}J^{-1}G^{1/2}}} + \frac{d}{4\mathrm{Tr}G^{1/2}J^{-1}G^{1/2}}, \tag{116}$$

and

$$\|a^{**}\| \leq \mathrm{Tr}a^{**} \leq \frac{d}{2\sqrt{\mathrm{Tr}G^{1/2}J^{-1}G^{1/2}}}. \tag{117}$$

Finally, we evaluate the calculation complexity of the tight CR bound within additive error ϵ by Theorems 16 and 17. Since a^{**}, S^{**} are close to $a^*(\mathcal{W}_R), S^*(\mathcal{W}_R)$, $\|a^*(\mathcal{W}_R)\|$ and $\|\Pi(a^*(\mathcal{W}_R), S^*(\mathcal{W}_R))\|$ have the same order as $\|a^{**}\|$ and $\|\Pi(a^{**}, S^{**})\|$, respectively. We assume that G is the identity matrix. Combining Theorems 16 and 17, we find that

$$\xi = O\left(nC_1^2(\frac{d}{2\sqrt{\mathrm{Tr}J^{-1}}})^2(\frac{d}{2\sqrt{\mathrm{Tr}J^{-1}}} + \frac{d}{4\mathrm{Tr}J^{-1}})\right). \tag{118}$$

Hence, the calculation complexity is upper bounded as

$$O\left(\frac{\left(nC_1^2(\frac{d}{2\sqrt{\mathrm{Tr}J^{-1}}})^2(\frac{d}{2\sqrt{\mathrm{Tr}J^{-1}}} + \frac{d}{4\mathrm{Tr}J^{-1}})\right)^d d^{d/2}}{\epsilon^d}(n^6 + d^2n^3 + d^4n^2)\right). \tag{119}$$

For simplicity, we consider the case when J is the identity matrix. The value $\mathrm{Tr}J^{-1}$ is d. The above value is simplified as

$$O\left(\frac{(nC_1^2 d^{3/2})^d d^{d/2}}{\epsilon^d}(n^6 + d^2n^3 + d^4n^2)\right)$$
$$= O\left(\frac{1}{\epsilon^d}n^d C_1^{2d}d^{2d}(n^6 + d^2n^3 + d^4n^2)\right). \tag{120}$$

7   Numerical lower bound for $S(D1)$

Here we numerically find a lower bound for $S(D1)$. Here n denotes the dimension of $\mathcal{H}$, and d is the number of parameters that we estimate. Here, we consider the metrology problem with probe state

$$\rho = I_n/n, \tag{121}$$

where I_n denotes an identity matrix of size n. We set the number of parameters to estimate to be and $d = 2$. Now let us define the algorithm `MakeRandomDs`(n) that generates size n random traceless matrices D_1 and D_2.

Algorithm `MakeRandomDs`
$[D_1, D_2] = $`MakeRandomDs`$(n)$

1. For $i = 1, 2$:

2.   Make a random matrix M of size n, with each matrix element chosen independently from the uniform distribution on $[0, 1]$.

3.   Set $D = M^\dagger M$.

4.   Set $D_i = D - \mathrm{Tr}D/n$.

We can interpret this as corresponding to a quantum model with true parameters θ_1 and θ_2, where if x is in the neighborhood of $\theta = (\theta_1, \theta_2)$, we have $\rho((x_1, x_2)) \approx I_n/n + D_1(x_1 - \theta_1) + D_2(x_2 - \theta_2)$. Physically, this could correspond to quantum parameter estimation on a set of states that are close to maximally mixed, which describes the scenario where noise dominates the quantum system.

Now we describe how we make the set $\mathcal{W}_R$. First we make the algorithm $\texttt{circlepoints}(N)$ which takes as input a positive integer N, and returns a set $\textsf{calS}$.

Algorithm `circlepoints`
$\textsf{calS} = \texttt{circlepoints}(N)$

1. Set $\textsf{calS} = $ a matrix with N rows and two columns, with every entry equal to zero.

2. For $j = 1 : N$

3.   Set $\theta = 2\pi j/N$

4.   Set $\textsf{calS}(j, 1) = \cos(\theta)$

5.   Set $\textsf{calS}(j, 2) = \sin(\theta)$

Next, we make $\mathcal{W}_R$ using the following algorithm. [3]

Algorithm `makeWR`
$\mathcal{W}_R = \texttt{makeWR}(N)$

1. Set $N = 70, k = 100$ and $\phi_0 = 1.2$.

2. Set $\mathcal{W}_R = \emptyset$

3. for $\textsf{idx} = 1 : (k + 1)$

4.   Set $\textsf{currN} = \max(\lceil (\textsf{idx}/k)^{1/4} N \rceil, 20)$

5.   Set $\textsf{calS} = \texttt{circlepoints}(\textsf{currN})$

6.   Set $j = \textsf{idx} - 1$

7.   Set $\phi = \phi_0 j/k$

8.   Set $\mathbf{1}$ as a column vector of ones that has the same number of rows as $\textsf{calS}$

9.   Set W as the matrix with three columns $[\cos(\phi)\mathbf{1}, \sin(\phi)\textsf{calS}]$ with N rows.

10.   Add every row of W into the set $\mathcal{W}_R$.

---

[3]We export information about $\mathcal{W}_R$ in a csv file `WR.csv` available on request.

We find that $\mathcal{W}_R$ has 5750 points.

For each $n = 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17$, we run Algorithm `MakeRandomDs` 50 times, and use the same $\mathcal{W}_R$. For each (ρ, D_1, D_2) that we obtain, we

1. calculate the NH bound, the HN bound, $S[D1, \mathcal{W}_R]$ and $\underline{S}[D1, \mathcal{W}_R]$.

To find $S[D1, \mathcal{W}_R]$ and $\underline{S}[D1, \mathcal{W}_R]$, we first solve $[D1, \mathcal{W}_R]$ and find its optimal solution (a^*, S^*). The optimal value is $S[D1, \mathcal{W}_R] = \mathrm{Tr}\,a^* + \mathrm{Tr}\,S^*$, and this is our upper bound to $S(D1)$. To evaluate $\underline{S}[D1, \mathcal{W}_R]$, which is our lower bound to $S(D1)$, we first evaluate $X^* = \Pi(a^*, S^*)(81)$. Next, we evaluate $c_2 = C_2(a^*)$ (80) and $\phi_0^* = \tan^{-1}(c_2)$. Our next step is to numerically approximate κ as defined in (82).

We implement our approximation of κ in Theorem 11 by the following method.

1. Set myrange $= \{-c_2 + 2c_2 j/999 | j = 0, \dots, 999\}$. (We can implement this in MatLab using linspace$(-c_2, c_2, 1000)$).

2. Set $\kappa = -\infty$.

3. For x in myrange:

4.   For y in myrange:

5.   Set $w = (1, x, y)$ as a column vector.

6.   If $\|(x, y)\| \leq c_2$:

7.   Set $m = \lambda_{\min}(\langle w | X^* | w \rangle)$, where $\lambda_{\min}(\cdot)$ gives the minimum eigenvalue of a Hermitian matrix.

8.   If $m < -\kappa$, set $\kappa = -m$.

This subroutine evaluates the smallest eigenvalue of the matrices $\langle (1, y) | X^* | (1, y) \rangle$ amongst the no more than 10^6 points of $y \in \mathbb{R}^2$, and $\|y\| \leq c_2$. Then we set our lower bound to $S(D1)$ to be $\underline{S}[D1, \mathcal{W}_R] = \mathrm{Tr}\,a^* + \mathrm{Tr}\,S^* - n\kappa$. If $\underline{S}[D1, \mathcal{W}_R] > S(P2)$, we set isgap $= 1$. Otherwise, we set isgap $= 0$.

Since (D_1, D_2) is generated randomly by `MakeRandomDs`, isgap is a binary random variable. We define f_n as the probability that isgap takes the value 1. Since f_n is an unknown parameter, we estimate f_n from our 50 numeric experiments that independently generates (D_1, D_2), which determines isgap in the above method. Let g_n be the number of times isgap $= 1$ out of the 50 experiments. We plot the values of g_n along with their error bars in in Fig. 5.

For fixed n, we plot the largest values of $\underline{S}[D1, \mathcal{W}_R]/S(P2)$ among our 50 experiments in Fig. 6. Values of $\underline{S}[D1, \mathcal{W}_R]/S(P2)$ that are strictly larger than 1 illustrate a positive gap between $S(P1)$ and $S(P2)$. We also plot the corresponding values of $S(P4)/S(P2)$ in Fig. 6.

As expected, when $n = 3$, $\underline{S}[D1, \mathcal{W}_R]$ is never greater than the NH bound given by $S(P2)$. However, for $n = 4, 5, 6, 7, 8, 9$, in the majority of our numerical experiments, $\underline{S}[D1, \mathcal{W}_R]$ exceeds the NH bound.

8 Applications

8.1 Learning parameters of Hamiltonian models

When one has a physical system on many qubits, the underlying Hamiltonian can always be written as $H = \sum_k a_k P_k$ where P_k are multiqubit Pauli operators. The Pauli operators
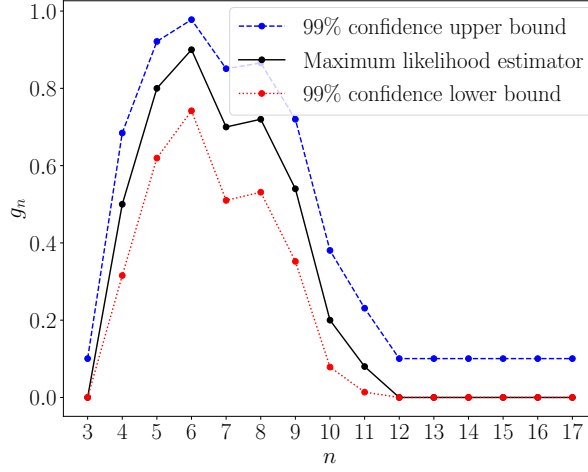
Figure 5: For each value of $n = \dim(\mathcal{H})$, we randomly generate 50 independent pairs of derivatives of probe states (D_1, D_2) according to Algorithm `MakeRandomDs`. Let β_n count how many of the 50 independent pairs (D_1, D_2) have $\underline{S}[D1, \mathcal{W}_R]/S(P2) > 1$. Then set $g_n = \beta_n/50$. Here g_n is a maximum likelihood estimator of f_n which is the asymptotic fraction of experiments which would have $\underline{S}[D1, \mathcal{W}_R]/S(P2) > 1$. When g_n is close to 1, this indicates that it is very likely that all random derivatives of probe states have a $C[G] > S(P2)$. Hence, the numerical evidence show that there is very often a gap between the tight bound and the NH bound, demonstrating the benefit of using the tight bound over the NH bound. We calculate the 99% confidence interval of our estimator g_n according to the Clopper–Pearson method (using the function binofit in MatLab.)
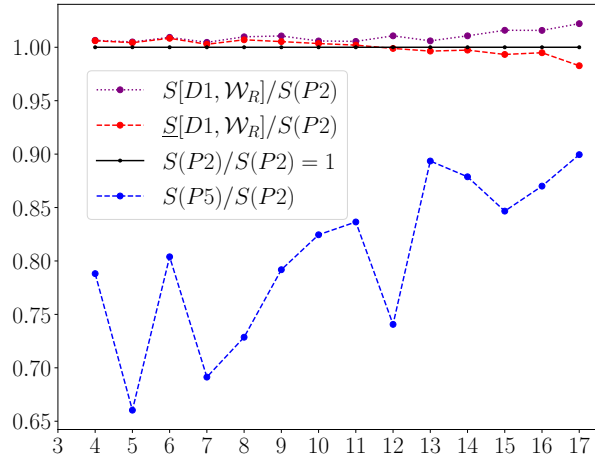
Figure 6: For each of our 50 numerical experiments for $n = \dim(\mathcal{H}) = 4, 5, \ldots, 16, 17$, we find the maximum $\underline{S}[D1, \mathcal{W}_R]/S(P2)$ and plot it. We also plot the corresponding values of $S[D1, \mathcal{W}_R]/S(P2)$ and $S(P5)]/S(P2)$. In Fig. 5, we see that there is very often a gap between the tight bound and the NH bound. Here, this figure shows us what the magnitude of this gap is. Since the gap is quite small, we can see that there is numerical evidence that the NH bound is a good approximation to the tight bound.

P_k are known, but often, we do not know the precise values of a_k. In a physical system with n qubits, there are up to 4^n coefficients for a_k to determine. However, in realistic physical systems comprising of qubits, the predominant interactions are between pairs of qubits, and the interaction strength decreases as the separation of the qubits increases. When the interaction between the qubits is dominated by nearest-neighbor interactions, and the qubits are arranged in a 2D or 3D array, the vast majority of the coefficients a_k have absolute values close to zero. In fact in this case, the number of unknown coefficients a_k that we like to estimate will be polynomial in n. This can arise for instance in physical systems where only spatially adjacent qubits have a non-negligible interaction. For example, we could have Heisenberg model with qubits interacting according to low-dimensional graph [51], and the goal is to find the coefficients of the non-negligible interaction terms.

In general, given the Hamiltonian $H = \sum_{k=1}^{d} a_k P_k$, we like to estimate the parameters $a_1, \ldots, a_d$ using a initial n-qubit probe state ρ. After we initialize the physical system as the probe state ρ, we allow the physical system to naturally evolve with time according to the quantum channel $\mathcal{N}_{a_1,\ldots,a_d}$. In the noiseless setting, the quantum channel $\mathcal{N}_{a_1,\ldots,a_d}$ describes evolution of quantum systems according to the Schrödinger equation, and is given by

$$\mathcal{N}_{a_1,\ldots,a_d}(\rho) = \exp(-i\sum_{k=1}^{d} a_k P_k)\rho\exp(i\sum_{k=1}^{d} a_k P_k). \tag{122}$$

In more realistic scenarios however, the quantum channel $\mathcal{N}_{a_1,\ldots,a_d}$ that models the evolution of the initial probe state, is not necessarily a unitary channel. In general, $\mathcal{N}_{a_1,\ldots,a_d}$ can be described using Lindblad's formalism [52], where the channels are generated by exponentiating Liouville operators. Such models have been considered for instance in Ref [53]. We can see that in both settings with and without noise, the channel $\mathcal{N}_{a_1,\ldots,a_d}$ is differentiable with respect to the parameters $a_1, \ldots, a_d$. In the language of multiparameter quantum estimation, we can define the model

$$\mathcal{M} = \{\mathcal{N}_{a_1,\ldots,a_d}(\rho) : a_1, \ldots, a_d \in \mathbb{R}\}, \tag{123}$$

which is differentiable set of quantum states. The task at hand is then to find the optimal quantum estimator that lets us learn the values of $a_1, \ldots, a_d$ while also minimizing the sum of the mean squared errors of $a_1, \ldots, a_d$.

Matsumoto [32] showed that when the model $\mathcal{M}$ is a set of pure states, then the corresponding tight bound is in fact equivalent to the HN bound. Since the noiseless setting corresponds to having $\mathcal{N}_{a_1,\ldots,a_d}$ as a unitary channel, having the initial probe state to be a pure state would make the model $\mathcal{M}$ comprise of only pure states. In this case, we can simply evaluate $S(P4)$, the HN bound, and know that $S(P4)$ is equal to the tight bound.

Now, if we consider the case where $\mathcal{N}_{a_1,\ldots,a_d}$ is a non-unitary channel which is generated by exponentiating Liouville operators, even if the initial probe state ρ is pure, the model $\mathcal{M}$ need not be a set of only pure states. In this scenario, we have no guarantee that the tight bound is equal to the HN bound. To find the optimal quantum estimator, we can nonetheless use our method for finding the asymptotically for the tight bound.

As a concrete example, consider a two-qubit Hamiltonian $H_{a,b} = aX \otimes X + bY \otimes Y$, where X and Y are the usual single qubit Pauli operators. Such a Hamiltonian $H_{a,b}$ models an XY Heisenberg interaction between a pair of qubits, and such interactions are ubiquitous in nature. Being able to estimate a and b is essential to characterize such Hamiltonians. We also consider having some dissipative jump operators L_j that describe noise in the

physical system. In this case, the Liouville operator is given by $\mathcal{L}_{a,b}$, where

$$\mathcal{L}_{a,b}(\rho) = -i[H_{a,b}, \rho] + \sum_j \gamma_j \left(L_j \rho L_j^\dagger - \frac{1}{2} \{ L_j^\dagger L_j, \rho \} \right), \tag{124}$$

$[H_{a,b}, \rho] = H_{a,b}\rho - \rho H_{a,b}$, $\{L_j^\dagger L_j, \rho\} = L_j^\dagger L_j \rho + \rho L_j^\dagger L_j$, and γ_j are real numbers. The quantum channel is

$$\mathcal{N}_{a,b}(\rho) = e^{\mathcal{L}_{a,b}}(\rho), \tag{125}$$

and the corresponding model is $\mathcal{M} = \{\exp(\mathcal{L}_{a,b}(\rho)) : a, b \in \mathbb{R}\}$. Our method can address such problems. As Section 7 illustrates, we are able to calculate tight two-sided bounds on the tight bound for such models with two-parameter estimation. We are also able to obtain corresponding near-optimal quantum estimators for the tight bound. Hence, for this example, we are able to determine the optimal uncorrelated measurement strategy to estimate a and b with the minimum mean-square error, which will be useful in characterizing interactions in spin-systems.

8.2   3D field sensing

The mathematical structure of the Hamiltonian learning problem in Section 8.1 also allows us to discuss the multiparameter quantum metrology of field sensing.

In the research area of 'field sensing' [54], we have a classical field that interacts with a physical system comprises of qubits, and the classical field interacts with each qubit in exactly the same way. The classical field is a 3D-vector (x, y, z) with three real components. The interaction Hamiltonian on n qubits is given by

$$\sum_{j=1}^n (xX_j + yY_j + zZ_j), \tag{126}$$

where X_j, Y_j and Z_j denote multiqubit Pauli operators that apply the Pauli X, Y and Z respectively on the jth qubit, and apply the identity operator on all other qubits. For example when we have only two qubits, the Hamiltonian is

$$H = x(X \otimes I + I \otimes X) + y(Y \otimes I + I \otimes Y) + z(Z \otimes I + I \otimes Z). \tag{127}$$

Just as in Section 8.1, we consider the evolution operator to be generated by a Liouville operator $\mathcal{L}$, to take into account the effects of noise in the quantum system. Namely, we consider $\mathcal{L}$ of the same form as (124), except that we replace the Hamiltonian H with what we consider in (127). In this case, the model is

$$\mathcal{M} = \{e^{\mathcal{L}}(\rho) : x, y, z \in \mathbb{R}\}, \tag{128}$$

where ρ is a fixed two-qubit density matrix. We could for example consider ρ as a GHZ state. Then we can numerically calculate upper and lower bounds on the tight bound for such models, and furthermore also also obtain corresponding near-optimal quantum estimators for the tight bound. Hence, it is possible to use our algorithm for determining optimal estimators in multiparameter quantum metrology for the extremely practical problem of field sensing.

9 Discussion

To summarize our results, we have presented a unified viewpoint to understand existing bounds for the Cramer-Rao bound from the viewpoint of conic linear programming. However in general, existing bounds such as the NH bound and HCR bound, are not the tight Cramer-Rao bound. Therefore, we proposed an algorithm to calculate the tight bound by using conic linear programming. Therefore, the main contribution of our paper is to propose an algorithm to calculate the tight bound. Our method also enables us to find the optimal POVM for the parameter estimation.

In more detail, we have characterized three types of bounds, HN bound, NH bound, and the tight bound as the solution of conic linear program with different cones in the same linear space. Their size relationship is characterized as the inclusion relation among their corresponding cones. Since the conic linear programing corresponding to the HN and NH bounds are given as an SDP, they can be efficiently solved. However, since the cone corresponding to the tight bound is the separable cone $\mathcal{S}_{SEP}$, i.e., the cone composed of positive semi-definite operators with separable form over a bipartite system, it is hard to solve the corresponding conic linear programing. In the second part of this paper, we have tackled the problem, i.e., how to efficiently solve this conic linear programing. For this aim, we have proposed a method to convert the conic linear programing with the separable cone $\mathcal{S}_{SEP}$ to an SDP with constraints labeled by unit vectors from a real vector space. We have derived upper and lower bounds on the tight bound by using the solution of the above converted SDP. From the optimal solution of the SDP, we give a concrete way to derive a corresponding uncorrelated measurement strategy which is asymptotically optimal. In addition, we have proposed the method to choose a strategic subset of the above unit vectors from design theory. Then, we have given the calculation complexity of the tight bound within an additive error of ϵ when the above method is applied. Then, we have numerically applied our method to our examples. In these examples, we have numerically shown that the tight bound is strictly larger than the NH bound, which shows the existence of the gap between the tight bound and the NH bound.

In fact, several problems of quantum information can be written with the separable cone $\mathcal{S}_{SEP}$. For example, the communication value can be calculated as a conic linear programming with the separable cone $\mathcal{S}_{SEP}$ [42, Proposition 2]. Also, the existence of entanglement can be written as a problem by using the separable cone $\mathcal{S}_{SEP}$ [14, 15]. We can expect that our method to approximately solve conic linear programing with the separable cone $\mathcal{S}_{SEP}$ can be extended to the above problems, and other similar problems [17, 18]. Such extensions are interesting for future study.

Further, after completing this research, the reviewer informed us about the method discussed in the reference [55] to calculate the conic linear programing with separable cone. The comparison between our method and this approach is an interesting topic, but beyond the scope of this paper. Also, as another application of our method includes the simultaneous estimation of phase and dephasing which was discussed in [56]. There is a possibility that our presented approach provides a new insight into this problem. This kind of research is another interesting future study.

Another open problem is whether our conic programming framework can also be used to unify multiparameter quantum metrology problems where instead of minimizing the mean square error, one minimizes a different regret function. For instance in the reference [36], minimization of an information regret function was considered. It will be interesting to see if our framework also applies in that setting.

Acknowledgement

A Applications of $C[G] = S(D0)$ (25)

A.1 One-parameter case

To consider how to choose a and S, we focus on the one-parameter case based on the idea by [37, 13]. Let L be the SLD, and J be the SLD Fisher Information. We simplify $(AB + BA)/2$ by $A \circ B$ Assume that $g(x,x) = x^2$. The condition (23) can simplified as

$$
\begin{aligned}
0 \leq & x^2 \rho - a x L \circ \rho - S \\
= & (x - \frac{a}{2}L)\rho(x - \frac{a}{2}L) - \frac{a^2}{4}L\rho L - S.
\end{aligned}
\tag{129}
$$

Hence, we have

$$
-\frac{a^2}{4}L\rho L \geq S
\tag{130}
$$

Now, we choose $S = -\frac{a^2}{4}L\rho L$. The value to be minimized is $a - \text{Tr}\frac{a^2}{4}L\rho L = a - \frac{a^2}{4}J = -(\frac{\sqrt{J}}{2}a - \frac{1}{\sqrt{J}})^2 + 1/J$. When $a = \frac{2}{J}$, the maximum $1/J$ is realized.

Indeed, the LHS of (26) is zero when M is the spectral decomposition of the Hermitian matrix $\frac{a}{2}L$. Hence, we can find that this measurement achieves the minimum value $1/J$.

However, in the general case, there is no simultaneous spectral decomposition of all SLDs. Hence, the problem is very difficult.

A.2 qubit case

Next, we focus on the qubit case based on the idea by [37, 13]. We choose the coordinate such that the SLD Fisher information matrix is identical. By applying suitable orthogonal transformation on the parameter space, the matrix G is a diagonal matrix such that $G_{i,j} = g_j \delta_{i,j}$ with $g_i \geq 0$. Let L_i be the SLD of the i-th parameter. Under the above condition, in the qubit case, we can check the following condition, as shown later:

$$
\frac{1}{2}(L_i\rho L_j + L_j\rho L_i) = I - \rho.
\tag{131}
$$

Since only the condition (131) is essential in the discussion of this section, the discussion of this section can be extended to the case of GPT (general probability theory) based on Lorentz cone [57, Appendix F].

We set $a_i^j = a\sqrt{g_j}\delta_i^j$ and $S = -\frac{a^2}{4}(I - \rho)$. The condition (23) can simplified as

$$\sum_j g_j(x^j)^2\rho - a\sum_j \sqrt{g_j}x^j L_j \circ \rho - S$$

$$= \left(\sqrt{\sum_j g_j(x^j)^2} - \frac{a}{2\sqrt{\sum_j g_j(x^j)^2}}\sum_j \sqrt{g_j}x^j L_j\right)\rho$$

$$\cdot \left(\sqrt{\sum_j g_j(x^j)^2} - \frac{a}{2\sqrt{\sum_j g_j(x^j)^2}}\sum_j \sqrt{g_j}x^j L_j\right)$$

$$\geq 0. \tag{132}$$

Hence, the objective value of the dual problem is $\text{Tr}\, a + \text{Tr}\, S = a\sum_j \sqrt{g_j} - \frac{a^2}{4} = -\frac{1}{4}(a - 2\sum_j \sqrt{g_j})^2 + (\sum_j \sqrt{g_j})^2$. Hence, we find that $(\sum_j \sqrt{g_j})^2$ is a lower bound for CR bound. Moreover, we expect that $(\sum_j \sqrt{g_j})^2$ is the solution with $a = 2\sum_j \sqrt{g_j}$.

This expectation can be checked by constructing a locally unbiasd estimator to realize the equality.

We denote the spectral decomposition of L_j as $\sum_k s_{j,k}E_{j,k}$. Define the vector $x(j,k) \in \mathbb{R}^d$ as follows. It has only non-zero element in the j-th entry. Its j-th entry is $s_{j,k}\frac{\sum_{j'} \sqrt{g_{j'}}}{\sqrt{g_j}}$. Then, we define the locally unbiasd estimator as follows. It takes the outcomes in the set $\{x(j,k)\}_{j,k}$. The POVM $\{M_{j,k}\}$ is defined as $M_{j,k} := \frac{\sqrt{g_j}}{\sum_{j'} \sqrt{g_{j'}}}E_{j,k}$. Here, we multiply the probabilistic factor $\frac{\sqrt{g_j}}{\sum_{j'} \sqrt{g_{j'}}}$ in the POVM element. Hence, to satisfy the locally unbiasedness condition, we need to multiply its inverse $\frac{\sum_{j'} \sqrt{g_{j'}}}{\sqrt{g_j}}$ in the measurement outcome. Hence, we can check that this POVM satisfies the condition for the locally unbiasedness condition.

In fact, when $x = x(j,k)$,

$$\sqrt{\sum_j g_j(x(j,k)^j)^2}$$

$$- \frac{2\sum_{j'} \sqrt{g_{j'}}}{2\sqrt{\sum_j g_j(x(j,k)^j)^2}}\sum_j \sqrt{g_j}x(j,k)^j L_j$$

$$= \left(\left(\sqrt{g_j}s_{j,k}\frac{\sum_{j'} \sqrt{g_{j'}}}{\sqrt{g_j}}\right) - \left(\sum_{j'} \sqrt{g_{j'}}\right)\text{sgn}\, s_{j,k}L_j\right)$$

$$= (\text{sgn}\, s_{j,k})\left(\sum_{j'} \sqrt{g_{j'}}\right)\left(s_{j,k} - L_j\right). \tag{133}$$

Hence, we find that

$$0 = \sum_{j,k} \text{Tr}\left(\sum_j g_j(x(j,k)^j)^2\rho\right.$$

$$\left. - a\sum_j \sqrt{g_j}x(j,k)^j L_j \circ \rho - S\right)M_{j,k}. \tag{134}$$

This equation shows the lower bound $(\sum_j \sqrt{g_j})^2$ is attainable.

Proof of (131). We choose σ_i as

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{135}$$

Without loss of generality, we can assume that $\rho = \frac{1}{2}(I + \alpha(\mu)\sigma_3)$ with $0 \le a \le 1$.

Notice that the SLD Fisher information matrix is identical. The condition (131) is covariant. That is, if the condition (131) under a coordinate, it holds even with another coordinate converted from orthogonal transformation. Hence, it is sufficient to check the condition (131) only under a specific coordinate.

Now, we consider the case that

$$L_1 = \sigma_1, \quad L_2 = \sigma_2, \quad L_3 = \frac{1}{\sqrt{1 - \alpha(\mu)^2}}(-\alpha(\mu)I + \sigma_3). \tag{136}$$

In this case, the SLD Fisher information matrix is identical, and the condition (131) holds. □

B Linear programming with general cone

We consider conic linear programming (EP) with a general cone. Let $\mathcal{X}$ and $\mathcal{Z}$ be topological real vector spaces. Let $\mathcal{P} \subset \mathcal{X}$ be a cone. Let A be a linear function from $\mathcal{X}$ to $\mathcal{Z}$. Given $c \in \mathcal{X}^*$ and $b \in \mathcal{Z}$, we consider the following minimization.

$$EP := \min_{x \in \mathcal{P}} \{c(x) | Ax = b\}. \tag{137}$$

As the duality problem, we consider the following maximization.

$$EP^* := \max_{w \in \mathcal{Z}^*} \{w(b) | - A^*w + c \in \mathcal{P}^*\} \tag{138}$$

When the pair of $x \in \mathcal{P}$ and $w \in \mathcal{Z}^*$ satisfy the constraints, we have $c(x) - w(b) = (-A^*w + c)(x) \ge 0$. Hence, the inequality

$$EP \ge EP^* \tag{139}$$

holds. The difference $EP - EP^*$ is called the duality gap, and its existence/non-existence is not trivial in general. To discuss this problem, we define two sets:

$$\mathcal{F} := \{(c(x), Ax)\}_{x \in \mathcal{P}} \tag{140}$$
$$\mathcal{E} := \mathbb{R} \times \{b\}. \tag{141}$$

Since $\mathcal{F} \cap \mathcal{E} \subset \overline{\mathcal{F}} \cap \mathcal{E}$ and $\overline{\mathcal{F}} \cap \mathcal{E}$ is a closed set, we have the inclusion relation

$$\overline{\mathcal{F} \cap \mathcal{E}} \subset \overline{\mathcal{F}} \cap \mathcal{E}. \tag{142}$$

The opposite inclusion relation is not trivial, and is related to the duality gap as follows.

Theorem 18 ([58, Proposition 4.2])**.** *When*

$$\overline{\mathcal{F} \cap \mathcal{E}} = \overline{\mathcal{F}} \cap \mathcal{E}, \tag{143}$$

we have

$$EP = EP^* \tag{144}$$

When $\mathcal{X}$ and $\mathcal{Z}$ are finite-dimensional, the relation (142) holds. Hence, there is no duality gap.

C   Proof of Theorem 6

This appendix aims to show the non-existence of the duality gap in the respective problems Pl for $l = 1, 2, 3, 4, 5$ by using Theorem 18 based on the notations given in Appendix B. Since the condition (143) holds in the finite-dimensional case, Theorem 6 immediately follows from Theorem 18. This point is in contrast with the proof of Proposition 1 by the paper [13] because the paper [13] had to manage a difficulty in the possibility of the existence of the duality gap even with a finite-dimensional space $\mathcal{H}$. The difficulty in [13] was caused by handling the space of POVMs with $\Omega = \mathbb{R}^d$ as an infinite-dimensional linear vector space. Hence, in our proof, we encounter the difficulty related to the potential existence of the duality gap only with an infinite-dimensional space $\mathcal{H}$.

Step 1: Preparation and structure of this proof.

We define $\mathcal{X}$ as the set of bounded operators on $\mathbb{R}^{d+1} \otimes \mathcal{H}$. We define $\mathcal{Z}$ as the tensor product of $\mathbb{R}^{d^2}$ and the set of bounded operators on $\mathcal{H}$. Then, the linear map A is defined as the left sides of (36) and (37). The element $b \in \mathcal{Z}$ is defined as the right sides of (36) and (37). The element $c \in \mathcal{X}^*$ is defined as (38). Then, for $l = 1$, we choose $\mathcal{P}$ as $\mathcal{S}_{SEP}$. For $l = 2$, we choose $\mathcal{P}$ as $\mathcal{S}_P$. For $l = 3$, we choose $\mathcal{P}$ as $\mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_{PPT} \cap \mathcal{B}''$. For $l = 4$, we choose $\mathcal{P}$ as $\mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}''$. For $l = 5$, we choose $\mathcal{P}$ as $\mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}''_\rho$. We consider the operator norm in the space $\mathcal{F}$ when $\mathcal{H}$ is infinite-dimensional.

We denote the SLD Fisher information matrix by $J_{i,j}$, and its inverse matrix by $J^{i,j}$. We define $L^j := \sum_i J^{i,j} L_i$. Hence, we have

$$\mathrm{Tr} L^j D_j = \delta_i^j. \tag{145}$$

Due to Theorem 18, it is sufficient to show (143). In the following, we show (143). Given an element $s \in \mathbb{R}$ with the condition $(s, b) \in \overline{\mathcal{F}} \cap \mathcal{E}$, we choose a sequence $\{X_n\}$ in $\mathcal{X}$ such that $\{(c(X_n), AX_n)\}$ converges to (s, b). It is sufficient to show that there exists a sequence $\{\dot{X}_n\}$ in $\mathcal{P}$ such that $A\dot{X}_n = b$ and $c(\dot{X}_n)$ converges to s, which shows that $\overline{\mathcal{F}} \cap \mathcal{E} \subset \overline{\mathcal{F} \cap \mathcal{E}}$.

Step 2: Definition of a new element $\dot{X}_n \in \mathcal{P}$.

We define the $d \times d$ matrix $a^{(n)}$ with component $a_j^{(n),i} := \mathrm{Tr}(\frac{1}{2}(|0\rangle\langle i| + |i\rangle\langle 0|) \otimes D_j) X_n$ and the operator $B_n := \mathrm{Tr}_{\mathcal{R}} |0\rangle\langle 0| \otimes I_{\mathcal{H}} X_n$ on $\mathcal{H}$. We define

$$\hat{X}_n := (I_{\mathcal{R}} \otimes B_n^{-1/2}) X_n (I_{\mathcal{R}} \otimes B_n^{-1/2}) \in \mathcal{P} \tag{146}$$

for any $l = 1, 2, 3, 4, 5$. Then, we have $\mathrm{Tr}_{\mathcal{R}} |0\rangle\langle 0| \otimes I_{\mathcal{H}} \hat{X}_n = I_{\mathcal{H}}$. The convergence $AX_n \to b$ implies the relations $a^{(n)} \to I_{\mathcal{R}}$ and $B_n \to I_{\mathcal{H}}$ in the sense of the operator norm. We define the $d \times d$ matrix $\hat{a}^{(n)}$ by the components, $\hat{a}_j^{(n),i} := \mathrm{Tr}(\frac{1}{2}(|0\rangle\langle i| + |i\rangle\langle 0|) \otimes D_j) \hat{X}_n$. Since $B_n \to I_{\mathcal{H}}$ in the sense of the operator norm, $B_n^{-1/2} D_j B_n^{-1/2} \to D_j$ in the sense of the trace norm. Thus, we have

$$\hat{a}^{(n)} \to I_{\mathcal{R}}, \tag{147}$$

$$c(\hat{X}_n) \to s. \tag{148}$$

We define the vector $\dot{a}^{(n),i} \in \mathbb{R}^d$ as $\dot{a}_j^{(n),i} := \hat{a}_j^{(n),i} - \delta_i^j$. We define $\epsilon_n := \max_i \|\dot{a}^{(n),i}\|$. The relation (147) implies

$$\epsilon_n \to 0. \tag{149}$$

For $i = 1, \dots, d$, we denote the spectral decomposition of the operator

$$Y_{n,i} := -\frac{1}{\epsilon_n} \sum_j \dot{a}_j^{(n),i} L^j \tag{150}$$

by $\sum_u x_u E_u$.

Then, we define the matrix $X_{n,i} := \sum_u \sum_{i'',i'} \hat{\theta}^{i'}(u)\hat{\theta}^{i''}(u)|i''\rangle\langle i'| \otimes E_u \in \mathcal{P}$, where $\hat{\theta}^{i'}(u) = \delta_{i',i}(1 - \delta_{i',0})u + \delta_{i',0}$ i.e.,

$$\hat{\theta}^{i'}(u) = \begin{cases} u & \text{for } i' = i \\ 1 & \text{for } i' = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{151}$$

Then, we have

$$\begin{aligned} X_{n,i} = \sum_{i'',i'} |i''\rangle\langle i'| \otimes \Big(\delta_{i',i}(1 - \delta_{i',0})Y_{n,i} + \delta_{i',0}I_{\mathcal{H}}\Big) \\ \cdot \Big(\delta_{i'',i}(1 - \delta_{i'',0})Y_{n,i} + \delta_{i'',0}I_{\mathcal{H}}\Big). \end{aligned} \tag{152}$$

We define

$$\dot{X}_n := (1 - d\epsilon_n)T_n\hat{X}_nT_n + \epsilon_n \sum_{i=1}^{d} X_{n,i} \tag{153}$$

where

$$T_n := \frac{1}{1 - d\epsilon_n}(I - |0\rangle\langle 0| \otimes I_{\mathcal{H}}) + |0\rangle\langle 0| \otimes I_{\mathcal{H}}. \tag{154}$$

Since $X_{n,i}$ belongs to $\mathcal{P}$, $\hat{X}_{n,i}$ belongs to $\mathcal{P}$ for any $l = 1, 2, 3, 4$. Hence, $T_n\hat{X}_nT_n$ belongs to $\mathcal{P}$ for any $l = 1, 2, 3, 4$. Thus, $\dot{X}_n$ belongs to $\mathcal{P}$ for any $l = 1, 2, 3, 4$.

Step 3: Asymptotic behavior of new element $\dot{X}_n \in \mathcal{P}$.

For $i = 1, \ldots, d$, the relations (145), (150), and (152) imply

$$\begin{aligned} \text{Tr}\Big[(\frac{1}{2}(|0\rangle\langle i'| + |i'\rangle\langle 0|) \otimes D_j)X_{n,i}\Big] &= -\frac{1}{\epsilon_n}\dot{a}_j^{(n),i}\delta_{i,i'} \\ \text{Tr}_{\mathcal{R}}\Big[(|0\rangle\langle 0| \otimes I_{\mathcal{H}})X_{n,i}\Big] &= I_{\mathcal{H}} \\ \text{Tr}\Big[(|i'\rangle\langle i''| \otimes \rho)X_{n,i}\Big] &= \delta_{i,i''}\delta_{i,i'}\frac{1}{\delta_n}\sum_{k,k'} J^{k,k'}\dot{a}_k^{(n),i}\dot{a}_{k'}^{(n),i}. \end{aligned} \tag{155}$$

Hence,

$$\begin{aligned} &\text{Tr}\Big[(\frac{1}{2}(|0\rangle\langle i'| + |i'\rangle\langle 0|) \otimes D_j)\dot{X}_n\Big] \\ &= \frac{1 - d\epsilon_n}{1 - d\epsilon_n}\hat{a}_j^{(n),i} - \frac{\epsilon_n}{\epsilon_n}\sum_i \dot{a}_j^{(n),i}\delta_{i,i'} = \delta_{i',j}, \end{aligned} \tag{156}$$

$$\text{Tr}_{\mathcal{R}}\Big[(|0\rangle\langle 0| \otimes I_{\mathcal{H}})\dot{X}_n\Big] = (1 - d\epsilon_n)I_{\mathcal{H}} + d\epsilon_n I_{\mathcal{H}} = I_{\mathcal{H}}, \tag{157}$$

$$\begin{aligned} &\text{Tr}\Big[(|i'\rangle\langle i''| \otimes \rho)\dot{X}_n\Big] \\ &= \frac{(1 - d\epsilon_n)^2}{1 - d\epsilon_n}\text{Tr}\Big[(|i'\rangle\langle i''| \otimes \rho)\hat{X}_n\Big] \\ &\quad + \epsilon_n \sum_i \delta_{i,i''}\delta_{i,i'}\frac{1}{\delta_n}\sum_{k,k'} J^{k,k'}\dot{a}_k^{(n),i}\dot{a}_{k'}^{(n),i} \\ &= (1 - d\epsilon_n)\text{Tr}\Big[(|i'\rangle\langle i''| \otimes \rho)\hat{X}_n\Big] \\ &\quad + \epsilon_n \sum_i \delta_{i,i''}\delta_{i,i'}\frac{1}{\delta_n}\sum_{k,k'} J^{k,k'}\dot{a}_k^{(n),i}\dot{a}_{k'}^{(n),i}. \end{aligned} \tag{158}$$

Using (158), we have

$$
\begin{aligned}
c(\dot{X}_n) &= \sum_{i',i''} G_{i',i''} \mathrm{Tr}\Big[\big(|i'\rangle\langle i''| \otimes \rho\big)\dot{X}_n\Big]\\
&= \sum_{i',i''} G_{i',i''} \bigg((1-d\epsilon_n)\mathrm{Tr}(|i'\rangle\langle i''| \otimes \rho)\hat{X}_n\\
&\quad + \epsilon_n \sum_i \delta_{i,i''}\delta_{i,i'} \frac{1}{\delta_n}\sum_{k,k'} J^{k,k'}\dot{a}_k^{(n),i}\dot{a}_{k'}^{(n),i}\bigg)\\
&= (1-d\epsilon_n)c(\hat{X}_n) + \frac{1}{\epsilon_n}\sum_i G_{i,i}\sum_{k,k'} J^{k,k'}\dot{a}_k^{(n),i}\dot{a}_{k'}^{(n),i}.
\end{aligned} \tag{159}
$$

Since

$$
\begin{aligned}
0 \le \frac{1}{\epsilon_n}\sum_i G_{i,i}\sum_{k,k'} J^{k,k'}\dot{a}_k^{(n),i}\dot{a}_{k'}^{(n),i} &\le \frac{d}{\epsilon_n}\|G\|\|J\| \max_i \|\dot{a}^{(n),i}\|^2\\
&= d\|G\|\|J\|\epsilon_n \to 0,
\end{aligned} \tag{160}
$$

using (148), (149), and (159), we have

$$
c(\dot{X}_n) \to \lim_{n\to\infty} c(\hat{X}_n) = s. \tag{161}
$$

Since $\dot{X}_n$ belongs to $\mathcal{P}$, the relations (156) and (157) guarantee $(c(\dot{X}_n), b) \in \mathcal{F} \cap \mathcal{E}$. Thus, (161) implies $(s, b) \in \overline{\mathcal{F} \cap \mathcal{E}}$. Hence, we obtain (143) for any $l = 1, 2, 3, 4$.

Step 4: Case of $l = 5$.

Now, we discuss the case of $l = 5$. We define the $d \times d$ antisymmetric matrix e^n as

$$
\begin{aligned}
e_j^{n,i} &:= \mathrm{Im}\,\mathrm{Tr}\Big[\big(|i\rangle\langle j| \otimes \rho\big)\dot{X}_n\Big]\\
&= \frac{1}{(1-d\epsilon_n)^2}\mathrm{Im}\,\mathrm{Tr}\Big[\big(|i\rangle\langle j| \otimes \rho\big)\hat{X}_n\Big]\\
&= \frac{1}{(1-d\epsilon_n)^2}\mathrm{Im}\,\mathrm{Tr}\Big[\big(|i\rangle\langle j| \otimes B_n^{-1/2}\rho B_n^{-1/2}\big)X_n\Big]
\end{aligned} \tag{162}
$$

for $i, j = 1, \ldots, d$. Since $B_n \to I_\mathcal{H}$ in the sense of the operator norm, $B_n^{-1/2}\rho B_n^{-1/2} \to \rho$ in the sense of the trace norm. Since $X_n \in \mathcal{P}$, $\mathrm{Im}\,\mathrm{Tr}(|i\rangle\langle j| \otimes \rho)X_n = 0$. Since $\epsilon_n \to 0$, the above discussions imply $e_j^{n,i} \to 0$. That is, we have

$$
\mathrm{Tr} G|e^{(n)}| \to 0. \tag{163}
$$

We define the operator $\tilde{X}_n$ as

$$
\tilde{X}_n := \dot{X}_n + \big(|e^{(n)}| - \sqrt{-1}e^{(n)}\big) \otimes I_\mathcal{H}. \tag{164}
$$

Since $|e^{(n)}| - \sqrt{-1}e^{(n)} \ge 0$ and $\dot{X}_n \in \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}''$, the operator $\tilde{X}_n$ belongs to $\mathcal{P} = \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H})_P \cap \mathcal{B}''_\rho$.

Since the $(0,0)$ and $(0,j)$ components of $|e^{(n)}| - \sqrt{-1}e^{(n)}$ are zero for $j = 1, \ldots, d$, the relation (156) implies

$$
\begin{aligned}
&\mathrm{Tr}\Big[\big(\tfrac{1}{2}(|0\rangle\langle i'| + |i'\rangle\langle 0|) \otimes D_j\big)\tilde{X}_n\Big]\\
&= \mathrm{Tr}\Big[\big(\tfrac{1}{2}(|0\rangle\langle i'| + |i'\rangle\langle 0|) \otimes D_j\big)\dot{X}_n\Big] = \delta_{i',j},
\end{aligned} \tag{165}
$$

and the relation (157) implies

$$\mathrm{Tr}_{\mathcal{R}}\left[(|0\rangle\langle 0| \otimes I_{\mathcal{H}})\tilde{X}_n\right] = \mathrm{Tr}_{\mathcal{R}}\left[(|0\rangle\langle 0| \otimes I_{\mathcal{H}})\dot{X}_n\right] = I_{\mathcal{H}}. \tag{166}$$

Since $c(\tilde{X}_n) = c(\dot{X}_n) + \mathrm{Tr}G|e^{(n)}|$, the combination of (161) and (163) implies

$$c(\tilde{X}_n) \to s. \tag{167}$$

Since $\dot{X}_n$ belongs to $\mathcal{P}$, the relations (165) and (166) guarantee $(c(\dot{X}_n), b) \in \mathcal{F} \cap \mathcal{E}$. Thus, (167) implies $(s, b) \in \overline{\mathcal{F} \cap \mathcal{E}}$. Hence, we obtain (143) for $l = 5$.

D   Proof for Theorem 7

The reference [13, Theorem 6] spends long pages to show Proposition 1 because the conic linear programming $P0$ is given as a conic linear programming in an infinite-dimensional space. However, due to Theorems 2 and 6, we can show Proposition 1 by showing the relation $S(D0) = S(D1)$.

Theorem 19. *We have*

$$S(D0) = S(D1). \tag{168}$$

Proof. The condition for $D0$ can be rewritten as

$$\min_{y \in \mathcal{H}} \min_{x \in \mathbb{R}^d} \left((x^T G x) y^\dagger \rho y - \sum_{i,j=1}^{d} a_i^j x^j y^\dagger D_j y - y^\dagger S y \right) \geq 0$$

$$\text{subject to } y^\dagger y = 1.$$

For $x \in \mathbb{R}^d$, we define $v(x) \in \mathcal{R}$ as $v(x) := \sum_{j=1}^{d} x_j|j\rangle + |0\rangle$. Then, the condition (169) is rewritten as

$$0 \leq \mathrm{Tr}\bigg[|v(x)\rangle\langle v(x)| \otimes |y\rangle\langle y|$$
$$\left(G \otimes \rho - \frac{1}{2} \sum_{1 \leq i,j \leq d} a_i^j(|0\rangle\langle i| + |i\rangle\langle 0|) \otimes D_j - |0\rangle\langle 0| \otimes S \right)\bigg] \tag{169}$$

for $y \in \mathcal{H}$ and $x \in \mathbb{R}^d$. This condition is rewritten as

$$0 \leq \mathrm{Tr}\bigg[|v\rangle\langle v| \otimes |y\rangle\langle y|$$
$$\left(G \otimes \rho - \frac{1}{2} \sum_{1 \leq i,j \leq d} a_i^j(|0\rangle\langle i| + |i\rangle\langle 0|) \otimes D_j - |0\rangle\langle 0| \otimes S \right)\bigg] \tag{170}$$

for $y \in \mathcal{H}$ and $v \in \mathcal{R}$ to satisfy $\langle v|0\rangle \neq 0$.

When (170) holds for $y \in \mathcal{H}$ and $v \in \mathcal{R}$ to satisfy $\langle v|0\rangle \neq 0$, it holds for $y \in \mathcal{H}$ and $v \in \mathcal{R}$ because of the following reason. We choose $v \in \mathcal{R}$ such that $\langle v|0\rangle = 0$. Then, we choose a sequence v_n such that $v_n \to v$ and $\langle v_n|0\rangle \neq 0$. Since (170) holds with v_n, by taking the limit, (170) holds even with v.

Therefore, the condition for $D0$ can be rewritten as the condition that the inequality (170) holds with any $y \in \mathcal{H}$ and any $v \in \mathcal{R}$. This condition is equivalent to (58). Hence, we conclude that the two dual problems $D0$ and $D1$ are equivalent. $\qquad\square$

E Evaluation of κ based on $\mathcal{W}_R$ constructed in Section 5.3

Since the lower bound of $S(D1)$ depends on the value κ, it is better to construct $\mathcal{W}$ by reflecting the structure of κ. Hence, we choose $\mathcal{W}_R$ as follows. We choose a subset $\mathcal{S} \subset S^{d-1}$. We define a subset $\mathcal{S}(\phi) := \{(\cos\phi, \sin\phi y)\}_{y\in\mathcal{S}}$. We choose $\phi_0 > 0$ as $\tan\phi_0 = C_2(a^*)$. Using k, we choose $\mathcal{W}_R$ as $\cup_{j=0}^{k}\mathcal{S}(\frac{\phi_0 j}{k})$.

Indeed, when $d = 2$, the choice of $\mathcal{S}$ is very easy because we can choose $\mathcal{S}$ as $\{(\cos\frac{2\pi l}{N}, \sin\frac{2\pi l}{N})\}_{l=1}^{N}$. In this case, we find that $\delta(\mathcal{S}) = \|(1,0) - (\cos 2\pi/N, \sin 2\pi/N)\| = \sqrt{2 - 2\cos 2\pi/N}$.

In this case, to evaluate the quantity κ, we employ Lemma 14. In this method, we need to evaluate the quantity $\delta(s, \mathcal{W}_R)$. For this aim, we define

$$\bar{\delta}(\mathcal{S}) := \max_{x\in\mathbb{R}^d: \|x\|=1} \min_{|w\rangle\in\mathcal{S}} \||x\rangle - |w\rangle\| \tag{171}$$

$$\bar{\delta}(s, \mathcal{W}_R) := \max_{x=(\cos\tan^{-1} s, (\sin\tan^{-1} s)z), z\in S^{d-1}} \min_{|w\rangle\in\mathcal{W}_R} \||x\rangle - |w\rangle\|. \tag{172}$$

To consider the relation between $\bar{\delta}(s, \mathcal{W}_R)$ and $\delta(s, \mathcal{W}_R)$, we focus on the relations

$$\||0\rangle\langle 0| - (\cos\theta|0\rangle + \sin\theta|1\rangle)(\cos\theta\langle 0| + \sin\theta\langle 1|)\|_1$$
$$= 2|\sin\theta| \tag{173}$$

$$\||0\rangle - (\cos\theta|0\rangle + \sin\theta|1\rangle)\| = \sqrt{2(1 - \cos\theta)} \cong |\sin\theta|. \tag{174}$$

Then, we define function $\eta : \sqrt{2(1 - \cos\theta)} \mapsto 2|\sin\theta|$. When $t > 0$ is small, the relations (173) and (174) implies $\eta(t) \cong 2t$. This function connects two quantities $\bar{\delta}(s, \mathcal{W}_R)$ and $\delta(s, \mathcal{W}_R)$ as $\delta(s, \mathcal{W}_R) = \eta(\bar{\delta}(s, \mathcal{W}_R))$. Then, we have the following lemma.

Lemma 20. *Under the above construction of $\mathcal{W}_R$, for $s = \tan\theta$, we have*

$$\bar{\delta}(s, \mathcal{W}_R)$$
$$\leq \min_{j\in\{0,1,\ldots,k\}} \left((\cos\theta - \cos\frac{\phi_0 j}{k})^2 \right.$$
$$\left. + \left(|\sin\theta|\bar{\delta}(\mathcal{S}) + \left|\sin\theta - \sin\frac{\phi_0 j}{k}\right| \right)^2 \right)^{\frac{1}{2}}. \tag{175}$$

This lemma implies that

$$\delta(s, \mathcal{W}_R)$$
$$\leq \min_{j\in\{0,1,\ldots,k\}} \eta\left(\left((\cos\theta - \cos\frac{\phi_0 j}{k})^2 \right.\right.$$
$$\left.\left. + \left(|\sin\theta|\bar{\delta}(\mathcal{S}) + \left|\sin\theta - \sin\frac{\phi_0 j}{k}\right| \right)^2 \right)^{\frac{1}{2}} \right). \tag{176}$$

Combining (176) and Lemma 14, we have

$$\kappa$$
$$\leq \|X^*\| \max_{s\in[0,C_2(a^*)]} (1 + s^2) \min_{j\in\{0,1,\ldots,k\}} \eta\left(\left((\cos\theta - \cos\frac{\phi_0 j}{k})^2 \right.\right.$$
$$\left.\left. + \left(|\sin\theta|\bar{\delta}(\mathcal{S}) + \left|\sin\theta - \sin\frac{\phi_0 j}{k}\right| \right)^2 \right)^{\frac{1}{2}} \right), \tag{177}$$

where $s = \tan\theta$.

Proof. Using the definition of the set $\mathcal{S}$, we get

$$\bar{\delta}(s, \mathcal{W}_R)$$

$$= \max_{x=(\cos\theta,\sin\theta z), z\in S^{d-1}} \min_{0\le j\le k} \min_{y\in\mathcal{S}} \left\| |x\rangle - (\cos\frac{\phi_0 j}{k}, \sin\frac{\phi_0 j}{k} y) \right\|$$

$$= \max_{z\in S^{d-1}} \min_{0\le j\le k} \min_{y\in\mathcal{S}} \left\| (\cos\theta, \sin\theta z) - (\cos\frac{\phi_0 j}{k}, \sin\frac{\phi_0 j}{k} y) \right\|$$

$$= \max_{\substack{z\in\mathbb{R}^d \\ \|z\|=1}} \min_{0\le j\le k} \min_{y\in\mathcal{S}} \left\| (\cos\theta, \sin\theta z) - (\cos\frac{\phi_0 j}{k}, \sin\frac{\phi_0 j}{k} y) \right\|$$

$$\le \min_{0\le j\le k} \max_{\substack{z\in\mathbb{R}^d \\ \|z\|=1}} \min_{y\in\mathcal{S}} \left\| (\cos\theta, \sin\theta z) - (\cos\frac{\phi_0 j}{k}, \sin\frac{\phi_0 j}{k} y) \right\|$$

$$= \min_{0\le j\le k} \max_{\substack{z\in\mathbb{R}^d \\ \|z\|=1}} \min_{y\in\mathcal{S}} \left((\cos\theta - \cos\frac{\phi_0 j}{k})^2 \right.$$

$$\left. + \left\| \sin\theta z - \sin\frac{\phi_0 j}{k} y \right\|^2 \right)^{\frac{1}{2}}, \tag{178}$$

where the inequality follows from the max-min inequality.

Now note that for $\|y\| = 1$,

$$\left\| \sin\theta z - \sin\frac{\phi_0 j}{k} y \right\|$$

$$= \left\| \sin\theta z - \sin\theta y + \sin\theta y - \sin\frac{\phi_0 j}{k} y \right\|$$

$$\le \left\| \sin\theta z - \sin\theta y \right\| + \left\| \sin\theta y - \sin\frac{\phi_0 j}{k} y \right\|$$

$$= |\sin\theta| \|z - y\| + \left| \sin\theta - \sin\frac{\phi_0 j}{k} \right|. \tag{179}$$

Hence

$$\bar{\delta}(s, \mathcal{W}_R)$$

$$\le \min_{0\le j\le k} \max_{\substack{z\in\mathbb{R}^d \\ \|z\|=1}} \min_{y\in\mathcal{S}} \left((\cos\theta - \cos\frac{\phi_0 j}{k})^2 \right.$$

$$\left. + \left(|\sin\theta| \|z - y\| + \left| \sin\theta - \sin\frac{\phi_0 j}{k} \right| \right)^2 \right)^{\frac{1}{2}}$$

$$= \min_{0\le j\le k} \left((\cos\theta - \cos\frac{\phi_0 j}{k})^2 \right.$$

$$\left. + \left(|\sin\theta| \bar{\delta}(\mathcal{S}) + \left| \sin\theta - \sin\frac{\phi_0 j}{k} \right| \right)^2 \right)^{\frac{1}{2}}. \tag{180}$$

$\square$

In the previous choice of $\mathcal{W}_R$, we use the same subset of $\mathcal{S}^{d-1}$ to construct $\mathcal{S}(\phi)$. We modify this construction. That is, to construct $\mathcal{S}(\frac{\phi_0 j}{k})$, depending on j, we choose a subset $\mathcal{S}_j \subset S^{d-1}$. We define a subset $\mathcal{S}(\frac{\phi_0 j}{k}) := \{(\cos\frac{\phi_0 j}{k}, \sin\frac{\phi_0 j}{k} y)\}_{y\in\mathcal{S}_j}$. We choose $\phi_0 > 0$ as $\tan\phi_0 = C_2(a^*)$. Using k, we choose $\mathcal{W}_R$ as $\cup_{j=0}^k \mathcal{S}(\frac{\phi_0 j}{k})$.

Then, we can show the following lemma in the same way as Lemma 20.

Lemma 21. *Under the above construction of $\mathcal{W}_R$, for $s = \tan\theta$, we have*

$$\bar{\delta}(s, \mathcal{W}_R)$$

$$\leq \min_{j \in \{0,1,\ldots,k\}} \left((\cos\theta - \cos\frac{\phi_0 j}{k})^2 \right.$$

$$\left. + \left(|\sin\theta|\bar{\delta}(\mathcal{S}_j) + \left| \sin\theta - \sin\frac{\phi_0 j}{k} \right| \right)^2 \right)^{\frac{1}{2}}. \tag{181}$$

This lemma implies that

$$\delta(s, \mathcal{W}_R)$$

$$\leq \min_{j \in \{0,1,\ldots,k\}} \eta\left(\left((\cos\theta - \cos\frac{\phi_0 j}{k})^2 \right.\right.$$

$$\left.\left. + \left(|\sin\theta|\bar{\delta}(\mathcal{S}_j) + \left| \sin\theta - \sin\frac{\phi_0 j}{k} \right| \right)^2 \right)^{\frac{1}{2}} \right). \tag{182}$$

Combining (182) and Lemma 14, we have

$$\kappa$$

$$\leq \|X^*\| \max_{s \in [0, C_2(a^*)]} (1 + s^2) \min_{j \in \{0,1,\ldots,k\}} \eta\left(\left((\cos\theta - \cos\frac{\phi_0 j}{k})^2 \right.\right.$$

$$\left.\left. + \left(|\sin\theta|\bar{\delta}(\mathcal{S}_j) + |\sin\theta - \sin\frac{\phi_0 j}{k}| \right)^2 \right)^{\frac{1}{2}} \right), \tag{183}$$

where $s = \tan\theta$.

F Proof of Theorem 17

F.1 Proof of (116)

Lemma 22. *For an Hermitian matrix X, we have*

$$\left\| \left(\frac{1}{2}(X\rho + \rho X) \right)^2 \right\| \leq \mathrm{Tr} X\rho X. \tag{184}$$

Proof. We choose a normalized vector $|\phi\rangle$ to satisfy

$$\left\| \left(\frac{1}{2}(X\rho + \rho X) \right)^2 \right\| = \langle\phi\left(\frac{1}{2}(X\rho + \rho X) \right)^2 |\phi\rangle. \tag{185}$$

We choose another normalized vector $|\psi\rangle$ as a constant times of $\left(\frac{1}{2}(X\rho + \rho X) \right)|\phi\rangle$. Then, we have

$$\langle\phi\left(\frac{1}{2}(X\rho + \rho X) \right)^2 |\phi\rangle = |\langle\psi\left(\frac{1}{2}(X\rho + \rho X) \right)|\phi\rangle|^2$$

$$= \left| \mathrm{Tr}\left(\frac{1}{2}(X\rho + \rho X) \right)|\phi\rangle\langle\psi| \right|^2. \tag{186}$$

We apply Schwartz inequality to the inner product $\langle X, Y \rangle := \text{Tr}\left[\frac{1}{2}(X\rho + \rho X)\right]Y^\dagger$. Then, we have

$$
\left|\text{Tr}\left[\frac{1}{2}(X\rho + \rho X)\right]|\phi\rangle\langle\psi|\right|^2
$$
$$
\leq \text{Tr}\left[\left(\frac{1}{2}(X\rho + \rho X)\right)X^\dagger\right]
$$
$$
\cdot \text{Tr}\left[\left(\frac{1}{2}(|\psi\rangle\langle\phi|\rho + \rho|\psi\rangle\langle\phi|)\right)|\phi\rangle\langle\psi|\right]
$$
$$
\leq \text{Tr}\left[X\rho X\right]\frac{1}{2}\text{Tr}\left[|\psi\rangle\langle\phi|\rho|\phi\rangle\langle\psi| + |\phi\rangle\langle\psi|\rho|\psi\rangle\langle\phi|)\right]
$$
$$
= \text{Tr}\left[X\rho X\right]\frac{1}{2}(\langle\phi|\rho|\phi\rangle + \langle\psi|\rho|\psi\rangle) \leq \text{Tr}\left[X\rho X\right]. \tag{187}
$$

$\square$

We define

$$
S(a) := \underset{S}{\text{argmax}}\{\text{Tr}S|(a, S) \text{ satisfies the condition (23).}\} \tag{188}
$$

Lemma 23. *Then, for a real number t, we have*

$$
S(ta) = t^2 S(a). \tag{189}
$$

Also, we have

$$
-S(a) \geq 0. \tag{190}
$$

Proof. The condition (23) with $x = 0$ implies (190).

Eq (189) can be shown as follows.

The condition (23) for (ta, S) is written as follows. Any vector $x \in \mathbb{R}^d$ satisfies

$$
(y^T G y)\rho - \sum_{i,j} a_i^j t x^i D_j - S \geq 0. \tag{191}
$$

By using a vector $y = t^{-1}x \in \mathbb{R}^d$, (191) is rewritten as

$$
t^2 (y^T G y)\rho - \sum_{i,j} a_i^j t^2 y^i D_j - S \geq 0. \tag{192}
$$

That is, it is rewritten as

$$
(y^T G y)\rho - \sum_{i,j} a_i^j y^i D_j - t^{-2}S \geq 0, \tag{193}
$$

which implies (189). $\square$

In the following, we assume that G is the identity matrix.

Lemma 24. *For a real number t, we have*

$$
-\text{Tr}S(a) \geq \frac{1}{d}\text{Tr}a^T J a. \tag{194}
$$

Proof. The condition (23) for $(a, S(a))$ is written as follows.

$$
\begin{aligned}
0 \leq & (x^i - \frac{1}{2}\sum_{j=1}^d a_i^j L_j)\rho(x^i - \frac{1}{2}\sum_{j=1}^d a_i^j L_j) \\
& + \sum_{i' \neq i}\left((x^{i'})^2\rho + \sum_j a_{i'}^j t^2 x^{i'} D_j\right) \\
& - \frac{1}{4}(\sum_{j=1}^d a_i^j L_j)\rho(\sum_{j'=1}^d a_i^{j'} L_{j'}) - S(a).
\end{aligned} \tag{195}
$$

We take the diagonalization as $\frac{1}{2}\sum_{j=1}^d a_i^j L_j = \sum_k c_k |\phi_k\rangle\langle\phi_k|$. We choose $x^{i'} = 0$ for $i' \neq i$ and $x^i = c_k$ in (195). Then, we have

$$
\begin{aligned}
0 \leq & (c_k - \frac{1}{2}\sum_{j=1}^d a_i^j L_j)\rho(c_k^i - \frac{1}{2}\sum_{j=1}^d a_i^j L_j) \\
& - \frac{1}{4}(\sum_{j=1}^d a_i^j L_j)\rho(\sum_{j'=1}^d a_i^{j'} L_{j'}) - S(a)\Big).
\end{aligned} \tag{196}
$$

Hence, we have

$$
\begin{aligned}
& \langle\phi_k|\left(-\frac{1}{4}(\sum_{j=1}^d a_i^j L_j)\rho(\sum_{j'=1}^d a_i^{j'} L_{j'}) - S(a)\right)|\phi_k\rangle \\
& \overset{(a)}{=}\langle\phi_k|\left((c_k - \frac{1}{2}\sum_{j=1}^d a_i^j L_j)\rho(c_k^i - \frac{1}{2}\sum_{j=1}^d a_i^j L_j)\right. \\
& \left. - \frac{1}{4}(\sum_{j=1}^d a_i^j L_j)\rho(\sum_{j'=1}^d a_i^{j'} L_{j'}) - S(a)\right)|\phi_k\rangle \\
& \overset{(b)}{\geq} 0,
\end{aligned} \tag{197}
$$

where (a) follows from the fact that c_k is the eigenvalue of $\frac{1}{2}\sum_{j=1}^d a_i^j L_j$ with the vector ϕ_k. Also, (b) follows from (196).

Taking the sum with respect to k, we have

$$
\begin{aligned}
& -\operatorname{Tr}\frac{1}{4}(\sum_{j=1}^d a_i^j L_j)\rho(\sum_{j'=1}^d a_i^{j'} L_{j'}) - \operatorname{Tr}S(a) \\
& = \sum_k \langle\phi_k|\left(-\frac{1}{4}(\sum_{j=1}^d a_i^j L_j)\rho(\sum_{j'=1}^d a_i^{j'} L_{j'}) - S(a)\right)|\phi_k\rangle \geq 0.
\end{aligned} \tag{198}
$$

Since (198) holds for $i = 1, \ldots, d$, we have

$$
-\operatorname{Tr}S(a) \geq \frac{1}{d}\sum_{i=1}^d \operatorname{Tr}\frac{1}{4}(\sum_{j=1}^d a_i^j L_j)\rho(\sum_{j'=1}^d a_i^{j'} L_{j'}) = \frac{1}{d}a^t J a. \tag{199}
$$

$\square$

Lemma 25.

$$
\|X(a, S(a))\| \leq \|G\| + (-d\operatorname{Tr}S(a))^{1/2} - \operatorname{Tr}S(a). \tag{200}
$$

Proof. We have

$$\left\| \frac{1}{2} \sum_{1 \leq i,j \leq d} a_i^j (|0\rangle\langle i| + |i\rangle\langle 0|) \otimes D_j \right\|^2$$

$$= \left\| \left(\frac{1}{2} \sum_{1 \leq i,j \leq d} a_i^j (|0\rangle\langle i| + |i\rangle\langle 0|) \otimes D_j \right)^2 \right\|$$

$$= \left\| \frac{1}{4} \left(|0\rangle\langle 0| \otimes \left(\sum_{i=1}^{d} (\sum_{j=1}^{d} a_i^j D_j)^2 \right) + \sum_{i'=1}^{d} |i'\rangle\langle i'| \otimes (\sum_{j=1}^{d} a_{i'}^j D_j)^2 \right) \right\|$$

$$\overset{(a)}{=} \left\| \sum_{i=1}^{d} (\sum_{j=1}^{d} a_i^j D_j)^2 \right\| \leq \sum_{i=1}^{d} \left\| (\sum_{j=1}^{d} a_i^j D_j)^2 \right\|$$

$$\overset{(b)}{\leq} \sum_{i=1}^{d} \mathrm{Tr}\left[(\sum_{j=1}^{d} a_i^j L_j) \rho (\sum_{j=1}^{d} a_i^j L_j) \right] = \mathrm{Tr}\, a^T J a \overset{(c)}{\leq} -d \mathrm{Tr} S(a), \tag{201}$$

where (a), (b), and (c) follow from the inequality $(\sum_{j=1}^{d} a_{i'}^j D_j)^2 \leq \sum_{i=1}^{d} (\sum_{j=1}^{d} a_i^j D_j)^2$ for $i' = 1, \ldots, d$, Lemma 22, and Lemma 24, respectively.

Hence,

$$\|X(a, S(a))\|$$

$$\leq \|G \otimes \rho\| + \left\| \frac{1}{2} \sum_{1 \leq i,j \leq d} a_i^j (|0\rangle\langle i| + |i\rangle\langle 0|) \otimes D_j \right\|$$

$$+ \||0\rangle\langle 0| \otimes S(a)\|$$

$$\leq \|G\| + \left\| \frac{1}{2} \sum_{1 \leq i,j \leq d} a_i^j (|0\rangle\langle i| + |i\rangle\langle 0|) \otimes D_j \right\| + \|S\|$$

$$\overset{(a)}{\leq} \|G\| + (-d\mathrm{Tr}S(a))^{1/2} - \mathrm{Tr}S(a), \tag{202}$$

where (a) follows from (190) and (201). $\qquad\square$

Lemma 26. *We have*

$$-\mathrm{Tr}S(a^{**}) \leq \frac{d}{4\mathrm{Tr}J^{-1}} \tag{203}$$

and

$$\mathrm{Tr}\, a^{**} \leq \frac{d}{2\mathrm{Tr}J^{-1}}. \tag{204}$$

Proof. We have

$$\max_{(a,S)} \{\mathrm{Tr}a + \mathrm{Tr}S | (a, S) \text{ satisfies the condition } (23).\}$$

$$= \max_{a} \mathrm{Tr}a + \mathrm{Tr}S(a)$$

$$= \max_{t} \max_{a} \{\mathrm{Tr}ta + \mathrm{Tr}S(ta) | \mathrm{Tr}a = 1\}$$

$$\overset{(a)}{=} \max_{t} \max_{a} \{t + t^2\mathrm{Tr}S(a) | \mathrm{Tr}a = 1\}$$

$$= \max_{a} \{\max_{t} t + t^2\mathrm{Tr}S(a) | \mathrm{Tr}a = 1\}$$

$$= \max_{a} \{\frac{-1}{4\mathrm{Tr}S(a)} | \mathrm{Tr}a = 1\}, \tag{205}$$

where (a) follows from Lemma 23.

In particular, when t equals $t_a := -\frac{1}{2\text{Tr}S(a)}$, the maximum $\max_t t + t^2\text{Tr}S(a)$ is realized. Then, since $\text{Tr}S(t_a a) \leq 0$, due to Lemma 23, we have

$$-\text{Tr}S(t_a a) = -t_a^2\text{Tr}S(a) = \frac{-1}{4\text{Tr}S(a)}. \tag{206}$$

Also, due to (205), using

$$a_* := \underset{a}{\text{argmax}}\left\{\frac{1}{4(\text{Tr}S(a))^2}\bigg|\text{Tr}a = 1\right\}, \tag{207}$$

we have

$$a^{**} = t_{a_*}a_*. \tag{208}$$

Hence, we have

$$-\text{Tr}S(a^{**}) \overset{(a)}{=} -\text{Tr}S(t_{a_*}a_*) \overset{(b)}{=} \frac{-1}{4\text{Tr}S(a_*)}$$

$$\overset{(c)}{=} \max_a\left\{\frac{-1}{4\text{Tr}S(a)}\bigg|\text{Tr}a = 1\right\} \overset{(d)}{\leq} \max_a\left\{\frac{d}{4\text{Tr}[a^T Ja]}\bigg|\text{Tr}a = 1\right\}, \tag{209}$$

where (a), (b), and (c) follow from (208), (206), and (207), respectively. In addition, (d) follows from the inequality

$$\frac{-1}{4\text{Tr}S(a)} \leq \frac{d}{4\text{Tr}[a^T Ja]}, \tag{210}$$

which is shown by Lemma 24.

We consider the inner product $\langle a, b\rangle := \text{Tr}[ab^T]$. Schwartz inequality for $J^{1/2}a$ and $J^{-1/2}$ implies that

$$\text{Tr}J^{-1}\text{Tr}[a^T Ja] \geq |\text{Tr}[J^{-1/2}J^{1/2}a]|^2 = |\text{Tr}a|^2. \tag{211}$$

The equality holds when a is a constant times of J^{-1}. Therefore, we have the relation

$$\max_a\left\{\frac{d}{4\text{Tr}[a^T Ja]}\bigg|\text{Tr}a = 1\right\} = \frac{d}{4\text{Tr}J^{-1}}. \tag{212}$$

The combination of (209) and (212) yields (203).

Also, (204) is shown as

$$\text{Tr}a^{**} \overset{(a)}{=} t_{a_*} \overset{(b)}{=} -\frac{1}{2\text{Tr}S(a_*)} \overset{(c)}{\leq} \frac{d}{2\text{Tr}[a_*^T Ja_*]} \overset{(d)}{\leq} \frac{d}{4\text{Tr}J^{-1}}, \tag{213}$$

where (b), (c), and (d) follows from the definition of t_a, (210), and (212), respectively. In addition, (a) follows from (208) and the relation $\text{Tr}a_* = 1$. $\qquad\square$

Proof of (116). When G is the identity matrix, (116) follows from Lemmas 25 and 26.

When G is a general positive semidefinite matrix, we choose a new parameter $\eta = G^{1/2}\theta$. Under the new parameter η, the weight matrix is the identity matrix, and the Fisher information matrix is $G^{-1/2}JG^{-1/2}$. Since $(G^{-1/2}JG^{-1/2})^{-1} = G^{1/2}J^{-1}G^{1/2}$, we obtain (116) with a general weight matrix G. $\qquad\square$

F.2 Proof of (117)

Lemma 27. *When G is the identity matrix, a^{**} is a positive semi-definite matrix.*

Proof. Assume that (a, S) satisfies the condition (23). We choose the polar decomposition of a as $a = |a|o$ with an orthogonal matrix o. Then, we have

$$\|y\|^2 \rho + \sum_j (|a|oy)^j D_j - S \geq 0 \tag{214}$$

for $y \in \mathbb{R}^d$. We choose $y' = oy$. Then, we have

$$\|y'\|^2 \rho + \sum_j (|a|y')^j D_j - S \geq 0. \tag{215}$$

Hence, $(|a|, S)$ also satisfies the condition (23) and $\mathrm{Tr}|a| \geq \mathrm{Tr}a$. In fact, the equality of the above inequality holds if and only if a is a positive semi-definite matrix. Therefore, we can conclude that the optimizer a^{**} is a positive semi-definite matrix. □

Proof of (117). Assume that G is the identity matrix. Theorem 17 and Lemma 27 guarantee that $\|a^{**}\| \leq \mathrm{Tr}a^{**} \leq \frac{d}{2\sqrt{\mathrm{Tr}J^{-1}}}$. Hence, we obtain (117).

When G is a general positive-semi definite matrix, we choose a new parameter $\eta = G^{1/2}\theta$. Then, in the same way as (116), we obtain (117). □

References

[1] C. Helstrom, Minimum mean-squared error of estimates in quantum statistics, Physics Letters A **25**, 101 (1967).

[2] C. W. Helstrom, *Quantum detection and estimation theory* (Academic press, 1976).

[3] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory* (Edizioni della Normale, 2011).

[4] H. Nagaoka, A new approach to Cramér-Rao bounds for quantum state estimation, IEICE Tech Report **IT 89-42**, 9 (1989), (Reprinted in [20]).

[5] M. Hayashi and K. Matsumoto, Asymptotic performance of optimal state estimation in qubit system, Journal of Mathematical Physics **49**, 102101 (2008).

[6] R. Demkowicz-Dobrzański, W. Górecki, and M. Guţă, Multi-parameter estimation beyond quantum fisher information, Journal of Physics A: Mathematical and Theoretical **53**, 363001 (2020).

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

[8] F. Albarelli, J. F. Friel, and A. Datta, Evaluating the Holevo Cramér-Rao bound for multiparameter quantum metrology, Phys. Rev. Lett. **123**, 200503 (2019).

[9] J. S. Sidhu, Y. Ouyang, E. T. Campbell, and P. Kok, Tight bounds on the simultaneous estimation of incompatible parameters, Phys. Rev. X **11**, 011028 (2021).

[10] H. Nagaoka, A generalization of the simultaneous diagonalization of hermitian matrices and its relation to quantum estimation theory, in *Asymptotic Theory Of Quantum Statistical Inference: Selected Papers*, edited by M. Hayashi (World Scientific, 2005) pp. 133–149.

Accepted in 〈 〉uantum 2023-08-21, click title to verify. Published under CC-BY 4.0.

50

[11] M. Hayashi, On simultaneous measurement of noncommutative observables. in development of infinite-dimensional non-commutative anaysis, in *Surikaisekikenkyusho (RIMS), Kokyuroku No. 1099, In Japanese* (Kyoto Univ., 1999) pp. 96–188.

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

[13] M. Hayashi, A linear programming approach to attainable cramér-rao type bounds and randomness condition (1997a), arXiv:quant-ph/9704044 [quant-ph] .

[14] L. Gurvits, Classical deterministic complexity of Edmonds' problem and quantum entanglement, STOC '03, 10 (2003).

[15] D. Bruß, Characterizing entanglement, Journal of Mathematical Physics **43**, 4237 (2002).

[16] R. Uola, T. Kraft, J. Shang, X.-D. Yu, and O. Gühne, Quantifying quantum resources with conic programming, Phys. Rev. Lett. **122**, 130404 (2019).

[17] R. Takagi, B. Regula, K. Bu, Z.-W. Liu, and G. Adesso, Operational advantage of quantum resources in subchannel discrimination, Phys. Rev. Lett. **122**, 140402 (2019).

[18] R. Takagi and B. Regula, General resource theories in quantum mechanics and beyond: Operational characterization via discrimination tasks, Phys. Rev. X **9**, 031053 (2019).

[19] S.-I. Amari and H. Nagaoka, *Methods of information geometry* (American Mathematical Soc., 2007).

[20] H. Masahito, ed., *Asymptotic theory of quantum statistical inference: selected papers* (World Scientific, 2005).

[21] D. Petz, *Quantum information theory and quantum statistics* (Springer Science & Business Media, 2007).

[22] J. Suzuki, Y. Yang, and M. Hayashi, Quantum state estimation with nuisance parameters, Journal of Physics A: Mathematical and Theoretical **53**, 453001 (2020).

[23] A. Fujiwara and H. Nagaoka, Quantum fisher metric and estimation for pure state models, Physics Letters A **201**, 119 (1995).

[24] A. Fujiwara and H. Nagaoka, An estimation theoretical characterization of coherent states, Journal of Mathematical Physics **40**, 4227 (1999).

[25] M. Hayashi, *Quantum information theory: Mathematical Foundation* (Springer, 2016) graduate Texts in Physics, First edition was published from Springer in 2006.

[26] M. Hayashi and K. Matsumoto, Statistical model with measurement degree of freedom and quantum physics, in *Asymptotic Theory Of Quantum Statistical Inference: Selected Papers*, edited by M. Hayashi (World Scientific, 2005) pp. 162–169, originally published in Japanese in Surikaiseki Kenkyusho Kokyuroku No. 1055, 1998.

[27] R. D. Gill and S. Massar, State estimation for large ensembles, Phys. Rev. A **61**, 042312 (2000).

[28] M. Hayashi, Comparison between the Cramer-Rao and the mini-max approaches in quantum channel estimation, Commun. Math. Phys. **304**, 689 (2011).

[29] Y. Yang, G. Chiribella, and M. Hayashi, Attaining the ultimate precision limit in quantum state estimation, Communications in Mathematical Physics **36 8**, 223 (2019).

[30] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, Compatibility in multiparameter quantum metrology, Phys. Rev. A **94**, 052108 (2016).

[31] Review of Holevo's research in the 1980s, private paper submitted for review, and to appear on arxiv soon (2023).

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

[33] M. Guţă and J. Kahn, Local asymptotic normality for qubit states, Physical Review A **73**, 052108 (2006).

[34] J. Kahn and M. Guţă, Local asymptotic normality for finite dimensional quantum systems, Communications in Mathematical Physics **289**, 597 (2009).

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

[36] X.-M. Lu and X. Wang, Incorporating heisenberg's uncertainty principle into quantum multiparameter estimation, Phys. Rev. Lett. **126**, 120503 (2021).

[37] M. Hayashi, A linear programming approach to attainable Cramer-Rao type bound, in *Quantum Communication, Computing, and Measurement*, edited by O. Hirota, A. S. Holevo, and C. M. Caves (Plenum, New York, 1997).

[38] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: necessary and sufficient conditions, Physics Letters A **223**, 1 (1996).

[39] B. M. Terhal, A family of indecomposable positive linear maps based on entangled quantum states, Linear Algebra and its Applications **323**, 61 (2001).

[40] M. Lewenstein, B. Kraus, P. Horodecki, and J. I. Cirac, Characterization of separable states and entanglement witnesses, Phys. Rev. A **63**, 044304 (2001).

[41] M. Navascués, Pure state estimation and the characterization of entanglement, Phys. Rev. Lett. **100**, 070503 (2008).

[(2011)] R. Demkowicz-Dobrzański, Optimal phase estimation with arbitrary a priori knowledge, Phys. Rev. A **83**, 061802 (2011).

[42] E. Chitambar, I. George, B. Doolittle, and M. Junge, The communication value of a quantum channel, IEEE Transactions on Information Theory **69**, 1660 (2023).

[43] K. Fujisawa, M. Kojima, and K. Nakata, Exploiting sparsity in primal-dual interior-point methods for semidefinite programming, Mathematical Programming **79**, 235 (1997).

[44] A. Ambainis and J. Emerson, Quantum t-designs: t-wise independence in the quantum world, in *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)* (2007) pp. 129–140, quant-ph/0701126 .

[45] M. Hayashi, *Group Representation for Quantum Theory* (Springer, 2017).

[46] C. An and Y. Xiao, Numerical construction of spherical t-designs by Barzilai-Borwein method, Applied Numerical Mathematics **150**, 295 (2020).

[47] P. Delsarte, J. Goethals, and J. Seidel, Spherical codes and designs, Geom Dedicata **6**, 363 (1977).

[48] M. S. Baladram, On explicit construction of simplex t-designs, Interdisciplinary Information Sciences **24**, 181 (2018).

[49] E. Bannai, E. Bannai, S. Suda, and H. Tanaka, On relative t-designs in polynomial association schemes, Electronic Journal of Combinatorics **22**, 1392 (2015).

[50] E. Bannai and E. Bannai, A survey on spherical designs and algebraic combinatorics on spheres, European Journal of Combinatorics **30**, 1392 (2009).

[51] Y. Ouyang, Computing spectral bounds of the heisenberg ferromagnet from geometric considerations, Journal of Mathematical Physics **60**, 071901 (2019).

[52] G. Lindblad, On the generators of quantum dynamical semigroups, Communications in Mathematical Physics **48**, 119 (1976).

[53] Y. Yang, G. Chiribella, and M. Hayashi, Quantum stopwatch: how to store time in a quantum memory, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **474**, 20170773 (2018).

[54] G. Tóth and I. Apellaniz, Quantum metrology from a quantum information science perspective, Journal of Physics A: Mathematical and Theoretical **47**, 424006 (2014).

[55] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Complete family of separability criteria, Phys. Rev. A **69**, 022308 (2004).

[56] M. D. Vidrighin, G. Donati, M. G. Genoni, X.-M. Jin, W. S. Kolthammer, M. Kim, A. Datta, M. Barbieri, and I. A. Walmsley, Joint estimation of phase and phase diffusion for quantum metrology., Nat Commun **5**, 3532 (2014).

[57] H. Arai and M. Hayashi, Pseudo standard entanglement structure cannot be distinguished from standard entanglement structure, New Journal of Physics **25**, 023009 (2023).

[58] R. M. Van Slyke and R. J.-B. Wets, A duality theory for abstract mathematical programs with applications to optimal control theory, Journal of Mathematical Analysis and Applications **22**, 679 (1968).