



This is a repository copy of *Finding the optimal probe state for multiparameter quantum metrology using conic programming*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/224192/>

Version: Preprint

Preprint:

Hayashi, M. and Ouyang, Y. orcid.org/0000-0003-1115-0074 (Submitted: 2024) Finding the optimal probe state for multiparameter quantum metrology using conic programming. [Preprint - arXiv] (Submitted)

<https://doi.org/10.48550/arXiv.2401.05886>

© The Author(s) 2024. For reuse permissions, please contact the Author(s).

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Finding the optimal probe state for multiparameter quantum metrology using 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*

The aim of the channel estimation is to estimate the parameters encoded in a quantum channel. For this aim, it is allowed to choose the input state as well as the measurement to get the outcome. Various precision bounds are known for the state estimation. For the channel estimation, the respective bounds are determined depending on the choice of the input state. However, determining the optimal input probe state and the corresponding precision bounds in estimation is a non-trivial problem, particularly in the multi-parameter setting, where parameters are often incompatible. In this paper, we present a conic programming framework that allows us to determine the optimal probe state for the corresponding multi-parameter precision bounds. The precision bounds we consider include the Holevo-Nagaoka bound and the tight precision bound that give the optimal performances of correlated and uncorrelated measurement strategies, respectively. Using our conic programming framework, we discuss the optimality of a maximally entangled probe state in various settings. We also apply our theory to analyze the canonical field sensing problem using entangled quantum probe states.

I. INTRODUCTION

Channel estimation [1–4] utilizes quantum resources to allow us to estimate parametrized quantum processes with unprecedented precision. In channel estimation, a quantum channel that embeds physical parameters of interest acts on an initial probe state. For instance, consider estimating the physical parameters associated with a classical field that interacts linearly with an ensemble of qubits. In such a scenario, the quantum channel that describes the dynamics of the quantum system embeds the unknown field parameters. In addition, we have the flexibility to select the initial state, or probe state, of the quantum system. With access to multiple queries of the quantum channel, the objective is to estimate the embedded physical parameters with maximum precision. A fundamental challenge in channel estimation is that of selecting the probe state that gives the optimal precision in estimation.

In quantum state estimation [5–9], we have a family of quantum states that depend continuously on physical parameters. We may interpret the channel estimation problem as a quantum state estimation problem, where the quantum states in the latter problem are the output of a quantum channel that acts on the probe state. In general, precision bounds for the quantum state estimation depend on the type of measurement strategy that we choose. Two such families are correlated and uncorrelated measurement strategies. Correlated measurement strategies operate jointly on infinite copies of quantum systems, whereas uncorrelated measurement strategies act on one copy at a time. In the single parameter setting, the upper bound to

the ultimate precision is the celebrated quantum Cramer Rao bound, which is tight for both correlated and uncorrelated measurement strategies [5–8, 10, 11]. For our channel estimation problem, we set the probe state to be entangled over systems A and C, and the quantum channel maps system A to B for both uncorrelated and correlated measurement strategies (see Figures 1 and 2).

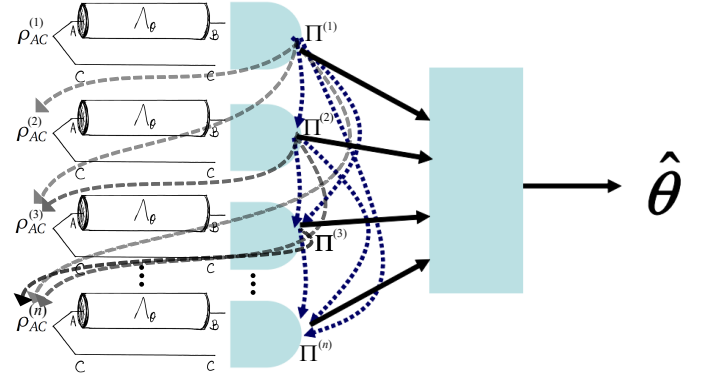


FIG. 1: Channel estimation by uncorrelated measurement and entangled probe states. In the j -th experiment for $j = 1, \dots, n$, we input the probe state $\rho_{AC}^{(j)}$ over systems A and C of equal size, and measure the output entangled state $(\Lambda_\theta \otimes \iota_C)(\rho_{AC}^{(j)})$ via the measurement $\Pi^{(j)}$. Then, we obtain j -th measurement outcome $\omega^{(j)}$. The probe state $\rho_{AC}^{(j)}$ and the measurement $\Pi^{(j)}$ are adaptively chosen by the previous measurement outcomes $\omega^{(1)}, \dots, \omega^{(j-1)}$. Finally, we decide our estimate $\hat{\theta}$ from the measurement outcomes $\omega^{(1)}, \dots, \omega^{(n)}$. The precision bound J_1 is the ultimate precision bound under the strategy of uncorrelated measurements.

* hmasahito@cuhk.edu.cn

† y.ouyang@sheffield.ac.uk

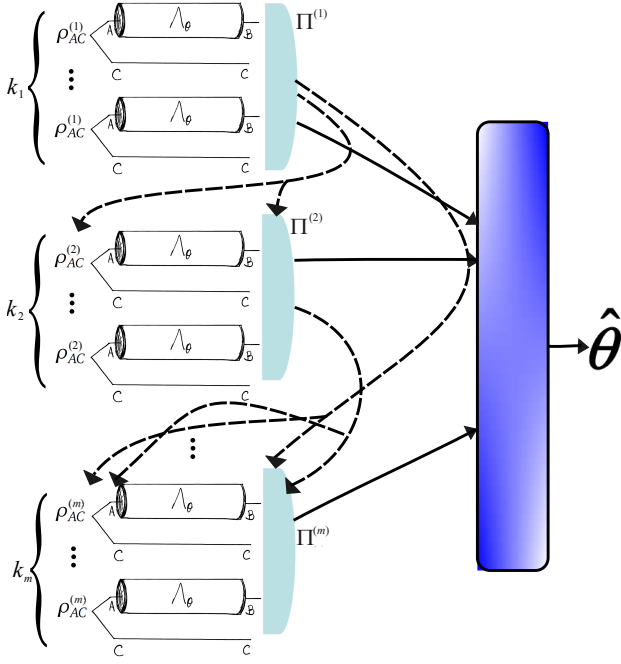


FIG. 2: **Channel estimation by correlated measurement and entangled probe states.** We have m steps, and the j -th step has k_j experiments. In the j -th step, we input the probe state $\rho_{AC}^{(j)}$ over systems A and C for k_j experiments. We measure the output entangled state $(\Lambda_\theta \otimes I_C)(\rho_{AC}^{(j)})^{\otimes k_j}$ via the measurement $\Pi^{(j)}$ and obtain j -th measurement outcome $\omega^{(j)}$. The probe state $\rho_{AC}^{(j)}$ and the measurement $\Pi^{(j)}$ are adaptively chosen by the previous measurement outcomes $\omega^{(1)}, \dots, \omega^{(j-1)}$. Finally, we decide our estimate $\hat{\theta}$ from the measurement outcomes $\omega^{(1)}, \dots, \omega^{(m)}$. In this scenario, we use the unknown channel for a total of $k_1 + \dots + k_m$ times. The precision bound J_5 is the ultimate precision bound under the strategy of correlated measurements.

In the multiparameter setting [12, 13], the theory for quantum state estimation is far richer than the single parameter case; there are several distinct Cramer-Rao type bounds on the optimal precision for both correlated and uncorrelated measurement strategies [14]. These bounds include the Holevo-Nagaoka (HN) bound [7, 8, 11, 15, 16]¹ for correlated measurement strategies and the symmetric logarithmic derivative (SLD) bound². For uncorrelated measurement strategies, the HN and SLD bounds are not tight. The ultimate precision bound, namely the tight bound, for uncorrelated measurement strategies was initially formulated in [17] as an optimization problem with uncountably infinite constraints, and was later reformu-

lated as a conic optimization of an operator with dimension proportional to the size of the probe state and over a certain separable cone in [14]. The Nagaoka-Hayashi (NH) bound [18–20] gives a simpler albeit weaker precision bound for uncorrelated measurement strategies. Recently, it was shown that the HN, SLD, tight, and NH bound can be obtained from the optimal value of conic programs with the same objective function, but optimized over different cones [14]. In Section III, we review multiparameter quantum state estimation, particularly with respect to the conic programming framework of Ref. [14].

From an operational point of view, the tight bound is the most significant for channel estimation, because it describes the ultimate precision bound using uncorrelated measurement strategies. In uncorrelated measurement strategies, after the preparation of each probe state and the action of the quantum channel, we measure the evolved probe state as depicted in Figure 1. We repeat the procedure for multiple copies of probe states, where the explicit form of each probe state and the corresponding measurement strategy on it is informed by all prior measurement outcomes.

We may also be interested in correlated measurement strategies for channel estimation, which are more challenging than uncorrelated measurement strategies to implement in practice. In this case, the relevant precision bound is the HN bound. For the HN bound, we are allowed to perform correlated measurements across multiple copies of probe states (see Figure 2). After each batch of correlated measurements, we can update our choice of probe states and measurement strategies.

Determining the optimal precision for the channel estimation problem entails a two-step optimization process. First, we fix the probe state, and find the precision according to the appropriate Cramer-Rao type bound. Namely, we consider a set of parametrized states when the multiparameter quantum channel acts on the probe state, and evaluate various Cramer-Rao type bounds on the set of parametrized states by solving an appropriate optimization problem. Second, we optimize various Cramer-Rao type bounds by changing the probe state.

The two-step optimization is a non-trivial problem. Even if the first optimization for a fixed probe state is efficient, the optimized precision bound is not necessarily a convex function of the probe state, and hence the second optimization over the probe state need not be so easy to perform. Remarkably, this challenging two-step optimization has been addressed in the single parameter setting [21, 22]. In particular, the references [21, 22] derived a semidefinite programming (SDP) form for the maximum SLD Fisher information in the channel estimation that embeds one parameter. However, the question of how to find both the optimal precision and optimal probe state for multi-parameter quantum channels remains unresolved.

The question of finding optimal probe states for estimating multiple parameters embedded in a quantum channel has recently been considered [23, 24]. For probe states without ancilla assistance, Ref. [24] numerically finds the optimal two-qubit probe state with the HN bound when

¹ In the literature it is often also called the Holevo Cramer-Rao bound.

² This is a direct generalization of the quantum Cramer Rao bound that is based on the SLD.

the channel models 3D-field sensing with independent and identical amplitude (i.i.d.) damping on the two qubits. Ref. [23] considers the HN bound with unitary quantum channels. This leaves open the questions of how to evaluate the optimal probe state and corresponding bounds for the tight, NH and SLD bound for the general problem of quantum channel estimation.

In the channel estimation problem where we estimate the parameters embedded in the quantum channel, the set of parameters is continuous. If we discretize the channel estimation problem, we would obtain the problem of discriminating a discrete set of quantum channels [3, 25–29]. Recently, the channel discrimination problem was formulated as a convex program, and this formulation made it possible to determine the optimal strategy to discriminate a pair of quantum channels [30]. However, it is not so clear how to extend this result to the continuous parameter setting that we require in the channel estimation problem.

In this paper, we give a framework that uses conic programs to find the optimal probe state for channel estimation with respect to multi-parameter bounds. The multi-parameter bounds include the SLD bound, the HN bound, the tight bound, and the NH bound. For the SLD bound, the HN bound, and the NH bound, we show that the conic programs are also in fact semidefinite programs (SDP) and hence can be efficiently solved. For the tight bound, we will require the same technique for optimizing over the separable cone as considered previously [14].

We furthermore apply our framework to unravel situations where the maximally entangled state is optimal, and also study both numerically and theoretically the canonical problem of field sensing in the presence of collective amplitude damping, and in the multiparameter setting.

With our resolution of the non-trivial problem of simultaneously optimizing the multiparameter Cramer-Rao bound and the corresponding probe state, we expect our theoretical contributions to have ramifications beyond the field-sensing application that we have explored. Namely, one will be able to apply our framework to determine optimal precision bounds and optimal probe states for a plethora of quantum sensing problems, where the unknown parameter is embedded in the underlying quantum channel, for instance, in many quantum imaging problems [31].

II. OVERVIEW OF OUR MAIN RESULTS

We establish the following main results.

*** A conic programming framework for channel estimation:-** The conic programming framework that we establish in Section IV proves that the optimal values of conic programs can correspond to Cramer-Rao type bounds for channel estimation. Our framework applies in the multiparameter setting. This main result is given Theorem 1, where we prove that the optimal values of the following are equivalent.

1. Fix the probe state, solve a minimization using the

conic programs in [14] to find the best precision bound. Next solve a maximization for the best probe state to use.

2. Solve a minimization for the conic programs that we introduce.

The optimal values of the above two problems give the precision bounds for channel estimation. Problem 1 is often not a convex program, and hence it is not clear how to solve it numerically. In contrast, in problem 2, the conic programs are convex programs, and we can solve them efficiently. Hence, our framework allows one to solve the channel estimation problem efficiently.

Moreover, the conic programs' optimal solutions also allow us to evaluate the corresponding optimal probe states. Namely, we define a matrix-valued function $\rho_A(Y)$ that maps the optimal solution Y^* of the conic program to a density matrix on system A (see Figures 1 and 2), and the purification of such a density matrix to the joint system AC gives the optimal probe state. Moreover, we discuss how to use the optimal solution of our conic program to calculate the optimal probe state. Hence, it becomes possible to determine the optimal probe state for channel estimation for the HN bound, the SLD bound, the NH bound, and the tight bound. Of operational significance are our results on the channel versions of the HN bound and the tight bound, which give the optimal performances of correlated and uncorrelated measurement strategies, respectively.

*** On the optimality of the maximally entangled probe state:-** There are many situations where the maximally entangled probe state is the optimal probe state to use in channel estimation. In Section V, we explore this possibility, and begin by considering the scenario where the input probe state is a maximally entangled state. We establish Theorem 5 and Theorem 6, which shows that a constraint on a dual matrix-valued variable is equivalent to the optimality of the maximally entangled state for the SLD bound and the HN bound respectively. These theoretical results allow us to find situations when the maximally entangled state is the optimal probe state for both the SLD bound and the HN bound. We consider the following examples.

▷ Noisy qubit channel with one parameter:- We consider the quantum channel as a unitary evolution afflicted with depolarizing noise, where the parameters are embedded into the unitary part of the channel. More precisely, we first consider a quantum channel that describes the mixture of a unitary qubit evolution and replacement by a completely mixed state. The unitary qubit evolution is generated by the product of the single parameter that we want to estimate and a Pauli matrix. We use Theorem 5 to establish Theorem 7, where we prove that the maximally entangled state is the optimal probe state, and furthermore, the tight bound, the NH bound, the SLD bound and the HN bound are all equivalent.

► **Generalized Pauli channels:-** On a qudit system, a generalized Pauli channel applies generalized Paulis randomly on the qudit state according an apriori determined probability distribution. The channel estimation task here is to estimate this probability distribution. We use Theorem 5 to establish Theorem 8, which shows that the optimal probe state for the channel estimation problem for generalized Pauli channels under the SLD bound is the maximally entangled state.

► **Noisy SU(2) channels:-** We consider a quantum channel that is the mixture of unitary evolution according to the spin- j representation of SU(2) unitary evolution according to the spin- j representation of SU(2) and replacement by a completely mixed state. We also give equal weightage to each of the three parameters that we estimate. We use Theorem 5 to establish Theorem 9, where we derive the analytical form for SLD bound. Furthermore, Theorem 9 shows that under purely SU(2) unitary evolution with zero noise, all the precision bounds J_1, \dots, J_5 are equivalent, and the maximally entangled state is the optimal probe state for all the precision bounds.

* **Noisy field sensing:-** In Section VI, we revisit the problem of quantum field sensing [32] in the multiparameter setting. In the noiseless setting, this is equivalent to the channel estimation problem for noisy SU(2) channels that we explore in Section VI. The noise model we consider here is different from that in Section VI, instead of a depolarizing type of channel, we consider noise introduced by collective amplitude damping [33, Eq. (7)]. The channel that we consider differs from [24] in two ways. First, we consider collective amplitude damping while [24] considers i.i.d. amplitude damping. Second, we model the channel using a master equation, considering collective amplitude damping that occurs during the SU(2) evolution, whereas the channel in [24] considers i.i.d. amplitude damping that occur after the unitary evolution. We give corresponding plots of various precision bounds against the noise parameter in Figure 4 and Figure 5.

Using the MatLab computer code, we numerically determine the optimal probe state for the NH bound, the HN bound, and the SLD bound, and numerically evaluate the corresponding precision bounds. We numerically ascertain that in the noiseless setting, the maximally entangled state on the symmetric subspace is the optimal probe state, which agrees with our result in Theorem 9.

When there is non-vanishing noise, we numerically ascertain that the SLD precision bound cannot be optimal. We furthermore prove this fact in Theorem 10, where in the proof, we calculate the expectations of the commutators of the symmetric logarithmic derivatives of different angular momentum generators.

* **Semi-definite programming formulations:-** To maximize the accessibility of our work, in Section D, we give formulate the mathematical optimization problems that correspond to J_2, J_3, J_4, J_5 as semidefinite programs to be used

with the CVX package and provide the corresponding MatLab code.

III. QUANTUM STATE ESTIMATION

In the quantum parameter estimation problem, given a set of parameters $\Theta \subseteq \mathbb{R}^d$, we consider d -parameter state family $\{\rho_\theta : \theta = (\theta^1, \dots, \theta^d) \in \Theta\}$ on the Hilbert space \mathcal{H} . Denoting the set of density matrices on \mathcal{H} by $\mathcal{S}(\mathcal{H})$, we note that the quantum states ρ_θ are elements of $\mathcal{S}(\mathcal{H})$, which means that they are positive semidefinite operators with unit trace. We define the true parameter to be θ_0 , and our goal is to construct an estimator $\hat{\theta}$ that will estimate θ_0 . We furthermore consider Θ as a continuous set, and where the quantum states ρ_θ are differentiable with respect to θ , so that we can define the partial derivatives ρ_θ . In particular, in the neighborhood of the true parameter θ_0 , we define $D_j := \frac{\partial}{\partial \theta^j} \rho_\theta|_{\theta_0}$, and $\rho := \rho_{\theta_0}$.

We can describe a measurement using a set of positive operators $\Pi = \{\Pi_x : x \in \mathcal{X}\}$ where the completeness condition $\sum_{x \in \mathcal{X}} \Pi_x = I$ holds, where \mathcal{X} is a set of classical labels. By the Born rule, when we perform a measurement Π on a quantum state ρ_θ , we will obtain the classical label x and the state $\Pi_x \rho_\theta / \text{Tr}(\Pi_x \rho_\theta)$ with probability $p_\theta(x) = \text{Tr}(\Pi_x \rho_\theta)$. Upon measurement, the classical label x is a random variable, and we denote $\mathbb{E}[f(x)|\Pi]$ as the expectation of $f(x)$, the function f of the classical label x , with probability distribution obtained according to the Born rule. Now given a measurement Π and an estimator $\hat{\theta}$, where the estimator $\hat{\theta}$ is to be a function of the classical label x , we denote $\hat{\Pi} = (\Pi, \hat{\theta})$ as an *estimator*. When the true parameter θ_0 is equal to θ , we define the mean-square error (MSE) matrix for the estimator $\hat{\Pi}$ as

$$\begin{aligned} V_\theta[\hat{\Pi}] &= \sum_{i,j=1}^d |i\rangle\langle j| \sum_{x \in \mathcal{X}} \text{Tr} \rho_\theta \Pi_x (\hat{\theta}^i(x) - \theta^i)(\hat{\theta}^j(x) - \theta^j) \\ &= \sum_{i,j=1}^d |i\rangle\langle j| \mathbb{E}_\theta [(\hat{\theta}^i(x) - \theta^i)(\hat{\theta}^j(x) - \theta^j) | \Pi]. \end{aligned}$$

In multiparameter quantum metrology, the objective is to find an optimal estimator $\hat{\Pi} = (\Pi, \hat{\theta})$ that in some sense minimizes the MSE matrix. More precisely, the quantity to minimize is $\text{Tr} G V_\theta[\hat{\Pi}]$, where a weight matrix G , a size d positive semidefinite matrix, quantifies the relative importance of the different parameters in our parameter estimation problem. If we give all parameters an equal importance, we may choose G to be the identity matrix.

When for all $i = 1, \dots, d$, the expectation of our estimator is equal to the true value of the parameter θ_0 when $\theta_0 = \theta$, we have the condition

$$\mathbb{E}_\theta[\hat{\theta}^i(x)|\Pi] = \sum_{x \in \mathcal{X}} \hat{\theta}^i(x) \text{Tr}[\rho_\theta \Pi_x] = \theta^i, \quad (1)$$

which means that our estimator $\hat{\Pi} = (\Pi, \hat{\theta})$ is unbiased at $\theta_0 = \theta$. When (1) holds for all $\theta \in \Theta$, we say that our

estimator is globally unbiased. Unfortunately, globally unbiased estimators need not exist, and hence we consider estimators that are unbiased in the neighborhood of the true parameter θ_0 . This motivates us to take partial derivatives on both sides of (1), and consider

$$\frac{\partial}{\partial \theta_j} \mathbb{E}_\theta [\hat{\theta}^i(x)|\Pi] = \sum_{x \in \mathcal{X}} \hat{\theta}^i(x) \text{Tr} D_j \Pi_x = \delta_{ij}^j, \quad (2)$$

where $D_j = \frac{\partial}{\partial \theta_j} \rho_\theta$. When (1) holds for all $i = 1, \dots, d$ for a fixed θ where $\theta_0 = \theta$, and when (2) holds for all $i, j = 1, \dots, d$, then we say that the estimator $\hat{\Pi} = (\Pi, \hat{\theta})$ satisfies the locally unbiased condition.

For any weight matrix $G = \sum_{i,j=1}^d g_{i,j} |i\rangle\langle j|$, the fundamental precision limit [14] is given by

$$C_\theta[G] := \min_{\hat{\Pi}: \text{l.u. at } \theta} \text{Tr}[GV_\theta[\hat{\Pi}]],$$

where the minimization is carried out for all possible estimators under the locally unbiasedness condition, which is indicated by l.u. at θ . Note that when we impose only the condition (2), the above minimum is attained by $\hat{\Pi}$ satisfying (1). Therefore, $C_\theta[G]$ can be considered as the minimum only with the condition (2). In the following, we consider the above minimization. Hence, it is sufficient to focus on the operators ρ_θ and $(D_j)_j$. Thus, the pair $(\rho_\theta, (D_j)_j)$ is called a model in the following.

As discussed in Ref. [14], any lower bound to weighted trace of the MSE matrix is a Cramer-Rao (CR) type bound. The fundamental precision limit $C_\theta[G]$ is one such lower bound which is tight, and hence referred to as the *tight CR bound* [14]. Operationally, we may attain the tight CR bound using an uncorrelated measurement strategy in the asymptotic setting. This means that we would attain this tight CR bound by performing optimal measurements independently on asymptotically many individual copies of the probe states. In quantum parameter estimation, one may also consider correlated measurement strategies, where we would allow joint measurements over multiple, and potentially infinite number of copies of probe states. The Holevo-Nagaoka (HN) bound [7, 8, 11, 15, 16] is a CR bound that describes the ultimate precision in this correlated measurement strategy scenario, and the HN bound can be strictly smaller than the tight CR bound [14].

While it is difficult to evaluate the tight CR bound exactly, one can nonetheless use a semidefinite program (SDP) to approximate it [14]. As the precision of the approximation increases, the complexity of the SDP also increases. Other CR-type bounds are more efficient to evaluate, such as the Nagaoka-Hayashi (NH) bound, the HN bound, and the SLD bound. Recently, Ref. [14] used the language of conic programming to clarify the relationship between these CR-type bounds. Namely, let us consider an operator X that we construct from any estimator $\hat{\Pi} = (\Pi, \hat{\theta})$, where

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

This operator $X(\Pi, \hat{\theta})$ acts on the vectors in $\mathcal{R} \otimes \mathcal{H}$, where $\mathcal{R} = \mathbb{R}^{d+1}$ is spanned by $d+1$ basis vectors $|0\rangle, |1\rangle, \dots, |d\rangle$.

We may write the trace of the weighted MSE matrix $\text{Tr} GV_\theta[\hat{\Pi}]$ using $X(\Pi, \hat{\theta})$. Namely,

$$\text{Tr} GV_\theta[\hat{\Pi}] = \text{Tr}(G \otimes \rho) X(\Pi, \hat{\theta}). \quad (3)$$

Next, note that the completeness condition $\sum_{x \in \mathcal{X}} \Pi_x = I_{\mathcal{H}}$ using $X(\Pi, \hat{\theta})$ implies that

$$\text{Tr}_{\mathcal{R}}(|0\rangle\langle 0| \otimes I_{\mathcal{H}}) X(\Pi, \hat{\theta}) = I_{\mathcal{H}}. \quad (4)$$

Hence, we may interpret (4) as a rewriting of the completeness condition $\sum_{x \in \mathcal{X}} \Pi_x = I_{\mathcal{H}}$. Next, we note that the condition (2) for a locally unbiased estimator guarantees

$$\text{Tr}\left(\frac{1}{2}(|0\rangle\langle i| + |i\rangle\langle 0|) \otimes D_j\right) X(\Pi, \hat{\theta}) = \delta_{i,j}. \quad (5)$$

Hence, we may interpret (5) as a rewriting of the locally unbiased condition.

Note that the operator $X(\Pi, \hat{\theta})$ has a tensor product structure. Namely, we may consider $X(\Pi, \hat{\theta})$ as an element of a cone generated by separable states on $\mathcal{R} \otimes \mathcal{H}$. More precisely, this separable cone \mathcal{S}^1 is the convex hull of the set of operators that are a tensor product of a real positive semidefinite matrix on \mathcal{R} and a complex positive semidefinite operators on \mathcal{H} with bounded norm. Hence, it is natural to consider the minimization:

$$S_1 := \min_{X \in \mathcal{S}^1} \{\text{Tr}(G \otimes \rho) X | (4), (5) \text{ hold.}\}$$

The reference [14] showed that S_1 is in fact equal to the tight CR type bound. The reference [14] also showed that if we consider the minimization of $\text{Tr}(G \otimes \rho) X$ subject to the conditions (4), (5), but over suitable cones that contain \mathcal{S}^1 , the optimal value can be made to be equal to the NH bound, the HN bound and the SLD bound. Given this phenomenon, it is instructive to revisit the cones over which we can optimize X .

We now proceed to define several cones over which we like to optimize X . Now let us define \mathcal{B} as the vector space spanned by the tensor product of real symmetric matrices on \mathcal{R} and bounded complex Hermitian matrices on \mathcal{H} . This means that we can write

$$\mathcal{B}(\mathcal{R}, \mathcal{H}) := \left\{ \sum_{j=0}^d \sum_{k=0}^d |k\rangle\langle j| \otimes X_{k,j} \mid X_{k,j} \in \mathcal{B}_{\text{sa}}(\mathcal{H}), X_{k,j} = X_{j,k} \right\},$$

where $\mathcal{B}_{\text{sa}}(\mathcal{H})$ denotes the set of self-adjoint (Hermitian) matrices on \mathcal{H} with bounded norm. When there is no confusion, $\mathcal{B}(\mathbb{R}^{d+1}, \mathcal{H})$ is simplified to \mathcal{B} . Let us consider the cone \mathcal{S}^2 as

$$\mathcal{S}^2 := \{X \in \mathcal{B} \mid \langle v|X|v \rangle \geq 0 \text{ for all } |v\rangle \in \mathbb{C}^{d+1} \otimes \mathcal{H}\}.$$

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

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

Symbol	Meaning
d	the number of parameters to be estimated
G	weight matrix, real positive semidefinite $d \times d$ matrix
θ_0	the parameters' true value, a vector in \mathbb{R}^d
$\hat{\theta}$	estimator of the d parameters, a vector in \mathbb{R}^d
$\Theta \subseteq \mathbb{R}^d$	set of all possible parameter vectors
Λ_θ	quantum channel (mapping system A to B) that embeds θ
T_θ	Choi matrix Λ_θ
T_{θ_0}	Choi matrix Λ_{θ_0} at the parameter's true value θ_0
ρ_{AC}	a probe state parametrized by θ
F_j	Choi matrix's partial derivative $\frac{\partial}{\partial \theta_j} T_\theta _{\theta=\theta_0}$.
S_1	CR-type bound: tight bound, optimized over cone $\mathcal{S}^1 = \mathcal{S}_{BC}^1$
S_2	CR-type bound: Nagaoka-Hayashi bound, optimized over cone $\mathcal{S}^2 = \mathcal{S}_{BC}^2$
S_3	CR-type bound: bound from optimization over PPT cone, optimized over cone $\mathcal{S}^3 = \mathcal{S}_{BC}^3$
S_4	CR-type bound: SLD bound, optimized over cone $\mathcal{S}^4 = \mathcal{S}_{BC}^4$
S_5	CR-type bound: Holevo-Nagaoka bound, optimized over cone $\mathcal{S}^5((\Lambda_\theta \otimes \iota_C)(\rho_{AC})) = \mathcal{S}_{BC}^5((\Lambda_\theta \otimes \iota_C)(\rho_{AC}))$
$\mathcal{S}^1, \mathcal{S}^2, \mathcal{S}^3, \mathcal{S}^4, \mathcal{S}^5((\Lambda_\theta \otimes \iota_C)(\rho_{AC}))$	cones in $\mathcal{R}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_A$ that correspond to the CR-type bounds S_1, S_2, S_3, S_4, S_5
$\tilde{S}_k := \min_{\rho_{AC}} \tilde{S}_k[\rho_{AC}]$	CR-type precision bound for the channel estimation problem using the optimal probe state ρ_{AC}
$\mathcal{S}_{BA}^1, \mathcal{S}_{BA}^2, \mathcal{S}_{BA}^3, \mathcal{S}_{BA}^4, \mathcal{S}_{BA}^5(T)$	cones in $\mathcal{R}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_A$ that are analogous to the cones in the conic optimizations of S_1, S_2, S_3, S_4, S_5
J_1, J_2, J_3, J_4, J_5	conic optimizations on cones $\mathcal{S}_{BA}^1, \mathcal{S}_{BA}^2, \mathcal{S}_{BA}^3, \mathcal{S}_{BA}^4, \mathcal{S}_{BA}^5(T)$
Y	optimization variable for the cones $\mathcal{S}_{BA}^1, \mathcal{S}_{BA}^2, \mathcal{S}_{BA}^3, \mathcal{S}_{BA}^4, \mathcal{S}_{BA}^5(T)$
$\rho_A(Y)$	When Y is the optimal solution of CR-type bound, the purification of $\rho_A(Y)$ is the optimal entangled state
$ \Phi\rangle\langle\Phi $	maximally entangled state, an example of a ρ_{AC}
\mathcal{K}	set of matrices on system A and B for which the partial trace on B is proportional to the identity matrix

TABLE I: Notations for the channel estimation problem.

Let us define $\mathcal{S}(\mathbb{C}^{d+1} \otimes \mathcal{H})_{\text{PPT}}$ as the set of self-adjoint operators on $\mathbb{C}^{d+1} \otimes \mathcal{H}$ with positive partial transpose, and define \mathcal{S}^3 as $\mathcal{S}(\mathbb{C}^{d+1} \otimes \mathcal{H})_{\text{PPT}} \cap \mathcal{B}''$. Likewise, we define the set $\mathcal{S}(\mathbb{C}^{d+1} \otimes \mathcal{H})_{\text{P}}$ as the set of positive semi-definite self-adjoint operators on $\mathbb{C}^{d+1} \otimes \mathcal{H}$, and define \mathcal{S}^4 as $\mathcal{S}(\mathbb{C}^{d+1} \otimes \mathcal{H})_{\text{P}} \cap \mathcal{B}''$. Then, for $k = 1, 2, 3, 4$, we define

$$S_k := \min_{X \in \mathcal{S}^k} \{\text{Tr}(G \otimes \rho)X | (4), (5) \text{ hold.}\}. \quad (6)$$

The relation

$$\mathcal{S}^1 \subset \mathcal{S}^2 \subset \mathcal{S}^3 \subset \mathcal{S}^4$$

implies

$$S_1 \geq S_2 \geq S_3 \geq S_4. \quad (7)$$

In addition, we introduce a linear constraint to the operator $X \in \mathcal{B}''$ as

$$\text{Tr}X(|j\rangle\langle i| - |i\rangle\langle j|) \otimes T = 0 \quad (8)$$

for $i, j = 1, 2, \dots, d$ and a trace-class self-adjoint operator T . We may simplify (8) to

$$\text{Im Tr}X(|j\rangle\langle i| \otimes T) = 0.$$

by noting that $\text{Tr}X(|i\rangle\langle j| \otimes T)$ is the complex conjugate of $\text{Tr}X(|j\rangle\langle i| \otimes T)$ because $\text{Tr}_R X(|i\rangle\langle j| \otimes I) = (\text{Tr}_R X(|j\rangle\langle i| \otimes I))^\dagger$. Using this linear constraint, we define the subspace \mathcal{B}_T'' of \mathcal{B}'' as

$$\mathcal{B}_T'' := \{X \in \mathcal{B}'' | (8) \text{ holds.}\}.$$

Next, given a density matrix ρ on \mathcal{H} , let us define $\mathcal{S}^5(\rho)$ as $\mathcal{S}(\mathbb{C}^{d+1} \otimes \mathcal{H})_{\text{P}} \cap \mathcal{B}_\rho''$. We consider the minimization:

$$S_5 := \min_{X \in \mathcal{S}^5(\rho)} \{\text{Tr}(G \otimes \rho)X | (4), (5) \text{ hold.}\}.$$

Note that only the cone $\mathcal{S}^5(\rho)$ depends on ρ , for the cones $\mathcal{S}^1, \mathcal{S}^2, \mathcal{S}^3, \mathcal{S}^4$ are independent of ρ . Since we have the relation

$$\mathcal{S}^2 \subset \mathcal{S}^5(\rho) \subset \mathcal{S}^4,$$

we have the following relations

$$S_4 \leq S_5 \leq S_2. \quad (9)$$

Since S_k depends on the model $(\rho, (D_j)_j)$, i.e., the probe state ρ and the partial derivatives of the probe state D_j for $k = 1, 2, 3, 4, 5$, we can also write S_k as

$$S_k[\rho, (D_j)_j]$$

to emphasize the CR-type bounds' dependence on ρ and D_j .

Ref. [14] showed that S_2 equals the Nagaoka-Hayashi bound (NH bound) studied in Ref. [18–20]. Also, Ref. [14] showed that S_4 equals the SLD bound, and S_5 equals the HN bound. In the single parameter case, i.e., $d = 1$, the SLD bound is attainable. Hence, we have the equality in (7) and (9), i.e., the equation:

$$S_1 = S_2 = S_3 = S_5 = S_4. \quad (10)$$

For further discussion, we prepare several notations. We use the notation $X \circ Y := \frac{1}{2}(XY + YX)$. We define the SLD

operator L_j on \mathcal{H} as

$$\rho \circ L_j = D_j.$$

We denote the SLD fisher information matrix by J_{SLD} . We define $L^i := \sum_{j=1}^d (J_{\text{SLD}}^{-1})^{i,j} L_j$ as a linear combination of SLD operators that depend on the i th row of the inverse SLD Fisher information matrix. These operators L^i satisfy the constraint $\text{Tr} D_j L^i = \text{Tr} D_j \sum_{j'=1}^d (J_{\text{SLD}}^{-1})^{j,j'} L_{j'} = \sum_{j'=1}^d (J_{\text{SLD}}^{-1})^{i,j'} \text{Tr} D_j L_{j'} = \sum_{j'=1}^d (J_{\text{SLD}}^{-1})^{i,j'} J_{\text{SLD},j,j'} = \delta_j^i$. When we need to clarify the dependence of the model $(\rho, (D_j)_j)$, L^i , L_i , and J_{SLD} are denoted by $L^i[\rho, (D_j)_j]$, $L_i[\rho, (D_j)_j]$, and $J_{\text{SLD}}[\rho, (D_j)_j]$, respectively.

Now, we consider the n -fold tensor product system $\mathcal{H}^{\otimes n}$. Given an operator X on \mathcal{H} , we define the operator $X^{(n)}$ on $\mathcal{H}^{\otimes n}$ as $X^{(n)} := \sum_{i=1}^n X_i^{(n)}$, where $X_i^{(n)} := I^{\otimes i-1} \otimes X \otimes I^{n-i}$. Then, we call the model $(\rho^{\otimes n}, (D_j^{(n)})_j)$ as the n -fold extension of the state model $(\rho, (D_j)_j)$. Then, we have

$$\begin{aligned} L_i[\rho^{\otimes n}, (D_j^{(n)})_j] &= (L_i[\rho, (D_j)_j])^{(n)} \\ J_{\text{SLD}}[\rho^{\otimes n}, (D_j^{(n)})_j] &= n J_{\text{SLD}}[\rho, (D_j)_j] \\ L^i[\rho^{\otimes n}, (D_j^{(n)})_j] &= \frac{1}{n} (L^i[\rho, (D_j)_j])^{(n)}. \end{aligned}$$

Also, the relation

$$S_k[\rho, (D_j)_j] = n S_k[\rho^{\otimes n}, (D_j^{(n)})_j]$$

holds for $k = 4$ [34] and for $k = 5$ [11, Lemma 4].

IV. CHANNEL ESTIMATION

In the previous section, the optimizations S_k depend on the quantum model comprising of density operators ρ and their partial derivatives with respect to the parameters to be estimated. Here, we consider the channel estimation problem, where we have a d -parameter channel family $\{\Lambda_\theta\}$, where the input system is \mathcal{H}_A and the output system is \mathcal{H}_B . As we can see, the quantum channel Λ_θ embeds the d parameters θ to be estimated. After the quantum channel Λ_θ maps a probe state on \mathcal{H}_A to an output state $\Lambda_\theta(\rho)$ on \mathcal{H}_B . With asymptotically many copies of quantum states $\Lambda_\theta(\rho)$, we can perform appropriate measurements, either in a correlated or uncorrelated setting, to obtain a probability distribution that depends on the parameters, from which we may construct the best informed estimator for the parameters.

In the channel estimation problem, it is typical to consider the purification of the probe state ρ to a pure state on $\mathcal{H}_A \otimes \mathcal{H}_C$. Here, the Hilbert space \mathcal{H}_C , isomorphic to \mathcal{H}_A , is an ancillary system that the quantum channel has not access to. Our measurement strategies however do have access the ancillary system. This setting follows the preceding paper [21] which studies the optimization of the one-parameter case under the assumption that the ancilla system is available. We remark that considering measurement strategies

that do not have access to this ancilla system is highly non-trivial, and beyond the scope of our current study. In what follows, we assume that we have full access to the ancilla system, which in turn means that the quantities of interest are

$$\bar{S}_k[\rho_{AC}] := S_k[(\Lambda_\theta \otimes \iota_C)(\rho_{AC}), (\frac{\partial}{\partial \theta_j} (\Lambda_\theta \otimes \iota_C)(\rho_{AC}))_j], \quad (11)$$

where ι_C denotes an identity channel on system \mathcal{H}_C . With complete access to the ancilla system, the corresponding CR-type precision bounds are then

$$\bar{S}_k := \min_{\rho_{AC}} \bar{S}_k[\rho_{AC}] \quad (12)$$

where the minimization is over all pure density operators on $\mathcal{H}_A \otimes \mathcal{H}_C$, and the accessible measurements for our estimation are applied to the tensor product system $\mathcal{H}_B \otimes \mathcal{H}_C$. Since the whole ancilla system is accessible, when ρ_{AC} is a mixed state, we are allowed to retake the system \mathcal{H}_C to contain the reference of the purification of ρ_{AC} . This explains why the range of the above minimization is limited to all pure density operators on $\mathcal{H}_A \otimes \mathcal{H}_C$.

It is not immediately obvious how one would solve the optimization in (12). Even if one solves the inner optimization for S_k , it is unclear if the subsequent optimization in ρ_{AC} is a tractable optimization problem, such as a convex problem. We overcome these difficulties. Namely, we construct conic programs with optimal values that are precisely equal to (12), which allows us to find (12) using only a single optimization program.

The first tool that we use is the Choi matrix of a quantum channel [35]. We denote the Choi matrix of Λ_θ by T_θ . Given any orthonormal basis $\{|e_j\rangle\}$ of \mathcal{H}_A , we can define the unnormalized maximally entangled state $|I\rangle := \sum_j |e_j\rangle |e_j\rangle$ on $\mathcal{H}_{A'} \otimes \mathcal{H}_A$, where $\mathcal{H}_{A'}$ is isomorphic to \mathcal{H}_A . Then, the Choi matrix T_θ of Λ_θ is an operator on $\mathcal{H}_B \otimes \mathcal{H}_A$ and is given by

$$T_\theta := (\Lambda_\theta \otimes \iota)(|I\rangle \langle I|).$$

Since $\rho = \text{Tr}_{A'}[|I\rangle \langle I| (I_{A'} \otimes \rho)]$ for any input state ρ on \mathcal{H}_A , where $I_{A'} = \sum_j |e_j\rangle \langle e_j|$ we have

$$\begin{aligned} \Lambda_\theta(\rho) &= \Lambda_\theta(\text{Tr}_{A'}[|I\rangle \langle I| (I_{A'} \otimes \rho)]) \\ &= \text{Tr}_{A'}[(\Lambda_\theta \otimes \iota)(|I\rangle \langle I| (I_B \otimes \rho))] = \text{Tr}_{A'}[T_\theta (I_B \otimes \rho)], \end{aligned} \quad (13)$$

where I_B denotes the identity operator on \mathcal{H}_B . Here, (13) allows us to rewrite $\Lambda_\theta(\rho)$ in terms of the Choi matrix T_θ and the input probe state ρ .

Next, we can consider the partial derivatives of $\Lambda_\theta(\rho)$ in terms of the Choi matrix derivatives. Namely,

$$\frac{\partial}{\partial \theta_j} \Lambda_\theta(\rho) = \text{Tr}_{A'}[\frac{\partial}{\partial \theta_j} T_\theta (I_B \otimes \rho)].$$

When the parameter θ is in the neighborhood of the true parameter θ_0 , we denote the corresponding Choi matrix T_{θ_0}

by T and its derivatives as $F_j := \frac{\partial}{\partial \theta_j} T_\theta|_{\theta=\theta_0}$. In this case, we can write

$$\frac{\partial}{\partial \theta_j} \Lambda_\theta(\rho)|_{\theta=\theta_0} = \text{Tr}_A[F_j(I_B \otimes \rho)]. \quad (14)$$

Since we allow access to an ancilla system \mathcal{H}_C for both preparation and measurement of the probe state, instead of ρ as the input state, we consider ρ_{AC} on $\mathcal{H}_A \otimes \mathcal{H}_C$ as the input state. When the quantum channel still maps \mathcal{H}_A to \mathcal{H}_B , then the output state that corresponds to our input state ρ_{AC} at the true parameter value θ_0 is

$$(\Lambda_\theta \otimes \iota_C)(\rho_{AC})|_{\theta=\theta_0} = \text{Tr}_A[(T \otimes I_C)(I_B \otimes \rho_{AC})], \quad (15)$$

and its derivatives are

$$\frac{\partial}{\partial \theta_j} (\Lambda_\theta \otimes \iota_C)(\rho_{AC})|_{\theta=\theta_0} = \text{Tr}_A[(F_j \otimes I_C)(I_B \otimes \rho_{AC})]. \quad (16)$$

In this setting, the joint system $\mathcal{H}_B \otimes \mathcal{H}_C$ is accessible for our measurement for our estimation. Then, the CR-type bound $\bar{S}_k[\rho_{AC}]$ of channel estimation problem as given in (11) can be written as

$$\begin{aligned} & \bar{S}_k[\rho_{AC}] \\ &= S_k[\text{Tr}_A[(T \otimes I_C)(I_B \otimes \rho_{AC})], (\text{Tr}_A[(F_j \otimes I_C)(I_B \otimes \rho_{AC})])_j] \end{aligned} \quad (17)$$

with $k = 1, 2, 3, 4, 5$. In the following, the pair $(T, (F_j)_j)$ is called the channel model. In particular, when we need to clarify the dependence of the channel model, $\bar{S}_k[\rho_{AC}]$ is denoted by $\bar{S}_k[T, (F_j)_j, \rho_{AC}]$.

Note that the operators in the left sides of (15) and (16) act on the space $\mathcal{H}_B \otimes \mathcal{H}_C$. In the conic programming formulation of (11), we require the introduction of the space $\mathcal{R}_C = \mathbb{C}^{d+1}$, and consider conic programming over the tripartite system $\mathcal{R}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. For $k = 1, 2, 3, 4$, we define the cones we consider are operators on $\mathcal{R}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ which are equal to S^k as explained in Section III under the choice $\mathcal{H} = \mathcal{H}_B \otimes \mathcal{H}_C$ as S_{BC}^k . Following the notation of Section III, we use X on $\mathcal{R}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ to optimize within the appropriate cones.

After solving (11), we need to perform a subsequent minimization over all states ρ_{AC} on the space $\mathcal{H}_A \otimes \mathcal{H}_C$, which is a challenging task. To overcome this challenge, we formulate new conic programs with optimal values equal to those given by (12).

Our formulation of new conic programs draws upon the insight that the space \mathcal{H}_C is in fact isomorphic to the space \mathcal{H}_A , and that we may consider optimizing over cones on $\mathcal{R}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_A$ instead of $\mathcal{R}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. This means that instead of considering the cones S_{BC}^k , we like to consider the cones S_{BA}^k , where we obtain S_{BA}^k by replacing \mathcal{H} with $\mathcal{H}_B \otimes \mathcal{H}_A$ in the definition of S^k .

One challenge in this idea is to be able to construct ρ_{AC} from the optimal solution of the cone. Denoting Y , an operator on $\mathcal{R}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_A$, as the optimization variable in this case, this challenge can be reduced to that of finding an appropriate density operator $\rho_A(Y)$ on \mathcal{H}_A that depends on Y ,

and subsequently purifying $\rho_A(Y)$ to ρ_{AC} . To formulate the objective function of the conic programs on Y , we revisit the optimization of (17). Note that, the objective function of (17) is

$$\begin{aligned} & \text{Tr}[G \otimes \text{Tr}_A[(T \otimes I_C)(I_B \otimes \rho_{AC})]X] \\ &= \text{Tr}[G \otimes ((T \otimes I_C)(I_B \otimes \rho_{AC}))(I_A \otimes X)] \\ &= \text{Tr}[(G \otimes T)\text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)]]. \end{aligned}$$

Identifying Y as $\text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)]$, the objective function of the our conic programs becomes

$$\text{Tr}[Y(G \otimes T)].$$

We establish an appropriate relation between the variable Y and some ρ_{AC} through the conditions:

(i): Given fixed Y on $\mathcal{R}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_A$, there exists a state ρ_A on \mathcal{H}_A such that

$$\text{Tr}_R[Y(|0\rangle\langle 0| \otimes I_{AB})] = I_B \otimes \rho_A. \quad (18)$$

(ii):

$$\frac{1}{2} \text{Tr}[Y((|0\rangle\langle j'| + |j'\rangle\langle 0|) \otimes F_j)] = \delta_{j,j'}$$

for $j, j' = 1, \dots, d$.

While the condition (ii) is a linear constraint, it is not immediately apparent how condition (i) can be written as a linear constraint. However, we point out that the condition (i) is in fact equivalent to the following linear constraint.

(i'): Let $\{|b\rangle\}$ be any orthonormal basis of \mathcal{H}_B . For $b \neq b'$, we have

$$\text{Tr}_{RB}[Y(|0\rangle\langle 0| \otimes I_A \otimes |b\rangle\langle b'|)] = 0, \quad (19)$$

$$\begin{aligned} & \text{Tr}_{RB}[Y(|0\rangle\langle 0| \otimes I_A \otimes |b\rangle\langle b|)] \\ &= \text{Tr}_{RB}[Y(|0\rangle\langle 0| \otimes I_A \otimes |b'\rangle\langle b'|)] \end{aligned} \quad (20)$$

as operators on \mathcal{H}_A . Also,

$$\text{Tr}[Y(|0\rangle\langle 0| \otimes I_A \otimes |b\rangle\langle b|)] = 1. \quad (21)$$

The equivalence between (i) and (i') is shown as follows. Since (i) \Rightarrow (i') is trivial, we show (i') \Rightarrow (i). We assume that (i'). We denote $\text{Tr}_B[\text{Tr}_R[Y(|0\rangle\langle 0| \otimes I_{AB})](I_A \otimes |b\rangle\langle b|)] = \text{Tr}_{RB}[Y(|0\rangle\langle 0| \otimes I_A \otimes |b\rangle\langle b|)]$ by $\rho_A(Y)$. Then, (21) guarantees that $\rho_A(Y)$ is a state. Also, (19) and (20) imply (18). We obtain the condition (i). Therefore, the condition (i) is replaced by the linear constraint (i'). Furthermore, when we impose the condition (i), the operator

$$\rho_A(Y) := \frac{1}{d_B} \text{Tr}_{RB}[Y(|0\rangle\langle 0| \otimes I_{AB})] \quad (22)$$

is a density matrix, where d_B is the dimension of \mathcal{H}_B . Given conditions (i') and (ii) along with the objective function

$\text{Tr}[Y(G \otimes T)]$, we can define our new conic programs. Namely for $k = 1, 2, 3, 4$, we define

$$J_k := \min_{Y \in \mathcal{S}_{BA}^k} \{\text{Tr}[Y(G \otimes T)] | Y \text{ satisfies (i'), (ii).}\}.$$

For $k = 5$, we define $\mathcal{S}_{BA}^5(T) := \mathcal{S}(\mathcal{R}_C \otimes \mathcal{H}_{AB})_P \cap \mathcal{B}_T''$, and define the conic linear programming.

$$J_5 := \min_{Y \in \mathcal{S}_{BA}^5(T)} \{\text{Tr}[Y(G \otimes T)] | Y \text{ satisfies (i'), (ii).}\}.$$

When we need to clarify the dependence of the channel model $(T, (F_j)_j)$, J_k is denoted by $J_k[T, (F_j)_j]$.

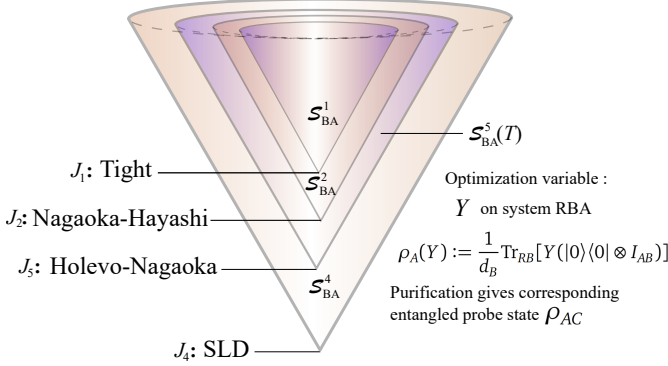


FIG. 3: . We depict the relationship between J_1, J_2, J_5, J_4 and their associated cones $\mathcal{S}_{BA}^1, \mathcal{S}_{BA}^2, \mathcal{S}_{BA}^5(T), \mathcal{S}_{BA}^4$. The optimal values of these optimizations, which are minimizations, are the tight bound, the Nagaoka-Hayashi bound, the Holevo-Nagaoka bound, and the SLD bound respectively. We pictorially illustrate that $J_1 \geq J_2 \geq J_5 \geq J_4$.

We denote the optimization variable for all of these minimizations as Y , which is a matrix supported on $\mathcal{R}_C \otimes \mathcal{H}_B \otimes \mathcal{H}_A$. From the optimal solution Y^* of any of these conic programmings, we can derive a corresponding $\rho_A(Y^*)$, whose purification to system AC yields the corresponding optimal probe state.

The main result of our paper is the following theorem.

Theorem 1. For $k = 1, 2, 3, 4, 5$, we have

$$J_k = \bar{S}_k. \quad (23)$$

This theorem allows calculation of precision bounds for channel estimation optimized over probe states as given by \bar{S}_k using the conic programs that correspond to J_k . Since the cones considered in J_k are analogous to the cones considered in S_k [14], we know how to solve J_k numerically. Furthermore for $k = 2, 3, 4, 5$, we can solve J_k via semidefinite programming (SDP). In contrast, the calculation of J_1 requires the minimization over a certain separable cone on $\mathcal{R}_C \otimes (\mathcal{H}_B \otimes \mathcal{H}_A)$. Such a type conic programming is discussed in Section IV of [14].

In addition, in the single parameter case, i.e., $d = 1$, the SLD bound is attainable. (10) implies

$$J_1 = J_2 = J_3 = J_5 = J_4. \quad (24)$$

First, J_1 has the operational meaning of the precision limit under the scenario shown in Fig. 1. We can choose our input state for each single input system individually. Our measurement can be done over the joint system of the single output system and the ancilla system of the single input. In addition, we allow adaptive improvement for the choice of the input state and the measurement.

Second, J_5 has the operational meaning of the precision limit under the scenario shown in Fig. 2. Again, we can choose our input state for each single input system individually. Our measurement can be done over the joint system of all output systems and all ancilla systems of the inputs. However, when our measurement can be done over the joint system of the initial k_1 output systems and the initial k_1 ancilla systems of the inputs, we can adaptively choose the next k_2 input states depending on the above measurement outcome. Then, depending on the above measurement outcome, our measurement can be done over the joint system of the next k_2 output systems and the next k_2 ancilla systems of the inputs. Such an adaptive improvement over several rounds is allowed.

Now we sketch Theorem 1's proof. First we can establish an upper bound on J_k in terms of \bar{S}_k in the following lemma.

Lemma 2. For $k = 1, 2, 3, 4, 5$, we have

$$J_k \leq \bar{S}_k. \quad (25)$$

To prove Lemma 2, we show that given any solution ρ_{AC} and X to \bar{S}_k , we can also construct a corresponding solution Y for J_k with the same value for the objective function. We prove the inequality opposite to (25), based on the discussion in Section III, we rewrite the constraints for the completeness condition and the locally unbiased condition as

$$\begin{aligned} \text{Tr}_R[X(|0\rangle\langle 0| \otimes I_{BC})] &= I_{BC} \\ \frac{1}{2} \text{Tr}[(I_A \otimes X)((|0\rangle\langle j'| + |j'\rangle\langle 0|) \otimes F_j \otimes I_C)(I_{RB} \otimes \rho_{AC})] &= \delta_{j,j'}. \end{aligned} \quad (26)$$

Then, we show the following lemma.

Lemma 3. For $k = 1, 2, 3, 4, 5$, we choose $Y \in \mathcal{S}_{BA}^k$ satisfying the conditions (i), (ii), and $\rho_A := \rho_A(Y) > 0$. We diagonalize ρ_A as $\sum_{j=1}^{d_A} s_j |\phi_j\rangle\langle \phi_j|$. We choose an orthonormal basis $\{\psi_j\}$ of \mathcal{H}_C . We define a unitary map $U : \phi_j \mapsto \psi_j$ from \mathcal{H}_A to \mathcal{H}_C . We choose ρ_{AC} as the pure state $\sum_{j=1}^{d_A} \sqrt{s_j} |\phi_j, \psi_j\rangle$, which is a purification of ρ_A . Then, we have

$$\text{Tr}[Y(G \otimes T)] \geq \bar{S}_k. \quad (28)$$

In addition, we choose X as

$$X := (U \otimes I_{RB})(\rho_A^{-1/2} \otimes I_{RB})Y(\rho_A^{-1/2} \otimes I_{RB})(U^\dagger \otimes I_{RB}). \quad (29)$$

For $k = 1, 2, 3, 4$, X is an element of \mathcal{S}_{BC}^k , and satisfies the conditions (26) and (27), and the following relation.

$$\text{Tr}[Y(G \otimes T)] = \text{Tr}[\text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)](G \otimes T)]. \quad (30)$$

For $k = 5$, X is an element of \mathcal{S}_{BC}^4 , and satisfies the conditions (26) and (27), (30), and

$$\text{Tr}[(I_A \otimes X)((|i\rangle\langle j'| - |j'\rangle\langle i|) \otimes T \otimes I_C)(I_{RB} \otimes \rho_{AC})] = 0. \quad (31)$$

Lemma 3 assumes that Y satisfies the full rank condition for $\rho_A(Y)$. If the minimization in J_k is achieved by an operator Y to satisfy this condition, we obtain Theorem 1. However, there is a possibility that the optimal operator Y does not satisfy this condition. To cover such a possibility, we need the following technical lemma.

Lemma 4. *We have*

$$J_k = \inf_{Y \in \mathcal{S}_{BA}^k} \left\{ \text{Tr}[Y(G \otimes T)] \left| \begin{array}{l} Y \text{ satisfies (i), (ii) and,} \\ \rho_A(Y) > 0, \text{ i.e.,} \\ \rho_A(Y) \text{ is full rank.} \end{array} \right. \right\} \quad (32)$$

for $k = 1, 2, 3, 4$, and

$$J_5 = \inf_{Y \in \mathcal{S}_{BA}^5(T)} \left\{ \text{Tr}[Y(G \otimes T)] \left| \begin{array}{l} Y \text{ satisfies (i), (ii) and,} \\ \rho_A(Y) \text{ is full rank.} \end{array} \right. \right\} \quad (33)$$

Notice that both sides in the definition (29) act on $\mathcal{H}_R \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ because U maps \mathcal{H}_A to \mathcal{H}_C . The combination of Lemma 4 and (28) in Lemma 3 yields the \geq part of (23) while Lemma 2 shows the \leq part of (23). Therefore, the combination of Lemmas 2, 3, and 4 shows Theorem 1. In addition, Lemmas 2, 3, and 4 are shown in Appendix A.

The pure state ρ_{AC} given in Lemma 3 is the input state to achieves the value $\text{Tr}[Y(G \otimes T)]$ in the respective bounds. The operator X given in Lemma 3 achieves the value $\text{Tr}[Y(G \otimes T)]$ in the respective conic linear programming.

V. CASE WITH MAXIMALLY ENTANGLED INPUT STATE

A. Notations

The channel model is given as the pair $(T^{\otimes n}, (F_j^{(n)})_j)$. Fixing the input state to be the maximally entangled state $|\Phi\rangle\langle\Phi|$, for $k = 1, 2, 3, 4, 5$, gives the following inequalities.

$$J_k[T, (F_j)_j] \leq \bar{S}_k[T, (F_j)_j, |\Phi\rangle\langle\Phi|]. \quad (34)$$

The aim of this section is to derive a necessarily and sufficient condition for the equality of the above inequality for $k = 4, 5$.

For this aim, we discuss when a maximally entangled input state realizes the minimum value in the respective bounds. In this section, we simplify $J_k[T, (F_j)_j]$ and $\bar{S}_k[T, (F_j)_j, |\Phi\rangle\langle\Phi|]$ to J_k and $\bar{S}_k[|\Phi\rangle\langle\Phi|]$, respectively. We employ the normalized Choi state $T_N := \frac{1}{d_A} T$ on \mathcal{H}_{AB} . Also, we employ $F_{j,N} := \frac{1}{d_A} F_j$. Then, we impose the following conditions to $Y_N := d_A Y$.

(i-N): Given fixed Y_N on $\mathcal{H}_R \otimes \mathcal{H}_B \otimes \mathcal{H}_A$, there exists a state ρ_A on \mathcal{H}_A such that

$$\text{Tr}_R[Y_N(|0\rangle\langle 0| \otimes I_{AB})] = I_B \otimes d_A \rho_A. \quad (35)$$

(ii-N):

$$\frac{1}{2} \text{Tr}[Y_N((|0\rangle\langle j'| + |j'\rangle\langle 0|) \otimes F_{N,j})] = \delta_{j,j'}$$

for $j, j' = 1, \dots, d$.

The condition (i-N) is equivalent to the following linear constraint.

(i'-N): Let $\{|b\rangle\}_{b=1}^{d_B}$ be any orthonormal basis of \mathcal{H}_B . For $b \in \{1, \dots, d_B - 1\}$ and $b' \in \{2, \dots, d_B - 1\}$ with $b > b'$ we have

$$\text{Tr}_{RB}[Y_N(|0\rangle\langle 0| \otimes I_A \otimes |b\rangle\langle b'|)] = 0, \quad (36)$$

$$\text{Tr}_{RB}[Y_N(|0\rangle\langle 0| \otimes I_A \otimes (|b\rangle\langle b| - |b+1\rangle\langle b+1|))] = 0 \quad (37)$$

as operators on \mathcal{H}_A . Also,

$$\text{Tr}[Y_N(|0\rangle\langle 0| \otimes I_A \otimes |1\rangle\langle 1|)] = d_A. \quad (38)$$

Then, J_k for $k = 1, 2, 3, 4$ is rewritten as

$$J_k = \min_{Y_N \in \mathcal{S}_{BA}^k} \{ \text{Tr}[Y_N(G \otimes T_N)] | Y_N \text{ satisfies (i'-N), (ii-N)}. \} \quad (39)$$

J_5 is rewritten as

$$J_5 = \min_{Y_N \in \mathcal{S}_{BA}^5(T_N)} \{ \text{Tr}[Y_N(G \otimes T_N)] | Y_N \text{ satisfies (i'-N), (ii-N)}. \} \quad (40)$$

When the input state is fixed as a maximally entangled state $|\Phi\rangle$, the bound is $\bar{S}_k[|\Phi\rangle\langle\Phi|]$. Then, we have $T_N = \text{Tr}_A[(T \otimes I_C)(I_B \otimes |\Phi\rangle\langle\Phi|)]$. Hence, by replacing the system \mathcal{H}_C by \mathcal{H}_A , we rewrite $\bar{S}_k[|\Phi\rangle\langle\Phi|]$ as

$$\begin{aligned} & \bar{S}_k[|\Phi\rangle\langle\Phi|] \\ &= \min_{Y_N \in \mathcal{S}_{BA}^k} \{ \text{Tr}[Y_N(G \otimes T_N)] | Y_N \text{ satisfies (i''), (ii-N)}. \} \end{aligned} \quad (41)$$

$$\begin{aligned} &= S_k[T_N, (F_{j,N})_j], \\ & \bar{S}_5[|\Phi\rangle\langle\Phi|] \\ &= \min_{Y_N \in \mathcal{S}_{BA}^5(T_N)} \{ \text{Tr}[Y_N(G \otimes T_N)] | Y_N \text{ satisfies (i''), (ii-N)}. \} \end{aligned} \quad (42)$$

$$= S_5[T_N, (F_{j,N})_j],$$

for $k = 1, 2, 3, 4$, where the condition (i'') is defined as

(i''):

$$\text{Tr}_R[Y_N(|0\rangle\langle 0| \otimes I_{AB})] = I_{AB}. \quad (43)$$

Therefore, we find that the difference between J_k and $\bar{S}_k[|\Phi\rangle\langle\Phi|]$ is characterized by the difference between the conditions (i'-N) and (i''). To discuss this difference, we focus on $S_k[T_N, (F_{j,N})_j]$, i.e., the model $(T_N, (F_{j,N})_j)$ in the following discussion.

B. Equality condition for $k = 4$

We choose $L_*^i := L^i[T_N, (F_{j,N})_j]$, and define $\vec{L}_* = (L_*^i)_i$. For a vector of Hermitian matrices $\vec{Z} = (Z^j)_j$, we define the operator

$$W_{\text{SLD}}(T_N, \vec{Z}) := \sum_{1 \leq i, j \leq d} G_{j,i} Z^i T_N Z^j.$$

We define the subset $\mathcal{K} \subset \mathcal{B}_{sa}(\mathcal{H}_{AB})$ as

$$\mathcal{K} := \{X \in \mathcal{B}_{sa}(\mathcal{H}_{AB}) | \text{Tr}_B X \text{ is a constant times of } I_A\}$$

This subset relates to dual matrix-valued variable that corresponds to the condition (i'-N) in the primal problem. We supply the precise description of the dual problem in Appendix B 1.

Then, we have the following theorem.

Theorem 5. *The following conditions are equivalent.*

(A1):

$$J_4 = S_4[|\Phi\rangle\langle\Phi|].$$

(A2): $\text{Tr}_B W_{\text{SLD}}(T_N, \vec{L}_*)$ belongs to \mathcal{K} .

C. Equality condition for $k = 5$

Given the weight matrix G , for simplicity, we choose the new parameter $\tilde{\theta} := \sqrt{G}\theta$. By using the estimator $\hat{\theta}$ of the parameter θ , the new parameter's estimator $\tilde{\theta}$ is given as $\sqrt{G}\hat{\theta}$. Hence, by using the covariance matrix $V_{\tilde{\theta}}[\hat{\Pi}]$ of the parameter $\tilde{\theta}$, the covariance matrix $\tilde{V}_{\tilde{\theta}}[\hat{\Pi}]$ of the new parameter $\tilde{\theta}$ is given as $\sqrt{G}V_{\tilde{\theta}}[\hat{\Pi}]\sqrt{G}$. That is, we have

$$\text{Tr}GV_{\tilde{\theta}}[\hat{\Pi}] = \text{Tr}\tilde{V}_{\tilde{\theta}}[\hat{\Pi}].$$

In other words, under the new parameter $\tilde{\theta}$, the weight matrix is given as the identity matrix I so that the analysis on the weight matrix I can recover the case with a general weight matrix G . Therefore, without loss of generality, we can assume that the weight matrix is the identity matrix I .

We define \vec{Z}_* as

$$\vec{Z}_* := \underset{\vec{Z}}{\text{argmin}} \{ \text{Tr} \Pi(\vec{Z}) G \otimes T_N | \text{Tr} D_j Z^i = \delta_j^i \},$$

where $\Pi(\vec{Z})^{i,j} = Z^i Z^j$, i.e., $\Pi(\vec{Z}) = \sum_{1 \leq i, j \leq d} |i\rangle\langle j| \otimes (Z^i)^\dagger Z^j = (\sum_{i=1}^d |i\rangle \otimes (Z^i)^\dagger)(\sum_{j=1}^d \langle i| \otimes Z^j)$. We choose $Z_*^0 := I_{AB}$. We define the matrix $V_*^{i,j} := \text{Tr} Z_*^i (Z_*^j)^\dagger T_N$. We define $C_* := \text{Im } V_* | \text{Im } V_*|^{-1}$.

Next, we define

$$W_{HN}(T_n, \vec{Z}) := \left(\sum_{i=1}^d Z^i T_N Z^i \right) - \sum_{1 \leq i, j \leq d} \sqrt{-1} C_*^{i,j} Z^j T_N Z^i$$

which corresponds to a dual matrix-valued variable.

Theorem 6. *The following conditions are equivalent.*

(B1):

$$J_5 = S_5[|\Phi\rangle\langle\Phi|].$$

(B2): $\text{Tr}_B W_{HN}(T_n, \vec{Z}_*)$ belongs to \mathcal{K} .

D. Examples

1. One-parameter case

We choose \mathcal{H}_A and \mathcal{H}_B as two dimensional systems spanned by $|0\rangle, |1\rangle$. We define the channel Λ_0 as a depolarizing channel:

$$\Lambda_{0,p}(\rho) := (1-p)\rho + p\rho_{\text{mix},B},$$

where $\rho_{\text{mix},B}$ is the completely mixed state on the output system \mathcal{H}_B . The channel Λ_θ is given as $\Gamma_{\theta,p}(\rho) := U_\theta \Lambda_{0,p}(\rho) U_\theta^\dagger$, where $U_\theta := \exp(i\theta\sigma_1)$ and $\sigma_1 = |0\rangle\langle 1| + |1\rangle\langle 0|$. Then, the following theorem holds.

Theorem 7. *When G is 1, we have the following relation*

$$\begin{aligned} & J_k[(\Gamma_{\theta,p} \otimes \iota)(|I\rangle\langle I|), (\frac{d}{d\theta} \Gamma_{\theta,p} \otimes \iota)(|I\rangle\langle I|)] \\ &= \bar{S}_k[(\Gamma_{\theta,p} \otimes \iota)(|I\rangle\langle I|), (\frac{d}{d\theta} \Gamma_{\theta,p} \otimes \iota)(|I\rangle\langle I|), |\Phi\rangle\langle\Phi|] \\ &= \frac{2-p}{8(1-p)^2} \end{aligned} \quad (44)$$

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

Proof. Due to (24), it is sufficient to show (44) with $k = 4$. Also, due to the symmetry, it is sufficient to show the case with $\theta = 0$. Then, $T_N = (1-p)|\Phi\rangle\langle\Phi| + p\rho_{\text{mix},AB}$ and $F_{1,N} = (1-p)[i\sigma_1, |\Phi\rangle\langle\Phi|]$, where $\rho_{\text{mix},AB}$ is the completely mixed state on \mathcal{H}_{AB} . Since

$$\begin{aligned} & [\sigma_1, |\Phi\rangle\langle\Phi|] \circ |\Phi\rangle\langle\Phi| = \frac{1}{2}[\sigma_1, |\Phi\rangle\langle\Phi|] \\ & [\sigma_1, |\Phi\rangle\langle\Phi|] \circ \rho_{\text{mix},AB} = \frac{1}{4}[\sigma_1, |\Phi\rangle\langle\Phi|], \end{aligned}$$

using $\frac{1}{2}(1-p) + \frac{1}{4}p = \frac{1}{4}(2-p)$, we have $L_1 = 4\frac{1-p}{2-p}[i\sigma_1, |\Phi\rangle\langle\Phi|]$. Since $\sigma_1^2 = I$, the SLD Fisher information in this model is

$$\begin{aligned} & \text{Tr} L_1^2 T_N = \text{Tr} (4\frac{1-p}{2-p})^2 ([i\sigma_1, |\Phi\rangle\langle\Phi|])^2 T_N \\ &= 16 \frac{(1-p)^2}{(2-p)^2} \text{Tr} (\sigma_1 |\Phi\rangle\langle\Phi| \sigma_1 + |\Phi\rangle\langle\Phi|) T_N \\ &= 16 \frac{(1-p)^2}{(2-p)^2} \text{Tr} \left((1-p)|\Phi\rangle\langle\Phi| + \frac{p}{4}(\sigma_1 |\Phi\rangle\langle\Phi| \sigma_1 + |\Phi\rangle\langle\Phi|) \right) \\ &= 16 \frac{(1-p)^2}{(2-p)^2} ((1-p) + \frac{p}{2}) = 16 \frac{(1-p)^2}{(2-p)^2} (1 - \frac{1}{2}p) \\ &= \frac{8(1-p)^2}{2-p}. \end{aligned}$$

Hence, we obtain the second equation in (44) with $k = 4$.

Then, $L_*^1 = \frac{2-p}{8(1-p)^2} \cdot 4 \frac{1-p}{2-p} [i\sigma_1, |\Phi\rangle\langle\Phi|] = \frac{1}{2(1-p)} [i\sigma_1, |\Phi\rangle\langle\Phi|]$. Since

$$[i\sigma_1, |\Phi\rangle\langle\Phi|]|\Phi\rangle\langle\Phi| = \sigma_1|\Phi\rangle\langle\Phi|\sigma_1$$

$$[i\sigma_1, |\Phi\rangle\langle\Phi|]\rho_{\text{mix},AB}[i\sigma_1, |\Phi\rangle\langle\Phi|] = \frac{1}{4}(\sigma_1|\Phi\rangle\langle\Phi|\sigma_1 + |\Phi\rangle\langle\Phi|),$$

we have

$$W_{\text{SLD}}(T_N, \tilde{L}_*) = L_*^1 T_N L_*^1$$

$$= \frac{1}{4(1-p)^2} ((1-p)\sigma_1|\Phi\rangle\langle\Phi|\sigma_1 + \frac{p}{4}(\sigma_1|\Phi\rangle\langle\Phi|\sigma_1 + |\Phi\rangle\langle\Phi|))$$

$$= \frac{1}{4(1-p)^2} ((1 - \frac{3p}{4})\sigma_1|\Phi\rangle\langle\Phi|\sigma_1 + \frac{p}{4}|\Phi\rangle\langle\Phi|).$$

Since $\text{Tr}_B \sigma_1 |\Phi\rangle\langle\Phi| \sigma_1 = \text{Tr}_B |\Phi\rangle\langle\Phi| = \frac{1}{2} I_A$, we have

$$\text{Tr}_B \frac{1}{4(1-p)^2} ((1 - \frac{3p}{4})\sigma_1|\Phi\rangle\langle\Phi|\sigma_1 + \frac{p}{4}|\Phi\rangle\langle\Phi|)$$

$$= \frac{1}{4(1-p)^2} ((1 - \frac{3p}{4}) + \frac{p}{4}) \frac{1}{2} I_A = \frac{2-p}{16(1-p)^2} I_A.$$

Thus, $W_{\text{SLD}}(T_N, \tilde{L}_*)$ belongs to \mathcal{K} . Theorem 5 guarantees the first equation in (44) for $k = 4$. \square

2. Generalized Pauli channel

We consider the generalized Pauli channel on the system $\mathcal{H} = \mathcal{H}_A = \mathcal{H}_B$, which is spanned by $\{|a\rangle\}_{a \in \mathbb{Z}_d}$. We define the operators $W(a, b)$ for $a, b \in \mathbb{Z}_d$ as the following unitary matrices on \mathcal{H} ;

$$X(a) := \sum_{j \in \mathbb{Z}_d} |j+a\rangle\langle j|, \quad Z(b) := \sum_{j \in \mathbb{Z}_d} \omega^{bj} |j\rangle\langle j|, \quad (45)$$

$$W(a, b) := X(a)Z(b), \quad (46)$$

where $\omega := \exp(2\pi i/d)$. We introduce a distribution family p_θ over \mathbb{Z}_d^2 . Then, we define the family of channels $\{\Lambda_\theta\}$ as

$$\Lambda_\theta(\rho) := \sum_{(a,b) \in \mathbb{Z}_d^2} p_\theta(a, b) W(a, b) \rho W(a, b)^\dagger. \quad (47)$$

We denote the Fisher information of the distribution family $\{p_\theta\}$. Then, as shown in [36, 37], we have the following theorem.

Theorem 8. *We have the following relations*

$$J_4[(\Lambda_\theta \otimes \iota)(|I\rangle\langle I|), (\frac{\partial}{\partial \theta^l} \Lambda_\theta \otimes \iota)(|I\rangle\langle I|)_l]$$

$$= \bar{S}_4[(\Lambda_\theta \otimes \iota)(|I\rangle\langle I|), (\frac{\partial}{\partial \theta^l} \Lambda_\theta \otimes \iota)(|I\rangle\langle I|)_l, |\Phi\rangle\langle\Phi|]$$

$$= \text{Tr} G J_\theta^{-1}. \quad (48)$$

Proof. When the input state is the maximally entangled state $|\Phi\rangle\langle\Phi|$, the output state is

$$\sum_{(a,b) \in \mathbb{Z}_d^2} p_\theta(a, b) W(a, b) |\Phi\rangle\langle\Phi| W(a, b)^\dagger. \quad (49)$$

Since $\{W(a, b)|\Phi\rangle\}_{a,b}$ forms an orthogonal basis on $\mathcal{H}^{\otimes 2}$, the state family can be considered as distribution family $\{p_\theta\}$. Then, we obtain the second equation.

Next, we show the first equation. We denote the logarithmic likelihood derivative for the j -th parameter of $\{p_\theta\}$ by $l_{\theta,j}$. We define the function l_θ^j as $l_\theta^j := \sum_{j'} (J_\theta^{-1})^{j,j'} l_{\theta,j'}$. Then, we have $L_*^j = \sum_{(a,b) \in \mathbb{Z}_d^2} l_\theta^j(a, b) W(a, b) |\Phi\rangle\langle\Phi| W(a, b)^\dagger$. Hence,

$$\text{Tr}_B W_{\text{SLD}}(T_N, \tilde{L}_*)$$

$$= \text{Tr}_B \sum_{(a,b) \in \mathbb{Z}_d^2} \left(\sum_{1 \leq j', j \leq d} G_{j,j'} l_\theta^{j'}(a, b) p_\theta(a, b) l_\theta^j(a, b) \right)$$

$$\cdot W(a, b) |\Phi\rangle\langle\Phi| W(a, b)^\dagger$$

$$= \sum_{(a,b) \in \mathbb{Z}_d^2} \sum_{1 \leq j', j \leq d} G_{j,j'} l_\theta^{j'}(a, b) p_\theta(a, b) l_\theta^j(a, b) \rho_{\text{mix}}, \quad (50)$$

where ρ_{mix} is the completely mixed state on \mathcal{H}_C because $\text{Tr}_B W(a, b) |\Phi\rangle\langle\Phi| W(a, b)^\dagger = \text{Tr}_A |\Phi\rangle\langle\Phi| = \rho_{\text{mix}}$. Hence, $\text{Tr}_B W_{\text{SLD}}(T_N, \tilde{L}_*)$ belongs to \mathcal{K} . Theorem 5 guarantees the first equation. \square

3. Spin j representation of $SU(2)$

Next, we consider spin j representation of $SU(2)$ over the Hilbert space \mathcal{H}_j . Here, $\sigma_{1,j}$, $\sigma_{2,j}$, and $\sigma_{3,j}$ are defined as the spin j representations of the generators of $SU(2)$ on \mathcal{H}_j . We set \mathcal{H}_A and \mathcal{H}_B to be \mathcal{H}_j . We define the channel Λ_0 as a depolarizing channel:

$$\Lambda_{0,p}(\rho) := (1-p)\rho + p\rho_{\text{mix},B}.$$

The channel Λ_θ is given as $\Lambda_{\theta,p}(\rho) := U_\theta \Lambda_{0,p}(\rho) U_\theta^\dagger$, where $U_\theta := \exp(i \sum_{k=1}^3 \theta^k \sigma_k)$.

Then, the following theorem holds.

Theorem 9. *When the weight matrix G is chosen to be I , we have the following relations*

$$J_4[(\Lambda_{\theta,p} \otimes \iota)(|I\rangle\langle I|), (\frac{\partial}{\partial \theta^l} \Lambda_{\theta,p} \otimes \iota)(|I\rangle\langle I|)_{l=1,2,3}]$$

$$= \bar{S}_4[(\Lambda_{\theta,p} \otimes \iota)(|I\rangle\langle I|), (\frac{\partial}{\partial \theta^l} \Lambda_{\theta,p} \otimes \iota)(|I\rangle\langle I|)_{l=1,2,3}, |\Phi\rangle\langle\Phi|]$$

$$= \frac{9(1 - \frac{4j^2+4j-1}{(2j+1)^2} p)}{8j(j+1)(1-p)^2}. \quad (51)$$

In particular, for $p = 0$, we have

$$\begin{aligned}
& J_4[(\Lambda_{\theta,0} \otimes \iota)(|I\rangle\langle I|), (\frac{\partial}{\partial \theta^l} \Lambda_{\theta,0} \otimes \iota)(|I\rangle\langle I|)]_{l=1,2,3} \\
&= J_1[(\Lambda_{\theta,0} \otimes \iota)(|I\rangle\langle I|), (\frac{\partial}{\partial \theta^l} \Lambda_{\theta,0} \otimes \iota)(|I\rangle\langle I|)]_{l=1,2,3} \\
&= \bar{S}_1[(\Lambda_{\theta,0} \otimes \iota)(|I\rangle\langle I|), (\frac{\partial}{\partial \theta^l} \Lambda_{\theta,0} \otimes \iota)(|I\rangle\langle I|)]_{l=1,2,3}, |\Phi\rangle\langle\Phi| \\
&= \frac{9}{8j(j+1)}. \tag{52}
\end{aligned}$$

The paper [38] studied a similar problem in the n -copy setting of the estimation of $SU(D)$. It maximized the trace of SLD Fisher information matrix by varying the input state. This maximization is a different problem from our problem.

Proof. First, we show (51). Due to the symmetry, it is sufficient to show the case with $\theta = 0$. Then, $T_N = (1-p)|\Phi\rangle\langle\Phi| + p\rho_{\text{mix},AB}$ and $F_{l,N} = (1-p)[i\sigma_{l,j}, |\Phi\rangle\langle\Phi|]$, where $\rho_{\text{mix},AB}$ is the completely mixed state on \mathcal{H}_{AB} . Since

$$\begin{aligned}
& [\sigma_{l,j}, |\Phi\rangle\langle\Phi|] \circ |\Phi\rangle\langle\Phi| = \frac{1}{2}[\sigma_{l,j}, |\Phi\rangle\langle\Phi|] \\
& [\sigma_{l,j}, |\Phi\rangle\langle\Phi|] \circ \rho_{\text{mix},AB} = \frac{1}{(2j+1)^2}[\sigma_{l,j}, |\Phi\rangle\langle\Phi|],
\end{aligned}$$

using $\frac{1}{2}(1-p) + \frac{1}{(2j+1)^2}p = \frac{1}{2} - \frac{4j^2+4j-1}{2(2j+1)^2}p$, we have $L_l = c_{p,j}[i\sigma_{l,j}, |\Phi\rangle\langle\Phi|]$ with $c_{p,j} := \frac{1-p}{\frac{1}{2} - \frac{4j^2+4j-1}{2(2j+1)^2}p}$. Since $\langle\Phi|\sigma_{l',j}\sigma_{l,j}|\Phi\rangle = \frac{2j(j+1)}{3}\delta_{l,l'}$ and $\langle\Phi|\sigma_{l',j}|\Phi\rangle = 0$, the SLD Fisher information matrix in this model is

$$\begin{aligned}
& \text{Tr} L_l L_{l'} T_N = \text{Tr} c_{p,j}^2 [i\sigma_{l,j}, |\Phi\rangle\langle\Phi|][i\sigma_{l',j}, |\Phi\rangle\langle\Phi|] T_N \\
&= c_{p,j}^2 \text{Tr}(\sigma_{l,j}|\Phi\rangle\langle\Phi|\sigma_{l',j} + \frac{2j(j+1)}{3}\delta_{l,l'}|\Phi\rangle\langle\Phi|) T_N \\
&= c_{p,j}^2 \text{Tr}\left((1-p)\frac{2j(j+1)}{3}\delta_{l,l'}|\Phi\rangle\langle\Phi| \right. \\
&\quad \left. + \frac{p}{(2j+1)^2}(\sigma_{l,j}|\Phi\rangle\langle\Phi|\sigma_{l',j} + \frac{2j(j+1)}{3}\delta_{l,l'}|\Phi\rangle\langle\Phi|)\right) \\
&= \tilde{c}_{p,j}\delta_{l,l'},
\end{aligned}$$

where $\tilde{c}_{p,j} := c_{p,j}^2 \frac{2j(j+1)}{3}((1-p) + \frac{2p}{(2j+1)^2})$. Thus, the SLD Fisher information matrix is $\tilde{c}_{p,j}I$. Hence, the SLD bound \bar{S}_4 is $\text{Tr}\tilde{c}_{p,j}^{-1}I = 3/\tilde{c}_{p,j}$. Since we have

$$\begin{aligned}
& \tilde{c}_{p,j} = \left(\frac{1-p}{\frac{1}{2} - \frac{4j^2+4j-1}{2(2j+1)^2}p}\right)^2 \frac{2j(j+1)}{3} \left(1 - \frac{4j^2+4j-1}{(2j+1)^2}p\right) \\
&= \frac{4(1-p)^2}{1 - \frac{4j^2+4j-1}{(2j+1)^2}p} \frac{2j(j+1)}{3} = \frac{8j(j+1)(1-p)^2}{3(1 - \frac{4j^2+4j-1}{(2j+1)^2}p)}.
\end{aligned}$$

Hence, we obtain the second equation in (51).

Then, $L_*^l = \frac{c_{p,j}}{\tilde{c}_{p,j}}[i\sigma_{l,j}, |\Phi\rangle\langle\Phi|]$. Since

$$[i\sigma_{l,j}, |\Phi\rangle\langle\Phi|]|\Phi\rangle\langle\Phi| = \sigma_{l,j}|\Phi\rangle\langle\Phi|\sigma_{l',j}, \tag{53}$$

$$\begin{aligned}
& [i\sigma_{l,j}, |\Phi\rangle\langle\Phi|]\rho_{\text{mix},AB}[i\sigma_{l',j}, |\Phi\rangle\langle\Phi|] \\
&= \frac{1}{(2j+1)^2}(\sigma_{l,j}|\Phi\rangle\langle\Phi|\sigma_{l',j} + \frac{2j(j+1)}{3}\delta_{l,l'}|\Phi\rangle\langle\Phi|),
\end{aligned}$$

we have

$$\begin{aligned}
& W_{\text{SLD}}(T_N, \vec{L}_*) = \sum_{l=1}^3 L_*^l T_N L_*^l \\
&= \left(\frac{c_{p,j}}{\tilde{c}_{p,j}}\right)^2 \sum_{l=1}^3 \left((1-p)\sigma_{l,j}|\Phi\rangle\langle\Phi|\sigma_{l,j} \right. \\
&\quad \left. + \frac{p}{(2j+1)^2}(\sigma_{l,j}|\Phi\rangle\langle\Phi|\sigma_{l,j} + \frac{2j(j+1)}{3}|\Phi\rangle\langle\Phi|) \right) \\
&= \left(\frac{c_{p,j}}{\tilde{c}_{p,j}}\right)^2 \left((1-p + \frac{p}{(2j+1)^2}) \left(\sum_{l=1}^3 \sigma_{l,j}|\Phi\rangle\langle\Phi|\sigma_{l,j} \right) \right. \\
&\quad \left. + \frac{3p}{(2j+1)^2} \frac{2j(j+1)}{3} |\Phi\rangle\langle\Phi| \right).
\end{aligned}$$

Since $\sum_{l=1}^3 \sigma_{l,j}^2$ is a Casimir operator, it is $2j(j+1)I$ on \mathcal{H}_j . Thus,

$$\begin{aligned}
& \text{Tr}_B \sum_{l=1}^3 \sigma_{l,j}|\Phi\rangle\langle\Phi|\sigma_{l,j} = \text{Tr}_B \sum_{l=1}^3 \sigma_{l,j}^2 |\Phi\rangle\langle\Phi| \\
&= \text{Tr}_B 2j(j+1)|\Phi\rangle\langle\Phi| = \frac{2j(j+1)}{2j+1} I_A.
\end{aligned}$$

Since $\text{Tr}_B |\Phi\rangle\langle\Phi| = \frac{1}{2j+1} I_A$, we have

$$\begin{aligned}
& \text{Tr}_B W_{\text{SLD}}(T_N, \vec{L}_*) \\
&= \left(\frac{c_{p,j}}{\tilde{c}_{p,j}}\right)^2 \left((1-p + \frac{p}{(2j+1)^2}) \frac{2j(j+1)}{2j+1} \right. \\
&\quad \left. + \frac{3p}{(2j+1)^2} \frac{2j(j+1)}{3} \frac{1}{2j+1} \right) I_A \\
&= \left(\frac{c_{p,j}}{\tilde{c}_{p,j}}\right)^2 \left((1-p + \frac{2p}{(2j+1)^2}) \frac{2j(j+1)}{2j+1} \right) I_A.
\end{aligned}$$

Thus, $W_{\text{SLD}}(T_N, \vec{L}_*)$ belongs to \mathcal{K} . Theorem 5 guarantees the first equation of (51).

Next, with $p = 0$, we show

$$\begin{aligned}
& \bar{S}_1[(\Lambda_{\theta,0} \otimes \iota)(|I\rangle\langle I|), (\frac{\partial}{\partial \theta^l} \Lambda_{\theta,0} \otimes \iota)(|I\rangle\langle I|)]_{l=1,2,3}, |\Phi\rangle\langle\Phi| \\
&= \bar{S}_4[(\Lambda_{\theta,0} \otimes \iota)(|I\rangle\langle I|), (\frac{\partial}{\partial \theta^l} \Lambda_{\theta,0} \otimes \iota)(|I\rangle\langle I|)]_{l=1,2,3}, |\Phi\rangle\langle\Phi|. \tag{54}
\end{aligned}$$

Then, $T_N = |\Phi\rangle\langle\Phi|$ and $F_{l,N} = (i\sigma_{l,j})|\Phi\rangle\langle\Phi| + |\Phi\rangle\langle\Phi|(i\sigma_{l,j})^\dagger$. Since $\langle\Phi|(i\sigma_{l,j})^\dagger(i\sigma_{l',j})|\Phi\rangle = \frac{2j(j+1)}{3}\delta_{l,l'}$, the vectors $((i\sigma_{l,j})|\Phi\rangle)_{l=1,2,3}$ are orthogonal. Thus, Theorem 3 of [39]

guarantees that the SLD bound is attainable, i.e., (54). Since $\hat{S}_1 \geq J_1 \geq J_4$, the combination of (51) and (54) implies the first and second equations in (52). The third equation in (52) follows from (51). \square

VI. APPLICATION TO FIELD SENSING

A. Model for field sensing

The canonical example that is widely considered in quantum metrology is ‘field sensing’, where a classical field interacts with an ensemble of qubits. When a 3D classical field interacts identically with n qubits, we can write the interaction Hamiltonian as

$$H_\theta = \theta_1 E^1 + \theta_2 E^2 + \theta_3 E^3$$

where $\theta = (\theta_1, \theta_2, \theta_3)$ is a real vector that is proportional to the 3D field that we wish to estimate, and E^1, E^2, E^3 are angular momentum operators defined on n -qubits, given by

$$\begin{aligned} E^1 &= \frac{1}{2}(\tau_1^{(1)} + \dots + \tau_1^{(n)}), \\ E^2 &= \frac{1}{2}(\tau_2^{(1)} + \dots + \tau_2^{(n)}), \\ E^3 &= \frac{1}{2}(\tau_3^{(1)} + \dots + \tau_3^{(n)}), \end{aligned}$$

$\tau_1 = |0\rangle\langle 1| + |1\rangle\langle 0|$, $\tau_3 = |0\rangle\langle 0| - |1\rangle\langle 1|$ are Pauli matrices that apply the bit-flip and phase-flip on a qubit, $\tau_2 = i\tau_1\tau_3$, and $\tau_j^{(k)}$ represents an n -qubit Pauli matrix that applies τ_j on the k th qubit, and identity operations everywhere else.

Amplitude damping operators A_j , which model energy loss, apply $\gamma|0\rangle\langle 1|$ on the j th qubit and the identity operator on other qubits. These operators arise because of a linear interaction between individual qubits and a Markovian zero-temperature bath[40]. Namely,

$$A_{j,\gamma} = \gamma I^{\otimes j-1} \otimes |0\rangle\langle 1| \otimes I^{\otimes (n-j)}.$$

The collective amplitude damping operator, also considered in [33, Eq. (7)], models collective energy loss, specifically because of a collective linear interaction between all n qubits and a Markovian zero-temperature bath, and is given by $\tilde{A}_\gamma := \sum_{j=1}^n A_{j,\gamma}$.

We model the evolution of an initial probe state using the master equation

$$\frac{d\rho}{dt} = \mathcal{L}_{\theta,\gamma}(\rho), \quad (55)$$

where t denotes time, and the operator \mathcal{L}_θ can be written as a linear operator, which is

$$\mathcal{L}_{\theta,\gamma}(\rho) = -i(H_\theta\rho - \rho H_\theta) + \tilde{A}_\gamma\rho\tilde{A}_\gamma^\dagger - \frac{1}{2}(\tilde{A}_\gamma^\dagger\tilde{A}_\gamma\rho + \rho\tilde{A}_\gamma^\dagger\tilde{A}_\gamma). \quad (56)$$

For a non-negative evolution time s , let $\tilde{\rho}_s$ denote the solution to the master equation (55). In particular, we can write

$$\tilde{\rho}_s = e^{s\mathcal{L}_{\theta,\gamma}}(\tilde{\rho}_0).$$

This means that we can write $\tilde{\rho}_s$ as the Taylor series

$$\tilde{\rho}_s = \tilde{\rho}_0 + \sum_{k=1}^{\infty} \frac{s^k (\mathcal{L}_{\theta,\gamma})^k (\tilde{\rho}_0)}{k!}.$$

In our application, we set the evolution time as $s = 1$. Hence, the channel that we consider in our channel estimation problem is

$$\Lambda_{\theta,\gamma}(\rho) = e^{\mathcal{L}_{\theta,\gamma}}(\rho) = \rho + \sum_{k=1}^{\infty} \frac{\mathcal{L}_{\theta,\gamma}^k(\rho)}{k!}. \quad (57)$$

Using $\Lambda_{\theta,\gamma}$, we can calculate the corresponding Choi matrix $T_{\theta,\gamma}$, and its derivatives about the true parameter θ_0 are

$$F_{j,\gamma} = \frac{\partial}{\partial \theta_j} T_{\theta,\gamma}|_{\theta=\theta_0}$$

for $j = 1, 2, 3$.

B. Numerical results

Here, we set the true parameter as $\theta_0 = (0, 0, 0)$ and the number of qubits as $n = 2, 3, 4, 5$. We investigate our channel estimation problem when γ varies from 0 to 1. In our calculations, we set the weight matrix G as the identity matrix I .

We numerically evaluate the channel estimation precision bounds $J_{2,\gamma}, J_{4,\gamma}, J_{5,\gamma}$, where

$$J_{k,\gamma} = J_k[T_{\theta_0,\gamma}, (F_{1,\gamma}, F_{2,\gamma}, F_{3,\gamma})].$$

The precision bounds $J_{2,\gamma}, J_{4,\gamma}, J_{5,\gamma}$ are given by the optimal values of semidefinite programs with corresponding optimal solutions given by $Y_{2,\gamma}, Y_{4,\gamma}, Y_{5,\gamma}$, respectively. Using the optimal solutions $Y_{2,\gamma}, Y_{4,\gamma}, Y_{5,\gamma}$, we calculate the corresponding probe states on \mathcal{H}_A given by $\rho_A(Y_{2,\gamma}), \rho_A(Y_{4,\gamma}), \rho_A(Y_{5,\gamma})$ respectively.

Let us denote $|\Phi\rangle$ as the maximally entangled state on the symmetric subspace, and consider the precision bounds that correspond to using $|\Phi\rangle$ as the input probe state for the channel estimation problem. These precision bounds are

$$\begin{aligned} S_{k,\gamma}^{\text{sym}} := & S_k \left[\Lambda_{\theta_0,\gamma} \otimes \iota_C(|\Phi\rangle\langle\Phi|), \right. \\ & \left. \left(\frac{\partial}{\partial \theta_j} \Lambda_{\theta,\gamma} \otimes \iota_C(|\Phi\rangle\langle\Phi|) \Big|_{\theta=\theta_0} \right)_{j=1,2,3} \right]. \end{aligned}$$

We also numerically evaluate $S_{2,\gamma}^{\text{sym}}, S_{4,\gamma}^{\text{sym}}, S_{5,\gamma}^{\text{sym}}$. When we have no access to an ancillary system \mathcal{H}_C , we may consider using the 3D-GHZ state [41]

$$|\psi_{3\text{D GHZ}}\rangle = \frac{\sum_{j=0}^1 |j\rangle^{\otimes n} + |+\rangle^{\otimes n} + |-\rangle^{\otimes n} + |i\rangle^{\otimes n} + |-i\rangle^{\otimes n}}{N},$$

Plots of multiparameter Cramér-Rao bounds against increasing levels of noise (γ)

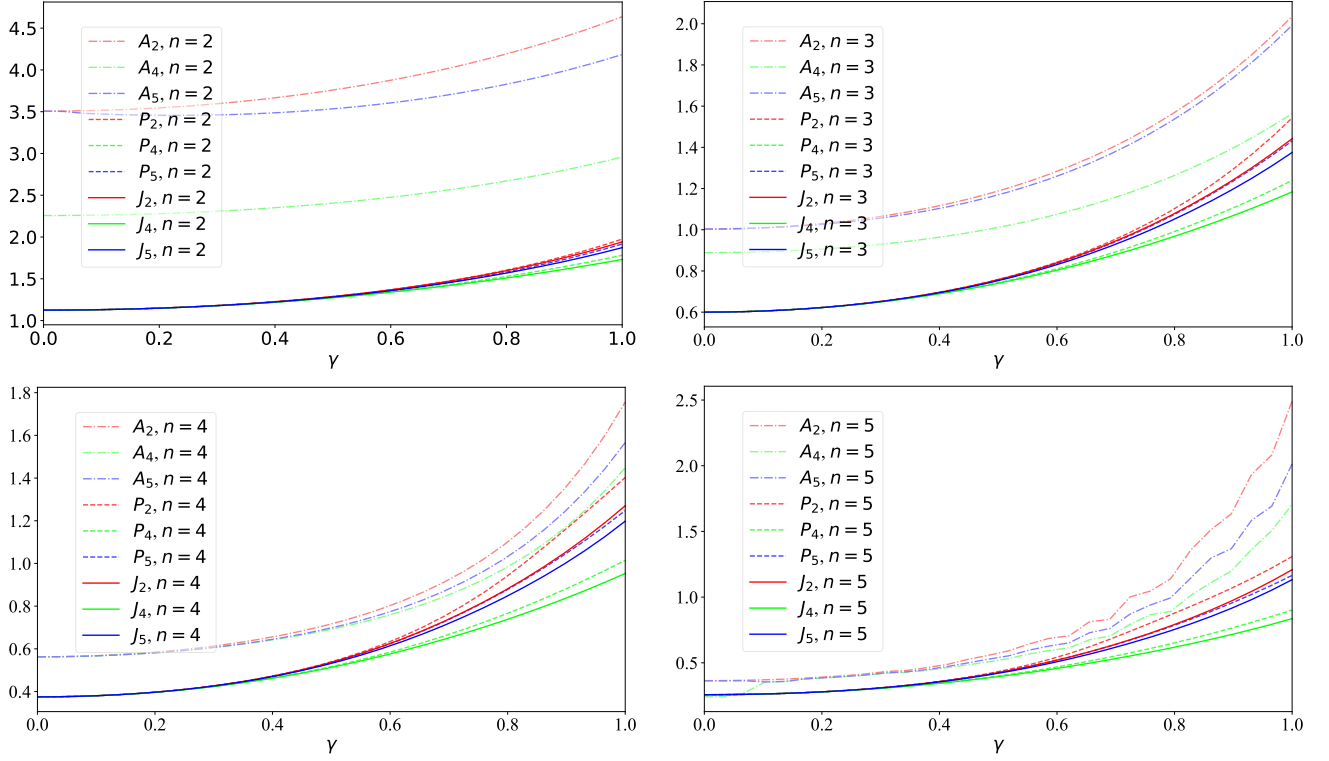


FIG. 4: We numerically calculate the CR-bounds for field sensing when $\theta_1 = \theta_2 = \theta_3 = 0$ for fixed number of qubits $n = 2, 3, 4, 5$. Here, $J_k := J_{k,\gamma}$ denote the CR-bounds using the optimal probe state with ancilla assistance. In contrast, $P_k := S_{k,\gamma}^{\text{sym}}$ denote the CR-bounds using the maximally entangled probe state on the symmetric subspace. We also use $A_k := S_{k,\gamma}^{3D}$ for the CR-type bounds using the 3D-GHZ state that do not require ancilla resistance. The vertical axis represents the CR-type lower bound on the sum of the variances of the field parameters θ_1, θ_2 and θ_3 , which corresponds to a weight matrix G equal to the identity matrix. The horizontal axis denotes the amount of amplitude damping noise γ .

Plots of multiparameter Cramér-Rao bounds against increasing number of qubits (n)

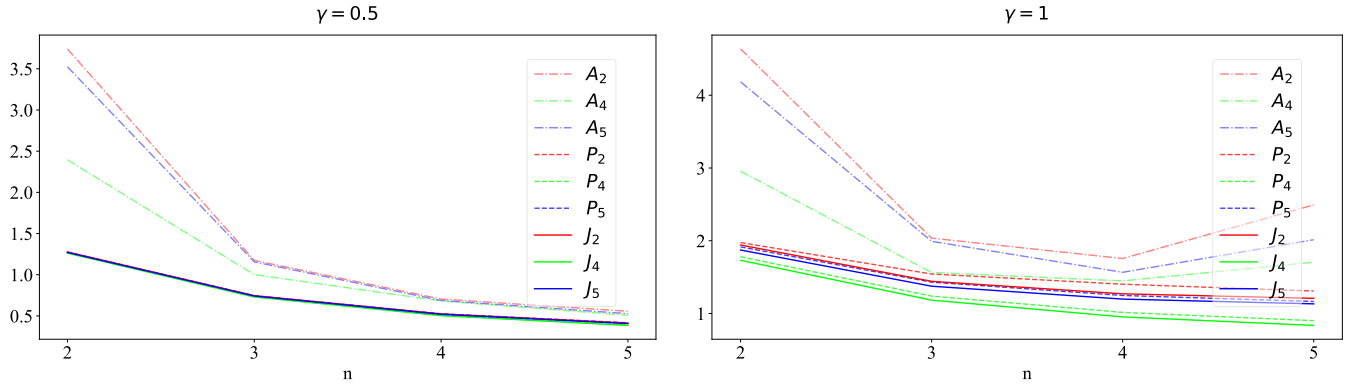


FIG. 5: We represent the data in Figure 4 differently, plotting the CR-bounds for fixed values of γ . The horizontal axis denotes the number of qubits n in the field sensing problem.

where $|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}$ and $|\pm i\rangle = \frac{|0\rangle \pm i|1\rangle}{\sqrt{2}}$, and N is the appropriate normalization factor. We denote the corresponding precision bounds for using the 3D-GHZ state without ancillary assistance as

$$S_{k,\gamma}^{3D} = S_k[\Lambda_{\theta_0}(\rho_{3DGHZ}), \frac{\partial}{\partial \theta_j} \Lambda_{\theta,\gamma}(\rho_{3DGHZ})|_{\theta=\theta_0}].$$

We also numerically calculate $S_{2,\gamma}^{3D}, S_{4,\gamma}^{3D}, S_{5,\gamma}^{3D}$.

We numerically find that when $\gamma = 0$, we have

$$J_{2,0} = J_{4,0} = J_{5,0} = S_{2,0}^{\text{sym}} = S_{4,0}^{\text{sym}} = S_{5,0}^{\text{sym}} = \frac{9}{n(n+2)}.$$

In fact, the analysis in Subsection VD3 showed that the same equality when the input state is limited to a state on the symmetric subspace. Our numerical analysis suggests that the support of the optimal input state is limited to the symmetric subspace.

Furthermore, when we solve the semidefinite programs corresponding to $J_{2,\gamma}, J_{4,\gamma}, J_{5,\gamma}$, we find that the corresponding optimal solutions $Y_{2,\gamma}, Y_{4,\gamma}, Y_{5,\gamma}$ have corresponding density matrices $\rho_A(Y_{2,\gamma}), \rho_A(Y_{4,\gamma}), \rho_A(Y_{5,\gamma})$ that are very close to the completely mixed state.

When $\gamma > 0$ we numerically find that

$$S_2^{3D} > S_2^{\text{sym}} > J_2, \quad S_4^{3D} > S_4^{\text{sym}} > J_4, \quad S_5^{3D} > S_5^{\text{sym}} > J_5. \quad (58)$$

In Figure 4, for different fixed values of $n = 2, 3, 4, 5$, we plot the precision bounds $S_{k,\gamma}^{3D}, S_{k,\gamma}^{\text{sym}}$ and $J_{k,\gamma}$ on the vertical axis and $\gamma \in [0, 1]$ on the horizontal axes. In Figure 5, for fixed $\gamma = 0.5$ or $\gamma = 1$, we plot the precision bounds A_k, P_k and J_k on the vertical axis, and the number of qubits n on the horizontal axis.

In this way, we confirm the suboptimality of the maximally entangled state by our numerical evaluation of $\rho_A(Y)$ using the optimal solution Y for the conic programs $J_{2,\gamma}, J_{4,\gamma}$ and $J_{5,\gamma}$ for γ . In particular, while $\rho_A(Y_{2,\gamma}), \rho_A(Y_{4,\gamma})$, and $\rho_A(Y_{5,\gamma})$ are still a diagonal matrices when $\gamma > 0$, they are not completely mixed states. Hence, their purifications cannot be the maximally entangled state. We can furthermore see the suboptimality of the maximally entangled state because $S_{k,\gamma}^{\text{sym}} \geq J_{k,\gamma}$ for $k = 2, 4, 5$ when $\gamma > 0$. Curiously, from Figure 4, the maximally entangled state on the symmetric subspace is nonetheless still quite close to optimal.

Motivating the above analysis, we prove the following theorem. When the support of the input state is limited to the symmetric subspace in the channel Λ_θ , we denote the obtained channel by $\Lambda_{\theta,\gamma}^{\text{sym}}$. Using this limitation for the inputs, we define $T_{\theta,\gamma}^{\text{sym}}$ and $F_{j,\gamma}^{\text{sym}}$ in the same way.

Theorem 10. *For $\gamma > 0$, we have*

$$\begin{aligned} J_4[T_{\theta_0,\gamma}^{\text{sym}}, (F_{j,\gamma}^{\text{sym}})_{j=1,2,3}] \\ &< S_{k,\gamma}^{\text{sym}} \\ &= S_4[\Lambda_{\theta_0,\gamma}^{\text{sym}} \otimes \iota_C(|\Phi\rangle\langle\Phi|), \\ &\quad \left(\frac{\partial}{\partial \theta_j} \Lambda_{\theta,\gamma}^{\text{sym}} \otimes \iota_C(|\Phi\rangle\langle\Phi|) \Big|_{\theta=\theta_0} \right)_{j=1,2,3}]. \end{aligned}$$

Since the relation $J_4[T_{\theta_0,\gamma}, (F_{j,\gamma})_{j=1,2,3}] \leq J_4[T_{\theta_0,\gamma}^{\text{sym}}, (F_{j,\gamma}^{\text{sym}})_{j=1,2,3}]$ holds, the above theorem means that the maximally entangled state on the symmetric subspace is not the optimal input state.

C. Methodology of numerical analysis

Now we use $\rho' = \Lambda_{\theta,\gamma}^{\approx}(\rho)$ to approximate $\Lambda_{\theta,\gamma}(\rho)$ according to the following procedure: Given any input state ρ , we define

$$\tilde{m}_*(\rho) = \min(100, \min\{j : \|(\mathcal{L}_{\theta,\gamma}^j(\rho)/j!\| < 10^{-12})\}).$$

Then we define

$$\Lambda_{\theta,\gamma}^{\approx}(\rho) := \sum_{j=0}^{m_*(\rho)} \frac{\mathcal{L}_{\theta,\gamma}^j(\rho)}{j!}.$$

On the state ρ_{AC} , we also define

$$\Lambda_{\theta,\gamma}^{\approx}(\rho_{AC}) := \sum_{j=0}^{m_*(\rho_{AC})} \frac{(\mathcal{L}_{\theta,\gamma}^j \otimes \iota_C)(\rho_{AC})}{j!}.$$

We obtain approximations of $T_{\theta,\gamma}$ with

$$T_{\theta,\gamma}^{\approx} = \Lambda_{\theta,\gamma}^{\approx}((n+1)|\Phi\rangle\langle\Phi|).$$

We also obtain approximations of F_1, F_2, F_3 according to the formula

$$\begin{aligned} F_{1,\gamma}^{\approx} &= \frac{T_{(10^{-12},0,0),\gamma}^{\approx} - T_{(0,0,0),\gamma}^{\approx}}{10^{-12}}, \\ F_{2,\gamma}^{\approx} &= \frac{T_{(0,10^{-12},0),\gamma}^{\approx} - T_{(0,0,0),\gamma}^{\approx}}{10^{-12}}, \\ F_{3,\gamma}^{\approx} &= \frac{T_{(0,0,10^{-12}),\gamma}^{\approx} - T_{(0,0,0),\gamma}^{\approx}}{10^{-12}}. \end{aligned}$$

With the above, we approximate $J_{2,\gamma}, J_{4,\gamma}, J_{5,\gamma}$ with

$$J_{k,\gamma}^{\approx} = J_k[T_{(0,0,0),\gamma}^{\approx}, (F_{j,\gamma}^{\approx})_{j=1,2,3}].$$

We also approximate P_2, P_4, P_5 with

$$S_{k,\gamma}^{\text{sym},\approx} = S_{k,\gamma}[T_{(0,0,0),\gamma}^{\approx}/(n+1), (F_{j,\gamma}^{\approx}/(n+1))_{j=1,2,3}].$$

We approximate $\Lambda_{(0,0,0)}(\rho_{3DGHZ})$ according to formula

$$Q_0 = \Lambda_{(0,0,0),\gamma}^{\approx}(\rho_{3DGHZ})$$

and approximate $\frac{\partial}{\partial \theta_j} \Lambda_\theta(\rho_{3DGHZ})|_{\theta=(0,0,0),\gamma}$ according to

$$\begin{aligned} Q_1^{\approx} &= \frac{\Lambda_{(10^{-12},0,0),\gamma}^{\approx}(\rho_{3DGHZ}) - \Lambda_{(0,0,0),\gamma}^{\approx}(\rho_{3DGHZ})}{10^{-12}}, \\ Q_2^{\approx} &= \frac{\Lambda_{(0,10^{-12},0),\gamma}^{\approx}(\rho_{3DGHZ}) - \Lambda_{(0,0,0),\gamma}^{\approx}(\rho_{3DGHZ})}{10^{-12}}, \\ Q_3^{\approx} &= \frac{\Lambda_{(0,0,10^{-12}),\gamma}^{\approx}(\rho_{3DGHZ}) - \Lambda_{(0,0,0),\gamma}^{\approx}(\rho_{3DGHZ})}{10^{-12}}. \end{aligned}$$

This allows us to approximate $S_{2,\gamma}^{3D}, S_{4,\gamma}^{3D}, S_{5,\gamma}^{3D}$ according to

$$S_{k,\gamma}^{3D,\approx} = S_k[Q_0, (Q_1, Q_2, Q_3)].$$

In our numerical evaluations of $S_{k,\gamma}^{3D,\approx}, S_{k,\gamma}^{\text{sym},\approx}, J_{k,\gamma}^{\approx}$, we evaluate the semidefinite programs according to MatLab code given in Appendix D.

D. Proof of Theorem 10

Next, we consider $\theta = (0, 0, 0)$, $\gamma > 0$, and where we fix the input probe state as the maximally entangled state $|\Phi\rangle\langle\Phi|$. In this case, after the channel acts on our input probe state, we have the output state $\rho_\gamma := (\Lambda_{0,\gamma} \otimes \iota_C)(|\Phi\rangle\langle\Phi|)$. Letting $F^u := E^u \otimes I_C$, we define

$$\begin{aligned} L_1 &:= L_1[\rho_\gamma, (i[\rho_\gamma, F^1], i[\rho_\gamma, F^2], i[\rho_\gamma, F^3])] \\ L_2 &:= L_2[\rho_\gamma, (i[\rho_\gamma, F^1], i[\rho_\gamma, F^2], i[\rho_\gamma, F^3])]. \end{aligned}$$

Then, we have the following lemma.

Lemma 11.

$$\begin{aligned} &\text{Tr}[L_1 \rho_\gamma L_2] - \text{Tr}[L_2 \rho_\gamma L_1] \\ &= \frac{-2in(n+2)}{3} (1 + n(n+2)/3) \gamma^2 + O(\gamma^4). \end{aligned} \quad (59)$$

Since $\text{Tr}[L_1 \rho_\gamma L_2] - \text{Tr}[L_2 \rho_\gamma L_1]$ is not equal to zero, this proves that the SLD bound $S_{k,\gamma}^{\text{sym}}$ is not a tight bound. That is, we obtain Theorem 10.

In the following, we show Lemma 11. For this aim, we handle the symmetric subspace. The symmetric subspace is described by using the Dicke basis $\{|D_0^n\rangle, \dots, |D_n^n\rangle\}$ where the Dicke states

$$|D_w^n\rangle := \sum_{\substack{(x_1, \dots, x_n) \in \{0,1\}^n \\ x_1 + \dots + x_n = w}} \frac{1}{\sqrt{\binom{n}{w}}} (|x_1\rangle \otimes \dots \otimes |x_n\rangle)$$

are uniform superpositions of all computational states in $\{|0\rangle, |1\rangle\}^{\otimes n}$ of constant Hamming weight w with unit norm. Then, the maximally entangled state $|\Phi\rangle$ on the symmetric subspace has the form $|\Phi\rangle = \frac{1}{\sqrt{n+1}} \sum_{u=0}^n |D_u^n\rangle \otimes |D_u^n\rangle$, and we introduce two orthogonal vectors

$$\begin{aligned} |\phi\rangle &:= ((I_A - \frac{1}{2} \tilde{A}_\gamma^* \tilde{A}_\gamma) \otimes I_C) |\Phi\rangle \\ &= \frac{1}{\sqrt{n+1}} \sum_{u=0}^n (I_A - \frac{1}{2} \tilde{A}_\gamma^* \tilde{A}_\gamma) |D_u^n\rangle \otimes |D_u^n\rangle, \end{aligned} \quad (60)$$

$$|\epsilon\rangle := (\tilde{A}_\gamma \otimes I_C) |\Phi\rangle = \frac{1}{\sqrt{n+1}} \sum_{u=0}^n \tilde{A}_\gamma |D_u^n\rangle \otimes |D_u^n\rangle. \quad (61)$$

The vectors $|\phi\rangle$ and $|\epsilon\rangle$ are orthogonal because of the following reason. First, according to (C7), when $u = 1, \dots, n$, $\tilde{A}_\gamma |D_u^n\rangle = \gamma \sqrt{(n-u+1)u} |D_{u-1}^n\rangle$, and when $u = 0$, $\tilde{A}_\gamma |D_u^n\rangle = 0$, and hence $\tilde{A}_\gamma |D_u^n\rangle = a_u |D_{u-1}^n\rangle$ for

some real number a_u . Second, according to (C7) $(I_A - \frac{1}{2} \tilde{A}_\gamma^* \tilde{A}_\gamma) |D_u^n\rangle = b_u |D_u^n\rangle$ for some real number b_u . Hence $\langle\phi|\epsilon\rangle = \frac{1}{n+1} \sum_{u=1}^n a_u b_u \langle D_{u-1}^n | D_u^n \rangle$. Since $|D_u^n\rangle$ and $|D_{u-1}^n\rangle$ are pairwise orthogonal for $u = 1, \dots, n$, it follows that $\langle\phi|\epsilon\rangle = 0$, and hence $|\phi\rangle$ and $|\epsilon\rangle$ are orthogonal.

It is easy to see that

$$\begin{aligned} \lambda_1 &:= \langle\phi|\phi\rangle = \frac{1}{n+1} \sum_{w=0}^n (1 - \gamma^2(n-w+1)w/2)^2 \\ &= 1 - \frac{\gamma^2}{6} n(n+2) + O(\gamma^4), \end{aligned} \quad (62)$$

$$\lambda_2 := \langle\epsilon|\epsilon\rangle = \frac{\gamma^2}{n+1} \sum_{w=0}^{n-1} (n-w)(w+1) = \frac{\gamma^2}{6} n(n+2). \quad (63)$$

Using $|\phi_1\rangle := |\phi\rangle/\sqrt{\lambda_1}$ and $|\phi_2\rangle := |\epsilon\rangle/\sqrt{\lambda_2}$, we define

$$\bar{\rho}_\gamma := |\phi\rangle\langle\phi| + |\epsilon\rangle\langle\epsilon| = \lambda_1 |\phi_1\rangle\langle\phi_1| + \lambda_2 |\phi_2\rangle\langle\phi_2|. \quad (64)$$

Hence, the two largest eigenvalues of $\bar{\rho}_\gamma$ are λ_1 and λ_2 , which are given by $1 - \frac{\gamma^2}{6} n(n+2) + O(\gamma^4)$ and $\frac{\gamma^2}{6} n(n+2)$ respectively. Using these, we define two quantities

$$\begin{aligned} f(1, 2) &:= \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} = 1 - \frac{\gamma^2}{3} n(n+2) + O(\gamma^4), \\ g &:= f(1, 2)^2 - 1. \end{aligned} \quad (65)$$

Then, we have

$$g = -\frac{2\gamma^2 n(n+2)}{3} + O(\gamma^4). \quad (66)$$

Introducing two operators

$$\begin{aligned} \bar{L}_1 &:= L_1[\bar{\rho}_\gamma, (i[\bar{\rho}_\gamma, F^1], i[\bar{\rho}_\gamma, F^2], i[\bar{\rho}_\gamma, F^3])] \\ \bar{L}_2 &:= L_2[\bar{\rho}_\gamma, (i[\bar{\rho}_\gamma, F^1], i[\bar{\rho}_\gamma, F^2], i[\bar{\rho}_\gamma, F^3])], \end{aligned}$$

we have the following two lemmas.

Lemma 12. Let $u, v = 1, 2, 3$. We have

$$\begin{aligned} \text{Tr}(\bar{L}_u \bar{\rho}_\gamma \bar{L}_v) &= -4 \sum_{l=1}^2 \langle\phi_l|F^u|\phi_l\rangle \langle\phi_l|F^v|\phi_l\rangle \\ &\quad + 4g\lambda_2 \langle\phi_1|F^u|\phi_2\rangle \langle\phi_2|F^v|\phi_1\rangle \\ &\quad + 4g\lambda_1 \langle\phi_2|F^u|\phi_1\rangle \langle\phi_1|F^v|\phi_2\rangle \\ &\quad + 4 \sum_{l=1}^2 \lambda_l \langle\phi_l|F^v F^u|\phi_l\rangle. \end{aligned} \quad (67)$$

Lemma 13. Let $\mu := n(n+2)/12$. Then

$$\langle\phi|F^1|\epsilon\rangle = \mu\gamma + O(\gamma^3), \quad (68)$$

$$\langle\phi|F^2|\epsilon\rangle = i\mu\gamma + O(\gamma^3), \quad (69)$$

$$\langle\phi|F^3|\phi\rangle = \mu\gamma^2, \quad (70)$$

$$\langle\epsilon|F^3|\epsilon\rangle = \mu\gamma^2 + O(\gamma^4). \quad (71)$$

Lemma 12 is shown in Appendix C3, and Lemma 13 is shown in Appendix C4.

Using (67) with $(u, v) = (1, 2)$, we find that

$$\begin{aligned}
& \text{Tr}[\bar{L}_1 \bar{\rho}_\gamma \bar{L}_2] - \text{Tr}[\bar{L}_2 \bar{\rho}_\gamma \bar{L}_1] \\
& \stackrel{(a)}{=} 4g\lambda_2 \langle \phi_1 | F^1 | \phi_2 \rangle \langle \phi_2 | F^2 | \phi_1 \rangle + 4g\lambda_1 \langle \phi_2 | F^1 | \phi_1 \rangle \langle \phi_1 | F^2 | \phi_2 \rangle \\
& \quad - 4g\lambda_2 \langle \phi_1 | F^2 | \phi_2 \rangle \langle \phi_2 | F^1 | \phi_1 \rangle - 4g\lambda_1 \langle \phi_2 | F^2 | \phi_1 \rangle \langle \phi_1 | F^1 | \phi_2 \rangle \\
& \quad + 4 \sum_{l=1}^2 \lambda_l \langle \phi_l | F^2 F^1 | \phi_l \rangle - 4 \sum_{l=1}^2 \lambda_l \langle \phi_l | F^1 F^2 | \phi_l \rangle \\
& \stackrel{(b)}{=} 4g\lambda_2 \langle \phi_1 | F^1 | \phi_2 \rangle \langle \phi_2 | F^2 | \phi_1 \rangle + 4g\lambda_1 \langle \phi_2 | F^1 | \phi_1 \rangle \langle \phi_1 | F^2 | \phi_2 \rangle \\
& \quad - 4g\lambda_2 \langle \phi_1 | F^2 | \phi_2 \rangle \langle \phi_2 | F^1 | \phi_1 \rangle - 4g\lambda_1 \langle \phi_2 | F^2 | \phi_1 \rangle \langle \phi_1 | F^1 | \phi_2 \rangle \\
& \quad - 4i(\lambda_1 \langle \phi_1 | F^3 | \phi_1 \rangle + \lambda_2 \langle \phi_2 | F^3 | \phi_2 \rangle) \\
& \stackrel{(c)}{=} 8g(\lambda_2 \langle \phi_1 | F^1 | \phi_2 \rangle \langle \phi_2 | F^2 | \phi_1 \rangle + \lambda_1 \langle \phi_2 | F^1 | \phi_1 \rangle \langle \phi_1 | F^2 | \phi_2 \rangle) \\
& \quad - 4i(\lambda_1 \langle \phi_1 | E^3 | \phi_1 \rangle + \lambda_2 \langle \phi_2 | E^3 | \phi_2 \rangle). \tag{72}
\end{aligned}$$

In (a) we use $\langle \phi_1 | F^2 | \phi_1 \rangle = 0$ and $\langle \phi_2 | F^2 | \phi_2 \rangle = 0$ so that the first summation in (67) vanishes. In (b) we use $\langle \phi_j | (F^2 F^1 - F^1 F^2) | \phi_j \rangle = -i \langle \phi_j | F^3 | \phi_j \rangle$. In (c) we use the following. First $i \langle \phi_1 | F^2 | \phi_2 \rangle \in \mathbb{R}$ implies that $\langle \phi_1 | F^2 | \phi_2 \rangle = -\langle \phi_2 | F^2 | \phi_1 \rangle$. Second, $\langle \phi_1 | F^1 | \phi_2 \rangle \in \mathbb{R}$ implies that $\langle \phi_1 | F^1 | \phi_2 \rangle = \langle \phi_2 | F^1 | \phi_1 \rangle$. Hence $\langle \phi_2 | F^1 | \phi_1 \rangle \langle \phi_1 | F^2 | \phi_2 \rangle = -\langle \phi_1 | F^1 | \phi_2 \rangle \langle \phi_2 | F^2 | \phi_1 \rangle$, which implies that $\langle \phi_2 | F^1 | \phi_1 \rangle \langle \phi_1 | F^2 | \phi_2 \rangle - \langle \phi_1 | F^1 | \phi_2 \rangle \langle \phi_2 | F^2 | \phi_1 \rangle = 2 \langle \phi_2 | F^1 | \phi_1 \rangle \langle \phi_1 | F^2 | \phi_2 \rangle$, which implies (c).

Then, we have

$$\begin{aligned}
& \text{Tr}[\bar{L}_1 \bar{\rho}_\gamma \bar{L}_2] - \text{Tr}[\bar{L}_2 \bar{\rho}_\gamma \bar{L}_1] \\
& \stackrel{(a)}{=} 8g(\langle \phi | F^1 | \epsilon \rangle \langle \epsilon | F^1 | \phi \rangle / \lambda_1 + \langle \epsilon | F^1 | \phi \rangle \langle \phi | F^2 | \epsilon \rangle / \lambda_2) \\
& \quad - 4i(\langle \phi | F^3 | \phi \rangle + \langle \epsilon | F^3 | \epsilon \rangle) \\
& \stackrel{(b)}{=} 8g(-i\mu^2 \gamma^2 + O(\gamma^4)) + 8g((i\mu^2 \gamma^2 + O(\gamma^4)) / (2\mu \gamma^2)) \\
& \quad - 4i(2\mu \gamma^2 + O(\gamma^4)) \\
& \stackrel{(c)}{=} 4i\mu g - 8i\mu \gamma^2 + O(\gamma^4) \\
& \stackrel{(d)}{=} (-32i\mu^2 - 8i\mu) \gamma^2 + O(\gamma^4) \\
& \stackrel{(e)}{=} -8i(4\mu^2 + \mu) \gamma^2 + O(\gamma^4) \\
& \stackrel{(f)}{=} \frac{-2in(n+2)}{3} (1 + n(n+2)/3) \gamma^2 + O(\gamma^4). \tag{73}
\end{aligned}$$

In (a), we use $|\phi\rangle = |\phi_1\rangle / \sqrt{\lambda_1}, |\epsilon\rangle = |\phi_2\rangle / \sqrt{\lambda_2}$, and (72). In (b), we use Lemma 13, and the fact that $\lambda_1 = \langle \phi | \phi \rangle, \lambda_2 = \langle \epsilon | \epsilon \rangle$ with (62) and (63). In (c), we collect terms of leading order in γ . In (d), we use (66), namely, $g = -8\mu \gamma^2 + O(\gamma^4)$. In (e), we substitute the value of $\mu = n(n+2)/12$.

In Appendix C5, we prove the following lemma.

Lemma 14. *Let $u, v = 1, 2, 3$. Then*

$$\text{Tr}[L_u \rho_\gamma L_v] - \text{Tr}[\bar{L}_u \bar{\rho}_\gamma \bar{L}_v] = O(\gamma^4). \tag{74}$$

Then from (73) and Lemma 14, we can get Lemma 11.

VII. DISCUSSIONS

We have shown that the channel estimation problems that correspond to S_j admit corresponding conic optimizations J_j that allow us to find the corresponding Cramér-Rao bound for the optimal entangled state. We illustrate the power of our framework with theoretical analysis on the scenario when the maximally entangled state is the optimal probe state, and also with numerical analysis for the often studied field sensing problem using quantum probe states.

We believe that the theory we develop here has applications that extend beyond the problem of field sensing using probe states. In fact, any problem where we estimate multiple incompatible parameters embedded within a quantum channel using entangled probe states stands to benefit from our theory. This encompasses many applications beyond that of quantum field sensing, such as in quantum imaging [31].

In a recent paper [42], it was shown that field sensing using quantum probe states in the face of a linear rate of errors can approach the Heisenberg limit if we use quantum error correction on appropriate permutation-invariant codes [43, 44]. However, the corresponding question of what can be done in the setting of using entangled probe states was not considered, and remains an interesting open problem.

Next we like to discuss the efficiency of evaluating the various bounds J_j . When $j = 2, \dots, 5$, the optimizations for J_j are efficiently solvable by SDPs. When $j = 1$, although the conic programming S_1 cannot be considered as a SDP, it can be approximately calculated with SDP by employing the concept of symmetric extension. Originally, symmetric extension was introduced to considering the membership problem for the separability [45, 46]. We consider the systems $\mathbb{C}^{d+1 \otimes n} \otimes \mathcal{H}, \mathbb{C}^{d+1} \otimes \mathcal{H}^{\otimes n}$ and define

$$\begin{aligned}
S_n^1 &:= \{X \in \mathcal{B}(\mathbb{C}^{d+1 \otimes n}, \mathcal{H}) | \text{Tr}_{j^c} X = \text{Tr}_{1^c} X, X \geq 0\} \\
\tilde{S}_n^1 &:= \{X \in \mathcal{B}(\mathbb{C}^{d+1}, \mathcal{H}^{\otimes n}) | \text{Tr}_{j^c} X = \text{Tr}_{1^c} X, X \geq 0\},
\end{aligned}$$

where Tr_{j^c} expresses the partial trace except for the j -th system on $\mathbb{C}^{d+1 \otimes n}$ or $\mathcal{H}^{\otimes n}$. The minimizer X_* of S_1 has a symmetric extension $X_{*,n}$ that belongs to S_n^1 and \tilde{S}_n^1 and satisfies $\text{Tr}_{1^c} X_{*,n} = X_*$. Due to the condition $\text{Tr}_{1^c} X_{*,n} = X_* \in S_1$, we have the following lower bounds of S_1 :

$$S_{1,n} := \min_{X \in S_n^1} \{\text{Tr}(G \otimes \rho) \text{Tr}_{1^c} X | \text{Tr}_{1^c} \text{ satisfies (4), (5)}\} \tag{75}$$

$$\tilde{S}_{1,n} := \min_{X \in \tilde{S}_n^1} \{\text{Tr}(G \otimes \rho) \text{Tr}_{1^c} X | \text{Tr}_{1^c} \text{ satisfies (4), (5)}\} \tag{76}$$

The above quantities can be calculated by SDP. Since an element X of S_n^1 or \tilde{S}_n^1 satisfies $\text{Tr}_{1^c} X \in S^2$, we have

$$S_1 \geq S_{1,n+1} \geq S_{1,n} \geq S_2, \quad S_1 \geq \tilde{S}_{1,n+1} \geq \tilde{S}_{1,n} \geq S_2.$$

Also, [46] showed that for any non-separable state ρ , there exists an integer n such that ρ does not belong to $S_{1,n}$. The speed of the convergence is studied in [47, 48].

Therefore, we have

$$S_1 = \lim_{n \rightarrow \infty} S_{1,n} = \lim_{n \rightarrow \infty} \tilde{S}_{1,n}.$$

In addition, this discussion can be applied to the approximate calculation of J_1 by SDP

As an application of our theory, we considered the canonical problem of field sensing. We expect that our theory can lend insight to many other problems that can be phrased

naturally in the channel estimation framework, such as for other applications in quantum imaging and sensing.

ACKNOWLEDGEMENT

MH is supported in part by the National Natural Science Foundation of China (Grant No. 62171212). YO acknowledges support from EPSRC (Grant No. EP/W028115/1).

-
- [1] B. Escher, R. L. de Matos Filho, and L. Davidovich, General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology, *Nature Physics* **7**, 406 (2011).
 - [2] M. Hayashi, Comparison between the Cramer-Rao and the mini-max approaches in quantum channel estimation, *Commun. Math. Phys.* **304**, 689 (2011).
 - [3] S. Pirandola, R. Laurenza, C. Lupo, and J. L. Pereira, Fundamental limits to quantum channel discrimination, *npj Quantum Information* **5**, 50 (2019).
 - [4] S. Zhou and L. Jiang, Asymptotic theory of quantum channel estimation, *PRX Quantum* **2**, 010343 (2021).
 - [5] C. Helstrom, Minimum mean-squared error of estimates in quantum statistics, *Physics Letters A* **25**, 101 (1967).
 - [6] C. W. Helstrom, *Quantum detection and estimation theory* (Academic press, 1976).
 - [7] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory* (Edizioni della Normale, 2011).
 - [8] H. Nagaoka, A new approach to Cramér-Rao bounds for quantum state estimation, IEICE Tech Report **IT 89-42**, 9 (1989), (Reprinted in [10]).
 - [9] M. Hayashi and K. Matsumoto, Asymptotic performance of optimal state estimation in qubit system, *Journal of Mathematical Physics* **49**, 102101 (2008).
 - [10] M. Hayashi, ed., *Asymptotic theory of quantum statistical inference: selected papers* (World Scientific, 2005).
 - [11] M. Hayashi and K. Matsumoto, Asymptotic performance of optimal state estimation in qubit system, *Journal of Mathematical Physics* **49**, 102101 (2008).
 - [12] 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).
 - [13] J. S. Sidhu and P. Kok, Geometric perspective on quantum parameter estimation, *AVS Quantum Science* **2**, 014701 (2020).
 - [14] M. Hayashi and Y. Ouyang, Tight Cramér-Rao type bounds for multiparameter quantum metrology through conic programming, *Quantum* **7**, 1094 (2023).
 - [15] F. Albarelli, J. F. Friel, and A. Datta, Evaluating the Holevo Cramér-Rao bound for multiparameter quantum metrology, *Phys. Rev. Lett.* **123**, 200503 (2019).
 - [16] 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).
 - [17] M. Hayashi, A linear programming approach to attainable Cramér-Rao type bounds and randomness condition (1997), [arXiv:quant-ph/9704044](https://arxiv.org/abs/quant-ph/9704044) [quant-ph].
 - [18] 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) p. 133–149.
 - [19] M. Hayashi, On simultaneous measurement of noncommutative observables (in Japanese), *Surikaiseikikenkyusho (RIMS) Kokyuroku (Development of Infinite-Dimensional Noncommutative Analysis)*, 96 (1999).
 - [20] 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).
 - [21] Q. Liu, Z. Hu, H. Yuan, and Y. Yang, Optimal strategies of quantum metrology with a strict hierarchy, *Phys. Rev. Lett.* **130**, 070803 (2023).
 - [22] A. Altherr and Y. Yang, Quantum metrology for non-markovian processes, *Phys. Rev. Lett.* **127**, 060501 (2021).
 - [23] W. Górecki, S. Zhou, L. Jiang, and R. Demkowicz-Dobrzański, Optimal probes and error-correction schemes in multiparameter quantum metrology, *Quantum* **4**, 288 (2020).
 - [24] J. Friel, P. Palittapongarnpim, F. Albarelli, and A. Datta, Attainability of the Holevo-Cramér-Rao bound for two-qubit 3d magnetometry, [arXiv preprint arXiv:2008.01502](https://arxiv.org/abs/2008.01502) 10.48550/arXiv.2008.01502 (2020).
 - [25] A. Acín, Statistical distinguishability between unitary operations, *Phys. Rev. Lett.* **87**, 177901 (2001).
 - [26] M. F. Sacchi, Optimal discrimination of quantum operations, *Phys. Rev. A* **71**, 062340 (2005).
 - [27] M. Hayashi, Discrimination of two channels by adaptive methods and its application to quantum system, *IEEE Transactions on Information Theory* **55**, 3807 (2009).
 - [28] Q. Zhuang and S. Pirandola, Ultimate limits for multiple quantum channel discrimination, *Phys. Rev. Lett.* **125**, 080505 (2020).
 - [29] M. M. Wilde, M. Berta, C. Hirche, and E. Kaur, Amortized channel divergence for asymptotic quantum channel discrimination, *Letters in Mathematical Physics* **110**, 2277 (2020).
 - [30] K. Nakahira and K. Kato, Generalized quantum process discrimination problems, *Phys. Rev. A* **103**, 062606 (2021).
 - [31] P.-A. Moreau, E. Toninelli, T. Gregory, and M. J. Padgett, Imaging with quantum states of light, *Nature Reviews Physics* **1**, 367 (2019).
 - [32] G. Tóth and I. Apellaniz, Quantum metrology from a quantum information science perspective, *Journal of Physics A: Mathematical and Theoretical* **47**, 424006 (2014).
 - [33] L.-M. Duan and G.-C. Guo, Optimal quantum codes for preventing collective amplitude damping, *Phys. Rev. A* **58**, 3491 (1998).
 - [34] H. Nagaoka, On the parameter estimation problem for quantum statistical models, In *Proceeding of 12th Symposium on*

- Information Theory and Its Applications (SITA), Inuyama, Tottori, Japan, December 6–9, 577 (1989).
- [35] M.-D. Choi, Completely positive linear maps on complex matrices, *Linear algebra and its applications* **10**, 285 (1975).
- [36] A. Fujiwara and H. Imai, Quantum parameter estimation of a generalized pauli channel, *Journal of Physics A: Mathematical and General* **36**, 8093 (2003).
- [37] M. Hayashi, Private communication to A. Fujiwara (2003).
- [38] H. Imai and A. Fujiwara, Geometry of optimal estimation scheme for su(d) channels, *Journal of Physics A: Mathematical and Theoretical* **40**, 4391 (2007).
- [39] 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).
- [40] I. L. Chuang, D. W. Leung, and Y. Yamamoto, Bosonic quantum codes for amplitude damping, *Phys. Rev. A* **56**, 1114 (1997).
- [41] T. Baumgratz and A. Datta, Quantum enhanced estimation of a multidimensional field, *Phys. Rev. Lett.* **116**, 030801 (2016).
- [42] Y. Ouyang and G. K. Brennen, Quantum error correction on symmetric quantum sensors, arXiv preprint arXiv:2212.06285 [10.48550/arXiv.2212.06285](https://arxiv.org/abs/2212.06285) (2022).
- [43] Y. Ouyang, Permutation-invariant quantum codes, *Physical Review A* **90**, 062317 (2014), 1302.3247.
- [44] Y. Ouyang, Permutation-invariant quantum coding for quantum deletion channels, in *2021 IEEE International Symposium on Information Theory (ISIT)* (2021) pp. 1499–1503.
- [45] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Distinguishing separable and entangled states, *Phys. Rev. Lett.* **88**, 187904 (2002).
- [46] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Complete family of separability criteria, *Phys. Rev. A* **69**, 022308 (2004).
- [47] M. Navascués, M. Owari, and M. B. Plenio, Power of symmetric extensions for entanglement detection, *Phys. Rev. A* **80**, 052306 (2009).
- [48] H. Fawzi, The set of separable states has no finite semidefinite representation except in dimension 3×2 , *Communications in Mathematical Physics* **386**, 1319 (2021).
- [49] Y. Ouyang, N. Shettell, and D. Markham, Robust quantum metrology with explicit symmetric states, *IEEE Transactions on Information Theory* **68**, 1809 (2022).
- [50] R. S. Varga, *Geršgorin and his circles*, 1st ed. (Springer-Verlag, 2004).
- [51] J. Fan, W. Wang, and Y. Zhong, An ℓ_∞ eigenvector perturbation bound and its application to robust covariance estimation, *Journal of Machine Learning Research* **18**, 1 (2018).

Appendix A: Proofs of Lemmas 2, 3, and 4

1. Proof of Lemma 2

Using the conditions (26) and (27), for $k = 1, 2, 3, 4$, we can write

$$S_k[\text{Tr}_A(T \otimes I_C)(I_B \otimes \rho_{AC}), (\text{Tr}_A(F_j \otimes I_C)(I_B \otimes \rho_{AC}))_j] \\ = \min_{X \in \mathcal{S}_{BC}^k} \left\{ \text{Tr}(G \otimes T \otimes I_C)(I_{RB} \otimes \rho_{AC})(I_A \otimes X) \left| \begin{array}{l} I_{BC} = \text{Tr}_R X |0\rangle\langle 0| \otimes I_{BC} \\ \frac{1}{2} \text{Tr}(I_A \otimes X)((|0\rangle\langle j'| + |j'\rangle\langle 0|) \otimes F_j \otimes I_C)(I_{RB} \otimes \rho_{AC}) = \delta_{j,j'} \end{array} \right. \right\}.$$

For $k = 5$, we have

$$S_5[\text{Tr}_A(T \otimes I_C)(I_B \otimes \rho_{AC}), (\text{Tr}_A(F_j \otimes I_C)(I_B \otimes \rho_{AC}))_j] \\ = \min_{X \in \mathcal{S}_{BC}^5} \left\{ \text{Tr}(G \otimes T \otimes I_C)(I_{RB} \otimes \rho_{AC})(I_A \otimes X) \left| \begin{array}{l} I_{BC} = \text{Tr}_R X |0\rangle\langle 0| \otimes I_{BC} \\ \frac{1}{2} \text{Tr}(I_A \otimes X)((|0\rangle\langle j'| + |j'\rangle\langle 0|) \otimes F_j \otimes I_C)(I_{RB} \otimes \rho_{AC}) = \delta_{j,j'} \\ \text{Tr}(I_A \otimes X)((|i\rangle\langle j'| - |j'\rangle\langle i|) \otimes T \otimes I_C)(I_{RB} \otimes \rho_{AC}) = 0 \end{array} \right. \right\}.$$

For $k = 1, 2, 3, 4$, we have (25) as

$$\begin{aligned}
& \min_{\rho_{AC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_C)} S_k[\text{Tr}_A(T \otimes I_C)(I_B \otimes \rho), (\text{Tr}_A(F_j \otimes I_C)(I_B \otimes \rho))_j] \\
&= \min_{\rho_{AC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_C)} \min_{X \in \mathcal{S}_{BC}^k} \left\{ \text{Tr}(G \otimes T \otimes I_C)(I_{RB} \otimes \rho_{AC})(I_A \otimes X) \left| \begin{array}{l} I_{BC} = \text{Tr}_R X |0\rangle\langle 0| \otimes I_{BC} \\ \frac{1}{2} \text{Tr}(I_A \otimes X)((|0\rangle\langle j'| + |j'\rangle\langle 0|) \otimes F_j \otimes I_C)(I_{RB} \otimes \rho_{AC}) = \delta_{j,j'} \end{array} \right. \right\} \\
&= \min_{\rho_{AC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_C)} \min_{X \in \mathcal{S}_{BC}^k} \left\{ \text{Tr}(G \otimes T) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] \left| \begin{array}{l} I_{BC} = \text{Tr}_R X |0\rangle\langle 0| \otimes I_{BC} \\ \frac{1}{2} \text{Tr}((|0\rangle\langle j'| + |j'\rangle\langle 0|) \otimes F_j) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] = \delta_{j,j'} \end{array} \right. \right\} \quad (\text{A1})
\end{aligned}$$

$$\begin{aligned}
& \stackrel{(a)}{\geq} \min_{\rho_{AC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_C)} \min_{X \in \mathcal{S}_{BC}^k} \left\{ \text{Tr}(G \otimes T) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] \left| \begin{array}{l} \text{Tr}_R \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)](|0\rangle\langle 0| \otimes I_{AB}) = (\text{Tr}_C \rho_{AC}) \otimes I_B \\ \frac{1}{2} \text{Tr}((|0\rangle\langle j'| + |j'\rangle\langle 0|) \otimes F_j) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] = \delta_{j,j'} \end{array} \right. \right\} \quad (\text{A2})
\end{aligned}$$

$$\stackrel{(b)}{\geq} \min_{Y \in \mathcal{S}_{BA}^k} \{ \text{Tr} Y G \otimes T | (i), (ii) \text{ hold} \} = J_k. \quad (\text{A3})$$

Here, step (a) is shown as follows. The condition $I_{BC} = \text{Tr}_R X |0\rangle\langle 0| \otimes I_{BC}$ implies the condition

$$\begin{aligned}
& \text{Tr}_R \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)](|0\rangle\langle 0| \otimes I_{AB}) = \text{Tr}_{RC}[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)](|0\rangle\langle 0| \otimes I_{ABC}) \\
&= \text{Tr}_C(I_B \otimes \rho_{AC}) \text{Tr}_R[(I_A \otimes X)](|0\rangle\langle 0| \otimes I_{ABC}) \\
&= \text{Tr}_C(I_B \otimes \rho_{AC})(I_A \otimes \text{Tr}_R[X(|0\rangle\langle 0| \otimes I_{BC})]) = \text{Tr}_C(I_B \otimes \rho_{AC})(I_A \otimes I_{BC}) = (\text{Tr}_C \rho_{AC}) \otimes I_B.
\end{aligned}$$

Hence, a pair (ρ_{AC}, X) in (A1) satisfies the condition in (A2). Thus, for a pair (ρ_{AC}, X) in (A1), we have $\text{Tr}(G \otimes T) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] \geq (\text{A2})$, which implies (a).

Step (b) is shown as follows. For a pair (ρ_{AC}, X) in (A2), we choose Y to be $\text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)]$. Since

$$\text{Tr}_R \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)](|0\rangle\langle 0| \otimes I_{AB}) = (\text{Tr}_C \rho_{AC}) \otimes I_B, \quad (\text{A4})$$

$\rho_A(Y)$ is calculated to be $\text{Tr}_C \rho_{AC}$. Then, the condition (A4) implies the condition (i). The condition $\frac{1}{2} \text{Tr}((|0\rangle\langle j'| + |j'\rangle\langle 0|) \otimes F_j) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] = \delta_{j,j'}$ implies the condition (ii). Also, we have $\text{Tr}(G \otimes T) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] \geq (\text{A3})$. The remaining issue is to show $Y \in \mathcal{S}_{BA}^k$ for $k = 1, 2, 3, 4$. For $k = 4$, it is sufficient to show $Y \geq 0$ and $Y \in \mathcal{B}_{ABC}''$. Since $X \in \mathcal{B}_{BC}''$ and $X \geq 0$, we have $\sqrt{I_{RB} \otimes \rho_{AC}}(I_A \otimes X)\sqrt{I_{RB} \otimes \rho_{AC}} \geq 0$ belongs to \mathcal{B}_{ABC}'' . Taking the trace on C , we obtain $Y \geq 0$ and $Y \in \mathcal{B}_{ABC}''$. For other cases, we need to show additional conditions. For $k = 1$, we need to show that $Y \geq 0$ has a separable form with the bipartition R and AB . Since $\sqrt{I_{RB} \otimes \rho_{AC}}(I_A \otimes X)\sqrt{I_{RB} \otimes \rho_{AC}}$ has a separable form with the bipartition R and ABC , Y satisfies this condition. For $k = 2$, we need to show that $Y \in \mathcal{B}_{AB}$. Since $X \in \mathcal{B}_{BC}$ implies $\sqrt{I_{RB} \otimes \rho_{AC}}(I_A \otimes X)\sqrt{I_{RB} \otimes \rho_{AC}} \in \mathcal{B}_{ABC}$, Y satisfies this condition. For $k = 3$, we need to show the following. When we apply the partial transpose on the system R to Y , it is positive semi-definite. Since $\sqrt{I_{RB} \otimes \rho_{AC}}(I_A \otimes X)\sqrt{I_{RB} \otimes \rho_{AC}}$ satisfies this property, Y satisfies this condition. Therefore, we obtain (b).

For $k = 5$, we have (25) as

$$\begin{aligned}
& \min_{\rho_{AC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_C)} S_5[\text{Tr}_A(T \otimes I_C)(I_B \otimes \rho), (\text{Tr}_A(F_j \otimes I_C)(I_B \otimes \rho))_j] \\
&= \min_{\rho_{AC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_C)} \min_{X \in \mathcal{S}_{BC}^4} \left\{ \text{Tr}(G \otimes T \otimes I_C)(I_{RB} \otimes \rho_{AC})(I_A \otimes X) \left| \begin{array}{l} I_{BC} = \text{Tr}_R X |0\rangle\langle 0| \otimes I_{BC} \\ \frac{1}{2} \text{Tr}(I_A \otimes X)((|0\rangle\langle j'| + |j'\rangle\langle 0|) \otimes F_j \otimes I_C)(I_{RB} \otimes \rho_{AC}) = \delta_{j,j'} \\ \text{Tr}(I_A \otimes X)((|i\rangle\langle j'| - |j'\rangle\langle i|) \otimes T \otimes I_C)(I_{RB} \otimes \rho_{AC}) = 0 \end{array} \right. \right\} \\
&= \min_{\rho_{AC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_C)} \min_{X \in \mathcal{S}_{BC}^4} \left\{ \text{Tr}(G \otimes T) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] \left| \begin{array}{l} I_{BC} = \text{Tr}_R X |0\rangle\langle 0| \otimes I_{BC} \\ \frac{1}{2} \text{Tr}((|0\rangle\langle j'| + |j'\rangle\langle 0|) \otimes F_j) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] = \delta_{j,j'} \\ \text{Tr}((|i\rangle\langle j'| - |j'\rangle\langle i|) \otimes T) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] = 0 \end{array} \right. \right\} \quad (\text{A5})
\end{aligned}$$

$$\stackrel{(a)}{\geq} \min_{\rho_{AC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_C)} \min_{X \in \mathcal{S}_{BC}^4} \left\{ \text{Tr}(G \otimes T) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] \left| \begin{array}{l} \text{Tr}_R \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)](|0\rangle\langle 0| \otimes I_B) = (\text{Tr}_C \rho_{AC}) \otimes I_B \\ \frac{1}{2} \text{Tr}((|0\rangle\langle j'| + |j'\rangle\langle 0|) \otimes F_j) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] = \delta_{j,j'} \\ \text{Tr}((|i\rangle\langle j'| - |j'\rangle\langle i|) \otimes T) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] = 0 \end{array} \right. \right\}$$

$$\stackrel{(b)}{\geq} \min_{Y \in \mathcal{S}_{BA}^5(T)} \{ \text{Tr} Y G \otimes T | (i), (ii) \text{ hold} \} = J_5. \quad (\text{A6})$$

Here, steps (a) and (b) can be shown in the same way as (A3).

2. Proof of Lemma 4

Since the proof of Lemma 4 is easier than the proof of Lemma 3, we show Lemma 4 here. For $k = 1, 2, 3, 4$, given $Y \in \mathcal{S}_{BA}^k$ satisfying the conditions (i), (ii) and $\epsilon > 0$, we define $Y'_\epsilon \in \mathcal{S}_{BA}^k$ as

$$Y'_\epsilon := ((\sqrt{1-\epsilon}|0\rangle\langle 0| + \sqrt{1-\epsilon}^{-1}(I_R - |0\rangle\langle 0|)) \otimes I_{AB})Y((\sqrt{1-\epsilon}|0\rangle\langle 0| + \sqrt{1-\epsilon}^{-1}(I_R - |0\rangle\langle 0|)) \otimes I_{AB}). \quad (\text{A7})$$

Then, we define $Y_\epsilon \in \mathcal{S}_{BA}^k$ as

$$Y_\epsilon := Y'_\epsilon + \frac{\epsilon}{d_A}|0\rangle\langle 0| \otimes I_{AB}.$$

Here, we relax the condition (i) as follows.

(i''): There exists a positive semi-definite operator $\rho_A(Y)$ on \mathcal{H}_A such that the relation

$$\text{Tr}_R Y(|0\rangle\langle 0| \otimes I_{AB}) = I_B \otimes \rho_A(Y) \quad (\text{A8})$$

holds as operators on $\mathcal{H}_A \otimes \mathcal{H}_B$. Here, we drop the condition $\text{Tr} \rho_A(Y) = 1$.

Since Y'_ϵ and $\frac{1}{d_A}|0\rangle\langle 0| \otimes I_{AB}$ satisfy the condition (i''), Y_ϵ also satisfies the condition (i''). We can define $\rho_A(Y'_\epsilon)$, $\rho_A(\frac{1}{d_A}|0\rangle\langle 0| \otimes I_{AB})$, and $\rho_A(Y_\epsilon)$. Since we have $\rho_A(Y'_\epsilon) = (1-\epsilon)\rho_A(Y)$ and $\rho_A(\frac{1}{d_A}|0\rangle\langle 0| \otimes I_{AB}) = \frac{1}{d_A}I_A$, we have $\rho_A(Y_\epsilon) = (1-\epsilon)\rho_A(Y) + \frac{\epsilon}{d_A}I_A > 0$.

Also, due to the definition (A7), for $j > 0$, $|0\rangle\langle j|$ component of Y'_ϵ and $|j\rangle\langle 0|$ component of Y'_ϵ are the same as $|0\rangle\langle j|$ component of Y and $|j\rangle\langle 0|$ component of Y , respectively. Thus, $|0\rangle\langle j|$ component of Y_ϵ and $|j\rangle\langle 0|$ component of Y_ϵ are the same as $|0\rangle\langle j|$ component of Y and $|j\rangle\langle 0|$ component of Y , respectively.

Therefore, Y_ϵ satisfies (i), (ii) and, $\rho_A(Y_\epsilon) > 0$. Also, we have $\lim_{\epsilon \rightarrow 0} \text{Tr} Y_\epsilon(G \otimes T) = \text{Tr} Y(G \otimes T)$. Therefore, we obtain (32).

For $k = 5$, in the same way as (32), we can show (33).

3. Proof of Lemma 3

For $k = 1, 2, 3, 4$, we choose $Y \in \mathcal{S}_{BA}^k$ satisfying the conditions (i), (ii), and $\rho_A := \rho_A(Y) > 0$. We diagonalize ρ_A as $\sum_{j=1}^{d_A} s_j |\phi_j\rangle\langle \phi_j|$. Hence, we have $s_j > 0$.

We choose a CONS (complete orthonormal system) $\{\psi_j\}$ of \mathcal{H}_C . We define a unitary map $U : \phi_j \mapsto \psi_j$ from \mathcal{H}_A to \mathcal{H}_C . We define the matrix $\rho_A^{-1/2} := \sum_{j=1}^{d_A} s_j^{-1/2} |\phi_j\rangle\langle \phi_j|$. In a way similar to the discussion after (A4), we can show that $X := (U \otimes I_{RB})(\rho_A^{-1/2} \otimes I_{RB})Y(\rho_A^{-1/2} \otimes I_{RB})(U^\dagger \otimes I_{RB})$ belongs to \mathcal{S}_{BC}^k for $k = 1, 2, 3, 4$. Notice that both sides in this definition act on $\mathcal{H}_R \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ because U maps \mathcal{H}_A to \mathcal{H}_C .

We choose ρ_{AC} as the pure state $\sum_{j=1}^{d_A} \sqrt{s_j} |\phi_j, \psi_j\rangle$, which is a purification of ρ_A .

Then,

$$\begin{aligned} \text{Tr}_R X(|0\rangle\langle 0| \otimes I_{BC}) &= \text{Tr}_R (U \otimes I_{RB})(\rho_A^{-1/2} \otimes I_{RB})Y(\rho_A^{-1/2} \otimes I_{RB})(U^\dagger \otimes I_{RB})|0\rangle\langle 0| \otimes I_{BC} \\ &= (U \otimes I_B)(\rho_A^{-1/2} \otimes I_B)\text{Tr}_R[Y|0\rangle\langle 0| \otimes I_{BA}](\rho_A^{-1/2} \otimes I_B)(U^\dagger \otimes I_B) \\ &= (U \otimes I_B)(\rho_A^{-1/2} \otimes I_B)I_B \otimes \rho_A(\rho_A^{-1/2} \otimes I_B)(U^\dagger \otimes I_B) \\ &= (U \otimes I_B)I_{BA}(U^\dagger \otimes I_B) = I_{BC}. \end{aligned} \quad (\text{A9})$$

For a matrix Z on \mathcal{H}_A , we have

$$\text{Tr}_C[\rho_{AC}(I_A \otimes (U\rho_A^{-1/2}Z(\rho_A^{-1/2}U^\dagger)))] = Z. \quad (\text{A10})$$

This can be shown as follows.

$$\begin{aligned} &\text{Tr}_C[\rho_{AC}(I_A \otimes (U\rho_A^{-1/2}Z(\rho_A^{-1/2}U^\dagger)))] \\ &= \text{Tr}_C\left[\sum_{i=1}^{d_A} \sqrt{s_i} |\phi_i, \psi_i\rangle \sum_{i'=1}^{d_A} \sqrt{s_{i'}} \langle \phi_{i'}, \psi_{i'}| \left(I_A \otimes \left(\sum_{j=1}^{d_A} s_j^{-1/2} |\psi_j\rangle\langle \phi_j| Z \sum_{j'=1}^{d_A} s_{j'}^{-1/2} |\psi_{j'}\rangle\langle \phi_{j'}| \right) \right)\right] \\ &= \left(\sum_{j=1}^{d_A} |\phi_j\rangle\langle \phi_j| Z \sum_{j'=1}^{d_A} |\phi_{j'}\rangle\langle \phi_{j'}| \right) = Z. \end{aligned} \quad (\text{A11})$$

We write the matrix Y as $\sum_{j,j',b,b'} |j, b\rangle\langle j', b'| \otimes Y_{j,b,j',b'}$ by using matrices $Y_{j,b,j',b'}$ on \mathcal{H}_A . Here, $\{|b\rangle\}$ is a CONS of \mathcal{H}_B . Applying (A10) to the matrix components $Y_{j,b,j',b'}$, we have

$$\begin{aligned}
& \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] \\
& \stackrel{(a)}{=} \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes ((U \otimes I_{RB})(\rho_A^{-1/2} \otimes I_{RB})Y(\rho_A^{-1/2} \otimes I_{RB})(U^\dagger \otimes I_{RB})))]) \\
& = \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes ((U \otimes I_{RB})(\rho_A^{-1/2} \otimes I_{RB})(\sum_{j,j',b,b'} |j, b\rangle\langle j', b'| \otimes Y_{j,b,j',b'}) (\rho_A^{-1/2} \otimes I_{RB})(U^\dagger \otimes I_{RB})))]) \\
& = \sum_{j,j',b,b'} |j, b\rangle\langle j', b'| \otimes \text{Tr}_C[\rho_{AC}(I_A \otimes (U \rho_A^{-1/2} Y_{j,b,j',b'} \rho_A^{-1/2} U^\dagger))] \\
& \stackrel{(b)}{=} \sum_{j,j',b,b'} |j, b\rangle\langle j', b'| \otimes Y_{j,b,j',b'} = Y,
\end{aligned} \tag{A12}$$

where (a) follows from the definition of X , and (b) follows from (A10).

Thus, we have

$$\frac{1}{2} \text{Tr}(|0\rangle\langle j'| + |j'\rangle\langle 0| \otimes F_j) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] \stackrel{(a)}{=} \frac{1}{2} \text{Tr}(|0\rangle\langle j'| + |j'\rangle\langle 0| \otimes F_j) Y \stackrel{(b)}{=} \delta_{j,j'}, \tag{A13}$$

where (a) follows from (A12), and (b) follows from the fact that Y satisfies the condition (ii).

Therefore, the relations (A9), (A12), and (A13) guarantee that $X \in \mathcal{S}_{BC}^k$ satisfies the conditions (26), (27), and (30), respectively. Thus, for $k = 1, 2, 3, 4$, we have

$$\begin{aligned}
& \text{Tr}Y(G \otimes T) \\
& \stackrel{(a)}{\geq} \min_{\rho_{AC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_C)} \min_{X \in \mathcal{S}_{BC}^k} \left\{ \text{Tr}(G \otimes T) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] \left| \begin{array}{l} \text{Tr}_R X(|0\rangle\langle 0| \otimes I_{BC}) = I_{BC} \\ \frac{1}{2} \text{Tr}(|0\rangle\langle j'| + |j'\rangle\langle 0| \otimes F_j \otimes I_C) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] = \delta_{j,j'} \end{array} \right. \right\} \\
& \stackrel{(b)}{=} \min_{\rho_{AC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_C)} S_k[\text{Tr}_A(T \otimes I_C)(I_B \otimes \rho), (\text{Tr}_A(F_j \otimes I_C)(I_B \otimes \rho))_j],
\end{aligned} \tag{A14}$$

where (a) follows from the fact that $X \in \mathcal{S}_{BC}^k$ satisfies the conditions (26), (27), and (30), and (b) follows from (A1). Therefore, we obtain (28).

For $k = 5$, we choose $Y \in \mathcal{S}_{BA}^5(T)$ satisfying the conditions (i), (ii) and $\rho_A(Y) > 0$. We choose ρ and X in the same way as the above. Since $Y \in \mathcal{S}_{BA}^5(T)$ satisfies the conditions (i), (ii), the relations (A9), (A13) hold. Since $Y \in \mathcal{S}_{BA}^5(T)$, in the same way as (A13), we have

$$\text{Tr}(|i\rangle\langle j'| - |j'\rangle\langle i| \otimes T \otimes I_C) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] = \text{Tr}(|i\rangle\langle j'| - |j'\rangle\langle i| \otimes T) Y = 0. \tag{A15}$$

Therefore, the relations (A9), (A12), (A13), and (A15) guarantee that $X \in \mathcal{S}_{BC}^4$ satisfies the conditions (26), (27), (31), and (30), respectively. Thus, we have

$$\begin{aligned}
& \text{Tr}Y(G \otimes T) \\
& \stackrel{(a)}{\geq} \min_{\rho_{AC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_C)} \min_{X \in \mathcal{S}_{BC}^4} \left\{ \text{Tr}(G \otimes T) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] \left| \begin{array}{l} \text{Tr}_R X(|0\rangle\langle 0| \otimes I_{BC}) = I_{BC} \\ \frac{1}{2} \text{Tr}(|0\rangle\langle j'| + |j'\rangle\langle 0| \otimes F_j \otimes I_C) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] = \delta_{j,j'} \\ \text{Tr}(|i\rangle\langle j'| - |j'\rangle\langle i| \otimes T) \text{Tr}_C[(I_{RB} \otimes \rho_{AC})(I_A \otimes X)] = 0. \end{array} \right. \right\} \\
& \stackrel{(b)}{=} \min_{\rho_{AC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_C)} S_5[\text{Tr}_A(T \otimes I_C)(I_B \otimes \rho), (\text{Tr}_A(F_j \otimes I_C)(I_B \otimes \rho))_j],
\end{aligned}$$

where (a) follows from the fact that $X \in \mathcal{S}_{BC}^k$ satisfies the conditions (26), (27), (31), and (30), and (b) follows from (A5). Therefore, we obtain (28).

Appendix B: Proofs of Theorems 5 and 6

1. Proof of Theorem 5

We consider the dual problem for the conic programming (39) and (40). For $(A, B, W) \in \mathbb{R}^{d \times d} \times \mathcal{T}_{sa}(\mathcal{H}_{AB})^d \times \mathcal{T}_{sa}(\mathcal{H}_{AB})$, we define

$$\begin{aligned} \Pi(A, B, W) \\ := G \otimes T_N - \frac{1}{2} \left(\sum_{1 \leq i, j \leq d} A_i^j (|0\rangle\langle i| + |i\rangle\langle 0|) \otimes F_{j,N} \right) + \sqrt{-1} \sum_{i=1}^d (|0\rangle\langle i| - |i\rangle\langle 0|) \otimes B_i - |0\rangle\langle 0| \otimes W. \end{aligned}$$

We have

$$J_4 = \max_{(A, B, W) \in \mathbb{R}^{d \times d} \times \mathcal{T}_{sa}(\mathcal{H}_{AB})^d \times (\mathcal{T}_{sa}(\mathcal{H}_{AB}) \cap \mathcal{K})} \left\{ \text{Tr} W + \sum_{j=1}^d A_j^j \mid \Pi(A, B, W) \geq 0 \right\}.$$

Here, the matrix $A \in \mathbb{R}^{d \times d}$ corresponds to the condition (ii-N) in the primal problem. The matrices $(B^1, \dots, B^d) \in \mathcal{T}_{sa}(\mathcal{H}_{AB})^d$ correspond to the condition $X_{k,0} \in \mathcal{B}_{sa}(\mathcal{H}_{A,B})$ for $k = 1, \dots, d$ in the primal problem. This condition is rewritten as $X_{k,0}^\dagger = X_{k,0}$.

The condition (i'-N) is composed of (36), (37), and (38). We denote the variable corresponding to the condition (36) with $b \in \{1, \dots, d_B - 1\}$ and $b' \in \{2, \dots, d_B - 1\}$ with $b > b'$ by the matrix $W_{b,b'}$ on \mathcal{H}_A . We denote the variable corresponding to the condition (37) with $b \in \{1, \dots, d_B - 1\}$ by the Hermitian matrix W_b . We denote the variable corresponding to the condition (38) by the real number w . Then, we choose the Hermitian matrix W on $\mathcal{H}_A \otimes \mathcal{H}_B$ as

$$W = \sum_{b > b'} W_{b,b'} \otimes |b\rangle\langle b'| + W_{b,b'}^\dagger \otimes |b'\rangle\langle b| + \sum_{b=1}^{d_B-1} W_b \otimes (|b\rangle\langle b| - |b+1\rangle\langle b+1|) + wI \otimes |1\rangle\langle 1|.$$

The above Hermitian matrix on $\mathcal{H}_A \otimes \mathcal{H}_B$ belongs to the subset \mathcal{K} , and any element of \mathcal{K} can be written as the above form. Thus, The matrix $W \in (\mathcal{T}_{sa}(\mathcal{H}_{AB}) \cap \mathcal{K})$ corresponds to the condition (i'-N) in the primal problem.

On the other hand, we have

$$\begin{aligned} S_4[T_N, (F_{j,N})_j] \\ = \max_{(A, B, W) \in \mathbb{R}^{d \times d} \times \mathcal{T}_{sa}(\mathcal{H}_{AB})^d \times \mathcal{T}_{sa}(\mathcal{H}_{AB})} \left\{ \text{Tr} W + \sum_{j=1}^d A_j^j \mid \Pi(A, B, W) \geq 0 \right\}. \end{aligned} \quad (\text{B1})$$

In S_4 , the condition for W is relaxed as $W \in (\mathcal{T}_{sa}(\mathcal{H}_{AB}) \cap \mathcal{K})$ because the condition (i'-N) is changed to $\text{Tr}_R[Y_N(|0\rangle\langle 0| \otimes I_{AB})] = I_{AB}$.

Therefore, if and only if the maximizer (A_*, B_*, W_*) in (B1) satisfies the condition $W_* \in \mathcal{K}$, we have

$$S_4[T_N, (F_{j,N})_j] = J_4. \quad (\text{B2})$$

Lemma 15. (1) The minimization (41) is attained when $Y_N^{i,j} = X_*^{i,j} := L_*^i L_*^j$, where L_*^0 is defined to be I_{AB} . (2) The maximization (B1) is attained when $A = A_* := -GJ_{\text{SLD}}^{-1}$, $B_j = B_{*,j} := -\frac{\sqrt{-1}}{2} \sum_{i=1}^d (GJ_{\text{SLD}}^{-1})_j^i [L_{*,i}, T_N]$ and W is $-W_{\text{SLD}}(T_N, \vec{L}_*)$.

The combination of (B2) and Lemma 15 implies Theorem 5.

Proof. We diagonalize the matrix G as $G_{i,j} = \sum_{j'=1}^d g_{j'} v_{j',i} v_{j',j}$, where $v_{j',j}$ is an orthogonal matrix. Then, we have

$$\begin{aligned}
& \Pi(A_*, \vec{B}_*, -W_{\text{SLD}}(\vec{L}_*)) \\
&= \sum_{1 \leq i,j \leq d} G_{i,j} |i\rangle\langle j| \otimes T_N - \sum_{j=1}^d G_{i,j} (|0\rangle\langle i| \otimes L_*^j T_N + |i\rangle\langle 0| \otimes T_N L_*^j) + \sum_{1 \leq i,j \leq d} G_{i,j} |i\rangle\langle j| \otimes L_*^i T_N L_*^j \\
&= \sum_{1 \leq i,j \leq d} G_{i,j} |i\rangle\langle j| \otimes T_N - \sum_{j=1}^d G_{i,j} (|0\rangle\langle j| \otimes L_*^i T_N + |i\rangle\langle 0| \otimes T_N L_*^j) + \sum_{1 \leq i,j \leq d} G_{i,j} |i\rangle\langle j| \otimes L_*^i T_N L_*^j \\
&= \sum_{1 \leq i,j \leq d} G_{i,j} \left((|0\rangle\langle 0| \otimes L_*^i - |i\rangle\langle 0| \otimes I_{AB}) (|0\rangle\langle 0| \otimes T_N) (|0\rangle\langle 0| \otimes L_*^j - |j\rangle\langle 0| \otimes I_{AB}) \right)^\dagger \\
&= \sum_{j'=1}^d g_{j'} \left(\sum_{i=1}^d v_{j',i} \left((|0\rangle\langle 0| \otimes L_*^i - |i\rangle\langle 0| \otimes I_{AB}) \right) (|0\rangle\langle 0| \otimes T_N) \left(\sum_{j=1}^d v_{j',j} \left((|0\rangle\langle 0| \otimes L_*^j - |j\rangle\langle 0| \otimes I_{AB}) \right) \right)^\dagger \right) \\
&\geq 0.
\end{aligned} \tag{B3}$$

Hence, the tuple $(A_*, \vec{B}_*, -W_{\text{SLD}}(\vec{L}_*))$ belongs to the range of the maximization given in the RHS of (B1).

Since we have

$$\begin{aligned}
& \text{Tr} X_* \Pi(A_*, \vec{B}_*, -W_{\text{SLD}}(\vec{L}_*)) \\
&= \text{Tr} \left[\left(|0\rangle\langle 0| \otimes I_{AB} + \sum_{j'=1}^d |j'\rangle\langle 0| \otimes L_*^{j'} \right) \left(|0\rangle\langle 0| \otimes I_{AB} + \sum_{j'=1}^d |j'\rangle\langle 0| \otimes L_*^{j'} \right)^\dagger \right. \\
&\quad \left. \sum_{1 \leq i,j \leq d} G_{i,j} \left((|0\rangle\langle 0| \otimes L_*^i - |i\rangle\langle 0| \otimes I_{AB}) (|0\rangle\langle 0| \otimes T_N) (|0\rangle\langle 0| \otimes L_*^j - |j\rangle\langle 0| \otimes I_{AB}) \right)^\dagger \right] \\
&= \sum_{1 \leq i,j \leq d} G_{i,j} \text{Tr} \left[\left(|0\rangle\langle 0| \otimes I_{AB} + \sum_{j'=1}^d |0\rangle\langle j'| \otimes L_*^{j'} \right) \left(|0\rangle\langle 0| \otimes L_*^i - |i\rangle\langle 0| \otimes I_{AB} \right) (|0\rangle\langle 0| \otimes T_N) \right. \\
&\quad \left. \left((|0\rangle\langle 0| \otimes L_*^j - |j\rangle\langle 0| \otimes I_{AB}) \right)^\dagger \left(|0\rangle\langle 0| \otimes I_{AB} + \sum_{j'=1}^d |j'\rangle\langle 0| \otimes L_*^j \right) \right] \\
&= \sum_{1 \leq i,j \leq d} G_{i,j} \text{Tr} \left((|0\rangle\langle 0| \otimes L_*^i - |0\rangle\langle 0| \otimes L_*^i) (|0\rangle\langle 0| \otimes T_N) (|0\rangle\langle 0| \otimes L_*^j - |0\rangle\langle 0| \otimes L_*^j) \right)^\dagger \\
&= 0,
\end{aligned} \tag{B4}$$

we have

$$\text{Tr}[X_*(G \otimes T_N)] = \text{Tr} W_{\text{SLD}}(\vec{L}_*) + \sum_{j=1}^d A_{*,j}^j.$$

Since X_* belongs to $\in \mathcal{S}_{BA}^4$, we have

$$\tilde{S}_4[|\Phi\rangle\langle\Phi|] = S_4[T_N, (F_{j,N})_j] = \text{Tr}[X_*(G \otimes T_N)] = \text{Tr} W_{\text{SLD}}(\vec{L}_*) + \sum_{j=1}^d A_{*,j}^j.$$

Therefore, we obtain both desired statements. \square

2. Proof of Theorem 6

We denote the set of $d \times d$ anti-symmetric matrices by $\mathbb{R}_{AS}^{d \times d}$. For $(A, C, B, W) \in \mathbb{R}^{d \times d} \times \mathbb{R}_{AS}^{d \times d} \times \mathcal{T}_{sa}(\mathcal{H}_{AB})^d \times \mathcal{T}_{sa}(\mathcal{H}_{AB})$, we define

$$\Pi(A, C, B, W) := \Pi(A, B, W) - \sqrt{-1} C \otimes T_N.$$

Also, we have

$$J_5 = \max_{(A,C,B,W) \in \mathbb{R}^{d \times d} \times \mathbb{R}_{AS}^{d \times d} \times \mathcal{T}_{sa}(\mathcal{H}_{AB})^d \times (\mathcal{T}_{sa}(\mathcal{H}_{AB}) \cap \mathcal{K})} \left\{ \text{Tr} W + \sum_{j=1}^d A_j^j \middle| \Pi(A, C, B, W) \geq 0 \right\}.$$

Here, the matrix $A \in \mathbb{R}^{d \times d}$, the matrices $(B^1, \dots, B^d) \in \mathcal{T}_{sa}(\mathcal{H}_{AB})^d$, and the matrix $W \in (\mathcal{T}_{sa}(\mathcal{H}_{AB}) \cap \mathcal{K})$ have the same meaning as in J_4 . The matrix $C \in \mathbb{R}_{AS}^{d \times d}$ corresponds to the condition (8) in the primal problem.

Also, we have

$$\begin{aligned} & S_5[T_N, (F_{j,N})_j] \\ &= \max_{(A,C,B,W) \in \mathbb{R}^{d \times d} \times \mathbb{R}_{AS}^{d \times d} \times \mathcal{T}_{sa}(\mathcal{H}_{AB})^d \times \mathcal{T}_{sa}(\mathcal{H}_{AB})} \left\{ \text{Tr} W + \sum_{j=1}^d A_j^j \middle| \Pi(A, C, B, W) \geq 0 \right\}. \end{aligned} \quad (\text{B5})$$

Then, if and only if the maximizer (A_*, C_*, B_*, W_*) in (B5) satisfies the condition $W_* \in \mathcal{K}$, we have

$$S_5[T_N, (F_{j,N})_j] = J_5. \quad (\text{B6})$$

Lemma 16. (1) The minimization (42) is attained when $X = X_* := \Pi(\vec{Z}_*) + (|\text{Im } V_*| - (\sqrt{-1} \text{Im } V_*) \otimes I_{AB})$.

(2) $(Z_*^i - \sqrt{-1} \sum_{j=1}^d C_*^{i,j} Z_*^j) \circ T_N$ is written as a linear sum of $F_{i',N}$ for $i = 1, \dots, d$. That is, there exist real numbers $A_{*,i}^{i'}$ such that $(Z_*^i - \sqrt{-1} \sum_{j=1}^d C_*^{i,j} Z_*^j) \circ T_N = \sum_{i'=1}^d A_{*,i}^{i'} F_{i',N}$.

(3) The maximization (B5) is attained when $A = -A_*$, $B_j = B_{*,j} := -\frac{\sqrt{-1}}{2} \sum_{i=1}^d (A_*)_{j,i}^i [Z_{*,i}, T_N]$, $C = C_*$, and $W = -W_{HN}(T_N, \vec{Z}_*)$.

The combination of (B6) and Lemma 16 implies Theorem 6.

Proof. (1) is known. We write the maximizer in (B5) by (A_o, C_o, B_o, W_o) , whose existence is guaranteed because of the finiteness of the dimension. Then, we have

$$\begin{aligned} 0 &= \text{Tr}(\Pi(\vec{Z}_*) + (|\text{Im } V_*| - (\sqrt{-1} \text{Im } V_*) \otimes I_{AB}) \Pi(A_o, C_o, B_o, W_o)) \\ &= \text{Tr} \Pi(\vec{Z}_*) \Pi(A_o, C_o, B_o, W_o) + \text{Tr}(|\text{Im } V_*| - (\sqrt{-1} \text{Im } V_*) \otimes I_{AB} \Pi(A_o, C_o, B_o, W_o)) \\ &\geq \text{Tr}(|\text{Im } V_*| - (\sqrt{-1} \text{Im } V_*) \otimes I_{AB} \Pi(A_o, C_o, B_o, W_o)) \geq 0. \end{aligned}$$

Hence, we have

$$\begin{aligned} 0 &= \text{Tr}(|\text{Im } V_*| - (\sqrt{-1} \text{Im } V_*) \otimes I_{AB} \Pi(A_o, C_o, B_o, W_o)) \\ &= \text{Tr}(|\text{Im } V_*| - (\sqrt{-1} \text{Im } V_*) \otimes I_{AB} (I_R - \sqrt{-1} C_o) \otimes T_N = \text{Tr}(|\text{Im } V_*| - (\sqrt{-1} \text{Im } V_*) (I_R - \sqrt{-1} C_o)) \\ &= \text{Tr}|\text{Im } V_*| - \text{Tr}(\text{Im } V_*) C_o. \end{aligned}$$

Hence, we have $C_o = C_*$.

Therefore, we have

$$0 = \text{Tr} \Pi(\vec{Z}_*) \Pi(A_o, C_*, B_o, W_o),$$

which implies

$$0 = \Pi(\vec{Z}_*) \Pi(A_o, C_*, B_o, W_o) = \left(\sum_{i=0}^d |i\rangle \otimes (Z_*^i)^\dagger \right) \left(\sum_{i=0}^d \langle i| \otimes Z_*^i \right) \Pi(A_o, C_*, B_o, W_o).$$

Hence, we have

$$0 = \left(\sum_{i=1}^d \langle i| \otimes Z^i \right) \Pi(A_o, C_*, B_o, W_o).$$

We choose the Hermitian matrix $R_i \in \mathcal{H}_{AB}$ as $\sum_{j=1}^d A_{*,i}^j D_j + \sqrt{-1} B_{o,i}$. Hence, $R_i + R_i^\dagger$ can be written as a linear sum of D_1, \dots, D_d . Hence,

$$\Pi(A_o, C_*, B_o, W_o) = \sum_{1 \leq i, j \leq d} (\delta^{i,j} - \sqrt{-1} C_*^{i,j}) |i\rangle \langle j| \otimes T_N + \sum_{j=1}^d (|0\rangle \langle j| \otimes R_j + |j\rangle \langle 0| \otimes R_j^\dagger) - |0\rangle \langle 0| \otimes W_o.$$

$$\Pi(\vec{Z}_*) = \sum_{1 \leq i, j \leq d} |i\rangle\langle j|$$

Hence,

$$\begin{aligned} 0 &= \left(\sum_{i=1}^d \langle i| \otimes Z_*^i \right) \Pi(A_o, C_*, B_o, W_o) \\ &= \sum_{i,j} (\delta^{i,j} - \sqrt{-1} C_*^{i,j}) \langle j| \otimes Z_*^i T_N + \sum_{j=1}^d \langle j| \otimes R_j + \sum_{j=1}^d \langle 0| \otimes Z_*^j R_j^\dagger - \langle 0| \otimes W_o. \end{aligned} \quad (\text{B7})$$

Thus,

$$R_j = - \sum_{i=1}^d (\delta^{i,j} - \sqrt{-1} C_*^{i,j}) \otimes Z_*^i T_N \quad (\text{B8})$$

$$W_o = \sum_{j=1}^d Z_*^j R_j^\dagger. \quad (\text{B9})$$

Thus, $\frac{1}{2}(R_i + R_i^\dagger) = -\sum_{i=1}^d (\delta^{i,j} - \sqrt{-1} C_*^{i,j}) \otimes Z_*^i \circ T_N$ can be written as a linear sum of D_1, \dots, D_d . Hence, the statement (2) is shown. Hence, we have $A = -A_*$ and $B_j = B_{*,j} := -\frac{\sqrt{-1}}{2} \sum_{i=1}^d (A_*)_{j,i}^i [Z_{*,i}, T_N]$. Combining (B8) and (B9), we obtain $W_o = -W_{HN}(T_N, \vec{Z}_*)$. \square

Appendix C: Calculations for the field sensing problem

1. Restriction to the symmetric subspace

Because our numerical calculation can be restricted to the symmetric subspace, we calculate the projections of operators that we need to implement on the symmetric subspace. We begin by considering the action of E^1, E^2, E^3 on the symmetric subspace.

Now let us denote $\Pi^{\text{sym}} := \sum_{u=0}^n |D_u^n\rangle\langle D_u^n|$. We now show the following lemma.

Lemma 17. Let $r_w := \frac{1}{2} \sqrt{(n-w)(w+1)}$. Then

$$\Pi^{\text{sym}} E^3 \Pi^{\text{sym}} = \sum_{w=0}^n \left(\frac{n}{2} - w \right) |D_w^n\rangle\langle D_w^n|, \quad (\text{C1})$$

$$\Pi^{\text{sym}} E^1 \Pi^{\text{sym}} = \sum_{w=0}^{n-1} r_w (|D_w^n\rangle\langle D_{w+1}^n| + |D_{w+1}^n\rangle\langle D_w^n|), \quad (\text{C2})$$

$$\Pi^{\text{sym}} E^2 \Pi^{\text{sym}} = \sum_{w=0}^{n-1} r_w (-i |D_w^n\rangle\langle D_{w+1}^n| + i |D_{w+1}^n\rangle\langle D_w^n|). \quad (\text{C3})$$

Proof of Lemma 17. From [49, Lemma 6] we know that

$$\begin{aligned} \langle D_w^n | \tau_3^{(1)} | D_w^n \rangle &= \frac{1}{\binom{n}{w}} K_0^0(0) K_w^n(1), \\ \langle D_{w+1}^n | \tau_1^{(1)} | D_w^n \rangle &= \frac{1}{\sqrt{\binom{n}{w} \binom{n}{w+1}}} K_0^1(1) K_w^{n-1}(0), \\ \langle D_{w+1}^n | \tau_2^{(1)} | D_w^n \rangle &= \frac{i}{\sqrt{\binom{n}{w} \binom{n}{w+1}}} K_1^1(0) K_w^{n-1}(0), \end{aligned}$$

where

$$K_k^n(z) = \sum_{j=0}^z \binom{z}{j} \binom{n-z}{k-j} (-1)^j$$

is a Krawtchouk polynomial. Making substitutions for the values of the Krawtchouk polynomials, we get

$$\begin{aligned} \langle D_w^n | \tau_3^{(1)} | D_w^n \rangle &= 1 - \frac{2w}{n}, \\ \langle D_{w+1}^n | \tau_1^{(1)} | D_w^n \rangle &= \frac{\sqrt{(n-w)(w+1)}}{n}, \\ \langle D_{w+1}^n | \tau_2^{(1)} | D_w^n \rangle &= i \frac{\sqrt{(n-w)(w+1)}}{n}. \end{aligned}$$

Hence it follows that

$$\langle D_w^n | E^3 | D_w^n \rangle = \frac{n}{2} - w, \quad (\text{C4})$$

$$\langle D_{w+1}^n | E^1 | D_w^n \rangle = \frac{1}{2} \sqrt{(n-w)(w+1)}, \quad (\text{C5})$$

$$\langle D_{w+1}^n | E^2 | D_w^n \rangle = \frac{i}{2} \sqrt{(n-w)(w+1)}. \quad (\text{C6})$$

Hence, the result follows. \square

Next, we calculate the projections of $A_{j,\gamma}$ onto the symmetric subspace.

Lemma 18. For all $j = 1, \dots, n$, we have

$$\Pi^{\text{sym}} A_{j,\gamma} \Pi^{\text{sym}} = \gamma \frac{1}{n} \sum_{w=0}^{n-1} \sqrt{(n-w)(w+1)} |D_w^n\rangle\langle D_{w+1}^n|.$$

Proof. Now consider $|0\rangle\langle 1| \otimes I^{\otimes n-1} = (\tau_1^{(1)} - i\tau_2^{(1)})/2$. Hence from (C5) and (C6) we have

$$\langle D_{w+1}^n | (|0\rangle\langle 1| \otimes I^{\otimes n-1}) | D_w^n \rangle = 0.$$

Also since $|1\rangle\langle 0| \otimes I^{\otimes n-1} = (\tau_1^{(1)} + i\tau_2^{(1)})/2$, we use (C5) and (C6) to find that

$$\langle D_{w+1}^n | (|1\rangle\langle 0| \otimes I^{\otimes n-1}) | D_w^n \rangle = \sqrt{(n-w)(w+1)}/n.$$

Hence the result follows. \square

Hence the projection of \tilde{A}_γ onto the symmetric subspace is given by

$$\gamma \sum_{w=0}^{n-1} \sqrt{(n-w)(w+1)} |D_w^n\rangle \langle D_{w+1}^n|. \quad (\text{C7})$$

We use the projections of E^1, E^2, E^3 and \tilde{A}_γ on the symmetric subspace in order to do our numerical calculations entirely within the symmetric subspace.

2. SLD equation

As a preparation of our proofs of Lemmas 12 and 14, here, we solve the SLD equation for a density matrix ρ that need not have full rank. This appendix addresses a general theory for SLD equations.

Lemma 19. *Let ρ be a Hermitian matrix with spectral decomposition $\rho = \sum_k \lambda_k |\phi_k\rangle\langle \phi_k|$, such that for all k , we have $\lambda_k \geq \lambda_{k+1}$. When $\lambda_k = 0$ for all $k > p$, then we can write $\rho = \sum_{k=1}^p \lambda_k |\phi_k\rangle\langle \phi_k|$, and the solution L to*

$$\frac{1}{2}(\rho L + L\rho) = i\rho E - iE\rho. \quad (\text{C8})$$

is given by

$$L = 2i \sum_{k,l} f(k,l) \langle \phi_k | E | \phi_l \rangle |\phi_k\rangle \langle \phi_l|$$

where

$$\begin{aligned} f(k,l) &:= 0, & k, l > p, \\ f(k,l) &:= -1, & k > p, l \leq p, \\ f(k,l) &:= 1, & l > p, k \leq p, \\ f(k,l) &:= \frac{\lambda_k - \lambda_l}{\lambda_k + \lambda_l}, & k, l \leq p. \end{aligned}$$

Proof of Lemma 19. Note that

$$\begin{aligned} \rho L &= 2i \sum_{k=1}^p \sum_{l=1}^p f(k,l) \langle \phi_k | E | \phi_l \rangle \lambda_k |\phi_k\rangle \langle \phi_l| \\ &= 2i \sum_{k=1}^p \sum_{l=1}^p \frac{\lambda_k - \lambda_l}{\lambda_k + \lambda_l} \langle \phi_k | E | \phi_l \rangle \lambda_k |\phi_k\rangle \langle \phi_l| \\ &\quad - 2i \sum_{k=1}^p \sum_{l>p} \langle \phi_k | E | \phi_l \rangle \lambda_k |\phi_k\rangle \langle \phi_l|, \\ L\rho &= 2i \sum_{l=1}^p \sum_{k=1}^p f(k,l) \langle \phi_k | E | \phi_l \rangle |\phi_k\rangle \langle \phi_l| \lambda_l \\ &= 2i \sum_{l=1}^p \sum_{k=1}^p \frac{\lambda_k - \lambda_l}{\lambda_k + \lambda_l} \langle \phi_k | E | \phi_l \rangle |\phi_k\rangle \langle \phi_l| \lambda_l \\ &\quad + 2i \sum_{l=1}^p \sum_{k>p} \langle \phi_k | E | \phi_l \rangle |\phi_k\rangle \langle \phi_l| \lambda_l. \end{aligned} \quad (\text{C9})$$

We hence get

$$\begin{aligned} &\frac{1}{2}(\rho L + L\rho) \\ &= i \sum_{k=1}^p \sum_{l=1}^p (\lambda_k - \lambda_l) \langle \phi_k | E | \phi_l \rangle \lambda_k |\phi_k\rangle \langle \phi_l| \\ &\quad - i \sum_{k=1}^p \sum_{l>p} \langle \phi_k | E | \phi_l \rangle \lambda_k |\phi_k\rangle \langle \phi_l| \\ &\quad + i \sum_{l=1}^p \sum_{k>p} \langle \phi_k | E | \phi_l \rangle \lambda_l |\phi_k\rangle \langle \phi_l|. \end{aligned}$$

Now,

$$\begin{aligned} iE\rho &= i \sum_{k=1}^p \sum_{l=1}^p |\phi_k\rangle \langle \phi_k | E | \phi_l \rangle \langle \phi_l| \lambda_l \\ i\rho E &= i \sum_{l=1}^p \sum_{k=1}^p \lambda_k |\phi_k\rangle \langle \phi_k | E | \phi_l \rangle \langle \phi_l|, \end{aligned} \quad (\text{C10})$$

and hence

$$\begin{aligned} -iE\rho + i\rho E &= -i \sum_{k=1}^p \sum_{l=1}^p |\phi_k\rangle \langle \phi_k | E | \phi_l \rangle \langle \phi_l| \lambda_l \\ &\quad + i \sum_{l=1}^p \sum_{k=1}^p \lambda_k |\phi_k\rangle \langle \phi_k | E | \phi_l \rangle \langle \phi_l|, \\ &= i \sum_{k=1}^p \sum_{l=1}^p (\lambda_k - \lambda_l) \langle \phi_k | E | \phi_l \rangle |\phi_k\rangle \langle \phi_l| \\ &\quad - i \sum_{k>p} \sum_{l=1}^p |\phi_k\rangle \langle \phi_k | E | \phi_l \rangle \langle \phi_l| \lambda_l \\ &\quad + i \sum_{l>p} \sum_{k=1}^p \lambda_k |\phi_k\rangle \langle \phi_k | E | \phi_l \rangle \langle \phi_l|. \end{aligned} \quad (\text{C11})$$

This proves the lemma. \square

Next we give a lemma that evaluates $\text{Tr}(L_j \rho L_k)$. We prove this using Lemma 19.

Lemma 20. *Let $\rho = \sum_l \lambda_l |\phi_l\rangle\langle\phi_l|$ where $|\phi_l\rangle$ are orthonormal vectors and $\lambda_l = 0$ for all $l > p$. Furthermore let L_j be the solution to $\frac{1}{2}(\rho L_j + L_j \rho) = i\rho E^j - iE^j \rho$ according to Lemma 19. Then*

$$\begin{aligned} & \text{Tr}(L_u \rho L_v) \\ &= 4 \sum_{k=1}^p \sum_{l=1}^p \lambda_l (f(k, l)^2 - 1) \langle \phi_k | E^u | \phi_l \rangle \langle \phi_l | E^v | \phi_k \rangle \\ & \quad + 4 \sum_{l=1}^p \lambda_l \langle \phi_l | E^v E^u | \phi_l \rangle. \end{aligned}$$

Proof of Lemma 20. From Lemma 19, note that

$$\begin{aligned} & \text{Tr}(L_u \rho L_v) \\ &= -4 \text{Tr} \sum_{k,l} f(k, l) \langle \phi_k | E^u | \phi_l \rangle | \phi_k \rangle \langle \phi_l | \\ & \quad \times \sum_{a=1}^p \lambda_a | \phi_a \rangle \langle \phi_a | \sum_{k', l'} f(k', l') \langle \phi_{k'} | E^v | \phi_{l'} \rangle | \phi_{k'} \rangle \langle \phi_{l'} |. \end{aligned}$$

We can simplify this expression to

$$\begin{aligned} & \text{Tr}(L_u \rho L_v) \\ &= -4 \sum_k \sum_{l=1}^p \lambda_l f(k, l) f(l, k) \langle \phi_k | E^u | \phi_l \rangle \langle \phi_l | E^v | \phi_k \rangle \\ &= 4 \sum_k \sum_{l=1}^p \lambda_l f(k, l)^2 \langle \phi_k | E^u | \phi_l \rangle \langle \phi_l | E^v | \phi_k \rangle \\ &= 4 \sum_{k=1}^p \sum_{l=1}^p \lambda_l f(k, l)^2 \langle \phi_k | E^u | \phi_l \rangle \langle \phi_l | E^v | \phi_k \rangle \\ & \quad + 4 \sum_{l=1}^p \lambda_l \langle \phi_l | E^v \left(\sum_{k>p} | \phi_k \rangle \langle \phi_k | \right) E^u | \phi_l \rangle \\ &= 4 \sum_{k=1}^p \sum_{l=1}^p \lambda_l f(k, l)^2 \langle \phi_k | E^u | \phi_l \rangle \langle \phi_l | E^v | \phi_k \rangle \\ & \quad + 4 \sum_{l=1}^p \lambda_l \langle \phi_l | E^v \left(I - \sum_{k=1}^p | \phi_k \rangle \langle \phi_k | \right) E^u | \phi_l \rangle \\ &= 4 \sum_{k=1}^p \sum_{l=1}^p \lambda_l (f(k, l)^2 - 1) \langle \phi_k | E^u | \phi_l \rangle \langle \phi_l | E^v | \phi_k \rangle \\ & \quad + 4 \sum_{l=1}^p \lambda_l \langle \phi_l | E^v E^u | \phi_l \rangle. \end{aligned}$$

□

3. Proof of Lemma 12

The goal of this section is to show Lemma 12, i.e., (67), which expresses $\text{Tr}[\bar{L}_u \bar{\rho}_\gamma \bar{L}_v]$ in terms of the eigenvectors

and eigenvalues of ρ_γ and the operators E^u and E^v . To achieve this goal, we treat (67) as a special case of Lemma 20. Then, we substituting ρ for $\bar{\rho}_\gamma$, and note that $\bar{\rho}_\gamma$ has two non-zero eigenvalues, so we can set $p = 2$. Then we get

$$\begin{aligned} & \text{Tr}(\bar{L}_u \bar{\rho}_\gamma \bar{L}_v) \\ & \stackrel{(a)}{=} 4 \sum_{k=1}^2 \sum_{l=1}^2 \lambda_l (f(k, l)^2 - 1) \langle \phi_k | F^u | \phi_l \rangle \langle \phi_l | F^v | \phi_k \rangle \\ & \quad + 4 \sum_{l=1}^2 \lambda_l \langle \phi_l | F^v F^u | \phi_l \rangle \end{aligned} \tag{C12}$$

$$\begin{aligned} & \stackrel{(b)}{=} 4(f(1, 1)^2 - 1) \lambda_1 \langle \phi_1 | F^u | \phi_1 \rangle \langle \phi_1 | F^v | \phi_1 \rangle \\ & \quad + 4(f(2, 2)^2 - 1) \lambda_2 \langle \phi_2 | F^u | \phi_2 \rangle \langle \phi_2 | F^v | \phi_2 \rangle \\ & \quad + 4g \lambda_1 \langle \phi_2 | F^u | \phi_1 \rangle \langle \phi_1 | F^v | \phi_2 \rangle \\ & \quad + 4g \lambda_2 \langle \phi_1 | F^u | \phi_2 \rangle \langle \phi_2 | F^v | \phi_1 \rangle \\ & \quad + 4 \sum_{l=1}^2 \lambda_l \langle \phi_l | F^v F^u | \phi_l \rangle \end{aligned} \tag{C13}$$

$$\begin{aligned} & \stackrel{(c)}{=} -4 \lambda_1 \langle \phi_1 | F^u | \phi_1 \rangle \langle \phi_1 | F^v | \phi_1 \rangle \\ & \quad - 4 \lambda_2 \langle \phi_2 | F^u | \phi_2 \rangle \langle \phi_2 | F^v | \phi_2 \rangle \\ & \quad + 4g \lambda_1 \langle \phi_2 | F^u | \phi_1 \rangle \langle \phi_1 | F^v | \phi_2 \rangle \\ & \quad + 4g \lambda_2 \langle \phi_1 | F^u | \phi_2 \rangle \langle \phi_2 | F^v | \phi_1 \rangle \\ & \quad + 4 \sum_{l=1}^2 \lambda_l \langle \phi_l | F^v F^u | \phi_l \rangle. \end{aligned} \tag{C14}$$

where in (a) we use $\bar{\rho}_\gamma = \lambda_1 |\phi_1\rangle\langle\phi_1| + \lambda_2 |\phi_2\rangle\langle\phi_2|$ and replace ρ with $\bar{\rho}_\gamma$, L^u with \bar{L}^u , and E^j with F^j in the SLD equation $\frac{1}{2}(\rho L_j + L_j \rho) = i\rho E^j - iE^j \rho$, in (b) we use $f(1, 2)^2 = f(2, 1)^2$ and $g = f(1, 2)^2 - 1$, in (c) we use $f(1, 1) = f(2, 2) = 0$. Since we show (C14), (67) follows.

4. Proof of Lemma 13

Here, we prove Lemma 13. Note that

$$\begin{aligned}
& \langle \phi | F^1 | \epsilon \rangle \\
& \stackrel{(a)}{=} \frac{1}{n+1} \sum_{u=0}^n \langle D_u^n | (I - \frac{1}{2} \tilde{A}_\gamma^\dagger \tilde{A}_\gamma) E^1 \tilde{A}_\gamma | D_u^n \rangle \\
& \stackrel{(b)}{=} \frac{\gamma}{n+1} \sum_{u=1}^n \langle D_u^n | (I - \frac{1}{2} \tilde{A}_\gamma^\dagger \tilde{A}_\gamma) E^1 \sqrt{(n-u+1)u} | D_{u-1}^n \rangle \\
& \stackrel{(c)}{=} \frac{\gamma}{2(n+1)} \sum_{u=1}^n \langle D_u^n | (I - \frac{1}{2} \tilde{A}_\gamma^\dagger \tilde{A}_\gamma) (n-u+1)u | D_u^n \rangle \\
& \stackrel{(d)}{=} \frac{\gamma}{2(n+1)} \sum_{u=1}^n \langle D_u^n | (1 - \frac{\gamma^2}{2} (n-u+1)u) (n-u+1)u | D_u^n \rangle \\
& \stackrel{(e)}{=} \frac{\gamma}{2(n+1)} \sum_{u=1}^n (1 - \frac{\gamma^2}{2} (n-u+1)u) (n-u+1)u \\
& \stackrel{(f)}{=} \frac{n\gamma}{12} (n+2) - \frac{n\gamma^3}{120} (n+2)(n^2+2n+2).
\end{aligned}$$

In (a) we use the definition of $|\phi\rangle$ and $|\epsilon\rangle$. In (b) we use (C7). In (c) we use (C2). In (d) we use (C7). In (e) we use the orthonormality of Dicke states. In (f) we perform the summation over u .

We can similarly calculate $\langle \phi | F^2 | \epsilon \rangle$. Furthermore, using the same ideas, we can calculate

$$\begin{aligned}
\langle \phi | F^3 | \phi \rangle &= \langle \Phi | F^3 | \Phi \rangle - \langle \Phi | \bar{E}^3 (A^\dagger A \otimes I_C) | \Phi \rangle + O(\gamma^4) \\
&= 0 - \gamma^2 \sum_{a=1}^n (n-a+1)a(n/2-a) + O(\gamma^4) \\
&= \gamma^2 n(n+2)/12 + O(\gamma^4), \tag{C15} \\
\langle \epsilon | F^3 | \epsilon \rangle &= \frac{1}{n+1} \sum_{a=0}^{n-1} (n-a)(a+1)(n/2-a) \\
&= \gamma^2 n(n+2)/12.
\end{aligned}$$

5. Proof of Lemma 14

Here we prove Lemma 14. First, we consider the spectral decomposition of ρ_γ given by

$$\rho_\gamma = \sum_{j=1}^{(n+1)^2} \lambda'_j |\phi'_j\rangle \langle \phi'_j|,$$

where $\lambda'_j \geq \lambda'_{j+1}$. Recall that the spectral decomposition of $\bar{\rho}_\gamma$ is given by

$$\bar{\rho}_\gamma = \lambda_1 |\phi_1\rangle \langle \phi_1| + \lambda_2 |\phi_2\rangle \langle \phi_2|.$$

Lemma 21. For eigenvalues λ_1, λ_2 and $\lambda'_1, \lambda'_2, \dots, \lambda'_{n+1}$, we have

$$\begin{aligned}
\lambda'_1 - \lambda_1 &= O(\gamma^4), \quad \lambda'_2 - \lambda_2 = O(\gamma^4), \\
\lambda'_j &= O(\gamma^4), \quad \text{for all } j \geq 3,
\end{aligned} \tag{C16}$$

and

$$\frac{(\lambda_2 - \lambda_1)^2}{(\lambda_2 + \lambda_1)^2} - 1 = c_2 \gamma^2 + O(\gamma^4), \tag{C17}$$

$$\frac{(\lambda'_2 - \lambda'_1)^2}{(\lambda'_2 + \lambda'_1)^2} - 1 = c_2 \gamma^2 + O(\gamma^4), \tag{C18}$$

$$\frac{(\lambda'_k - \lambda'_1)^2}{(\lambda'_k + \lambda'_1)^2} - 1 = O(\gamma^4), \quad \text{for all } k \geq 3, \tag{C19}$$

$$\frac{(\lambda'_k - \lambda'_2)^2}{(\lambda'_k + \lambda'_2)^2} - 1 = O(\gamma^2), \quad \text{for all } k \geq 3, \tag{C20}$$

where $c_2 = 2n(n+2)/3$.

Proof. The relations (C16) follow from the Gersgorin circle theorem [50],

The relation (C17) follows from (66). Now, we have

$$\begin{aligned}
& \left(\frac{(\lambda'_2 - \lambda'_1)^2}{(\lambda'_2 + \lambda'_1)^2} - 1 \right) - \left(\frac{(\lambda_2 - \lambda_1)^2}{(\lambda_2 + \lambda_1)^2} - 1 \right) \\
& \stackrel{(a)}{=} \frac{(\lambda'_2 - \lambda'_1)^2}{(\lambda'_2 + \lambda'_1)^2} - \frac{(\lambda'_2 - \lambda'_1)^2}{(\lambda_2 + \lambda_1)^2} + \frac{(\lambda'_2 - \lambda'_1)^2}{(\lambda_2 + \lambda_1)^2} - \frac{(\lambda_2 - \lambda_1)^2}{(\lambda_2 + \lambda_1)^2} \\
& \stackrel{(b)}{=} (\lambda'_2 - \lambda'_1)^2 ((\lambda'_2 + \lambda'_1)^{-2} - (\lambda_2 + \lambda_1)^{-2}) \\
& \quad + ((\lambda'_2 - \lambda'_1)^2 - (\lambda_2 - \lambda_1)^2) (\lambda_2 + \lambda_1)^{-2} \\
& \stackrel{(c)}{=} (\lambda'_2 - \lambda'_1)^2 O(\gamma^4) + ((\lambda'_2 - \lambda'_1)^2 - (\lambda_2 - \lambda_1)^2) (1 + O(\gamma^4)) \\
& \stackrel{(d)}{=} (1 + O(\gamma^2)) O(\gamma^4) + O(\gamma^4) (1 + O(\gamma^4)) = O(\gamma^4), \tag{C21}
\end{aligned}$$

where in (a) we use a telescoping sum, in (b) we collect terms in the telescoping sum, in (c) we use $\lambda_1 + \lambda_2 = 1 + O(\gamma^4)$, $\lambda'_1 + \lambda'_2 = 1 + O(\gamma^4)$, which implies that $(\lambda_1 + \lambda_2)^{-2} = 1 + O(\gamma^4)$ and $(\lambda'_1 + \lambda'_2)^{-2} = 1 + O(\gamma^4)$, in (d) we use $(\lambda'_2 - \lambda'_1)^2 = (\lambda_2 - \lambda_1)^2 + O(\gamma^4)$ because of (C16) and also $\lambda'_2 - \lambda'_1 = 2c_2 \gamma^2 - 1 + O(\gamma^4)$. Hence, from (C17) and (C21), we have (C18).

Since $\lambda'_1 = 1 - c_2 \gamma^2 + O(\gamma^4)$ and $\lambda'_k = O(\gamma^4)$ for all $k \geq 3$, we have (C19). Since $\lambda'_2 = c_2 \gamma^2 + O(\gamma^4)$ and $\lambda'_k = O(\gamma^4)$ for all $k \geq 3$, we have (C20). \square

We prepare several lemmas.

Lemma 22. We have

$$\rho_\gamma - \bar{\rho}_\gamma = O(\gamma^4).$$

Proof. We have

$$\begin{aligned}
\rho_\gamma &\stackrel{(a)}{=} (\Lambda_{0,\gamma} \otimes \iota_C)(|\Phi\rangle\langle\Phi|) \\
&\stackrel{(b)}{=} (e^{\mathcal{L}_0} \otimes \iota_C)(|\Phi\rangle\langle\Phi|) \\
&\stackrel{(c)}{=} |\Phi\rangle\langle\Phi| + (\tilde{A}_\gamma \otimes I_C)|\Phi\rangle\langle\Phi|(\tilde{A}_\gamma \otimes I_C)^\dagger \\
&\quad - \frac{1}{2}(\tilde{A}_\gamma^\dagger \tilde{A}_\gamma \otimes I_C)|\Phi\rangle\langle\Phi| - \frac{1}{2}|\Phi\rangle\langle\Phi|(\tilde{A}_\gamma^\dagger \tilde{A}_\gamma \otimes I_C) + O(\gamma^4) \\
&\stackrel{(d)}{=} |\Phi\rangle\langle\Phi| + (\tilde{A}_\gamma \otimes I_C)|\Phi\rangle\langle\Phi|(\tilde{A}_\gamma \otimes I_C)^\dagger \\
&\quad - \frac{1}{2}(\tilde{A}_\gamma^\dagger \tilde{A}_\gamma \otimes I_C)|\Phi\rangle\langle\Phi| - \frac{1}{2}|\Phi\rangle\langle\Phi|(\tilde{A}_\gamma^\dagger \tilde{A}_\gamma \otimes I_C) \\
&\quad + \frac{1}{4}(\tilde{A}_\gamma^\dagger \tilde{A}_\gamma \otimes I_C)|\Phi\rangle\langle\Phi|(\tilde{A}_\gamma^\dagger \tilde{A}_\gamma \otimes I_C) + O(\gamma^4) \\
&\stackrel{(e)}{=} |\phi\rangle\langle\phi| + |\epsilon\rangle\langle\epsilon| + O(\gamma^4) \stackrel{(f)}{=} \bar{\rho}_\gamma + O(\gamma^4), \tag{C22}
\end{aligned}$$

where the derivation of each step are given as follows. Step (a) follows from the definition of ρ_γ . Step (b) follows from (57). Step (c) follows from (56) and (57). Step (d) follows from the fact that $(\tilde{A}_\gamma^\dagger \tilde{A}_\gamma \otimes I_C)|\Phi\rangle\langle\Phi|(\tilde{A}_\gamma^\dagger \tilde{A}_\gamma \otimes I_C)$ is $O(\gamma^4)$. Step (e) follows from (60) and (61). Step (f) follows from (64). \square

Next, we have the perturbation bound for the eigenvectors of ρ_γ .

Lemma 23.

$$\| |\phi'_1\rangle\langle\phi'_1| - |\phi_1\rangle\langle\phi_1| \|_1 = O(\gamma^4) \tag{C23}$$

$$\| |\phi'_2\rangle\langle\phi'_2| - |\phi_2\rangle\langle\phi_2| \|_1 = O(\gamma^2). \tag{C24}$$

Proof. Step 1: First, we show (C24). Let us use the following argument according to [51, Eq (7)]. Let A be a Hermitian matrix with nonzero eigenvalues a_1, \dots, a_r with corresponding eigenvectors v_1, \dots, v_r , where $a_j \geq a_{j+1}$. Define the gap of A as

$$g(A) = \min_{j=1, \dots, r} a_j - a_{j+1}$$

Let $A + E$ be a Hermitian matrix with eigenvalues μ_j , and eigenvectors w_j , where $\mu_j \geq \mu_{j+1}$. Then for all $j = 1, \dots, r$, we have

$$\| |v_j\rangle - |w_j\rangle \| \leq 2\sqrt{2}\|E\|_2/g(A), \tag{C25}$$

where $\|\cdot\|_2$ is the Hilbert-Schmidt norm. Now the smallest non-zero eigenvalue of ρ_γ is $c\gamma^2$ for some $c > 0$ (see (63)) and Lemma 22, i.e., $\rho_\gamma - \bar{\rho}_\gamma = O(\gamma^4)$. Hence,

$$\begin{aligned}
&\| |\phi_2\rangle\langle\phi_2| - |\phi'_2\rangle\langle\phi'_2| \|_1 \leq \sqrt{1 - |\langle\phi'_2|\phi_2\rangle|^2} \\
&\leq \| |\phi_2\rangle - |\phi'_2\rangle \| \\
&\leq 2\sqrt{2}\|\rho_\gamma - \bar{\rho}_\gamma\|_2/g(\rho_\gamma) = O(\gamma^4/\gamma^2) = O(\gamma^2),
\end{aligned}$$

which shows (C24).

Step 2: Toward the proof of (C23), for simplicity in notation, we define $B := \tilde{A}_\gamma \otimes I_C$ and

$$|\phi^{(2)}\rangle := (I_{AC} - \frac{1}{2}B^\dagger B + \frac{1}{8}B^\dagger B B^\dagger B)|\Phi\rangle$$

$$|\epsilon^{(2)}\rangle := (B - \frac{1}{4}B B^\dagger B - \frac{1}{4}B^\dagger B B)|\Phi\rangle.$$

The vectors $|\phi^{(2)}\rangle$ and $|\epsilon^{(2)}\rangle$ are orthogonal for the same reason that $|\phi\rangle$ and $|\epsilon\rangle$ are orthogonal. Let us define

$$\rho_\gamma^{(2)} := |\phi^{(2)}\rangle\langle\phi^{(2)}| + |\epsilon^{(2)}\rangle\langle\epsilon^{(2)}|.$$

The aim of this step is to show the equation

$$\rho_\gamma = \rho_\gamma^{(2)} + O(\gamma^6). \tag{C26}$$

Since $\theta = 0$, we have

$$\rho_\gamma = |\Phi\rangle\langle\Phi| + \sum_{j=1}^{\infty} \frac{1}{j!} (\mathcal{L}_{\theta,\gamma} \otimes \iota_C)^j (|\Phi\rangle\langle\Phi|).$$

We note that

$$\rho_\gamma = |\Phi\rangle\langle\Phi| + \sum_{j=2}^{\infty} \frac{1}{j!} (\mathcal{L}_{\theta,\gamma} \otimes \iota_C)^j (|\Phi\rangle\langle\Phi|) + O(\gamma^6). \tag{C27}$$

Now

$$\begin{aligned}
&(\mathcal{L}_{\theta,\gamma} \otimes \iota_C)^2 (|\Phi\rangle\langle\Phi|) \\
&= BB|\Phi\rangle\langle\Phi|B^\dagger B^\dagger - \frac{1}{2}(BB^\dagger B|\Phi\rangle\langle\Phi|B^\dagger + B|\Phi\rangle\langle\Phi|B^\dagger BB^\dagger) \\
&\quad - \frac{1}{2}B^\dagger BB|\Phi\rangle\langle\Phi|B^\dagger + \frac{1}{4}B^\dagger BB^\dagger B|\Phi\rangle\langle\Phi| + \frac{1}{4}B^\dagger B|\Phi\rangle\langle\Phi|B^\dagger B \\
&\quad - \frac{1}{2}B|\Phi\rangle\langle\Phi|B^\dagger B^\dagger B + \frac{1}{4}B^\dagger B|\Phi\rangle\langle\Phi|B^\dagger B + \frac{1}{4}|\Phi\rangle\langle\Phi|B^\dagger BB^\dagger B \\
&= |\phi^{(2)}\rangle\langle\phi^{(2)}| + |\epsilon^{(2)}\rangle\langle\epsilon^{(2)}| = \rho_\gamma^{(2)}. \tag{C28}
\end{aligned}$$

The combination of (C27) and (C28) implies (C26).

Step 3: Now, define

$$\lambda_1^{(2)} := \langle\phi^{(2)}|\phi^{(2)}\rangle$$

$$\lambda_2^{(2)} := \langle\epsilon^{(2)}|\epsilon^{(2)}\rangle$$

$$|\phi_1^{(2)}\rangle := |\phi^{(2)}\rangle / \sqrt{\lambda_1^{(2)}}$$

$$|\phi_2^{(2)}\rangle := |\epsilon^{(2)}\rangle / \sqrt{\lambda_2^{(2)}},$$

The aim of this step is to show

$$\| |\phi_1^{(2)}\rangle - |\phi_1\rangle \| = O(\gamma^4). \tag{C29}$$

Now we have

$$\begin{aligned}
\lambda_1/\lambda_1^{(2)} &= \lambda_1/(\lambda_1 + \lambda_1^{(2)} - \lambda_1) \\
&= (1 + (\lambda_1^{(2)} - \lambda_1)/\lambda_1)^{-1} \\
&\stackrel{(a)}{=} (1 + O(\gamma^4)(1 + O(\gamma^2))^{-1})^{-1} \\
&= (1 + O(\gamma^4))^{-1} \\
&= 1 + O(\gamma^4)^{-1} \tag{C30}
\end{aligned}$$

where in (a) we use $(\lambda_1^{(2)} - \lambda_1) = O(\gamma^4)$ and $\lambda_1 = 1 + O(\gamma^2)$.

Since $|\phi_1\rangle = |\phi\rangle / \sqrt{\lambda_1}$ where $|\phi\rangle$ is defined in (60), we see that

$$\begin{aligned} \sqrt{\lambda_1}|\phi_1\rangle &= (I_A - \frac{1}{2}\tilde{A}_\gamma^\dagger\tilde{A}_\gamma) \otimes I_C |\Phi\rangle \\ &\stackrel{(a)}{=} (I_A \otimes I_C - \frac{1}{2}(\tilde{A}_\gamma \otimes I_C)^\dagger(\tilde{A}_\gamma \otimes I_C))|\Phi\rangle \\ &\stackrel{(b)}{=} (I_{AC} - \frac{1}{2}B^\dagger B)|\Phi\rangle, \end{aligned} \quad (C31)$$

where (a) is because we distribute the tensor product, and (b) is because $I_A \otimes I_C = I_{AC}$ and $B = \tilde{A}_\gamma \otimes I_C$.

Now we can write

$$\begin{aligned} |\phi_1^{(2)}\rangle &= (\lambda_1^{(2)})^{-1/2} (I_{AC} - \frac{1}{2}B^\dagger B + \frac{1}{8}B^\dagger B B^\dagger B) |\Phi\rangle \\ &\stackrel{(a)}{=} (\lambda_1^{(2)})^{-1/2} \sqrt{\lambda_1} |\phi_1\rangle + (\lambda_1^{(2)})^{-1/2} \frac{1}{8} B^\dagger B B^\dagger B |\Phi\rangle \\ &\stackrel{(b)}{=} (1 + O(\gamma^4)) |\phi_1\rangle + (1 + O(\gamma^2)) \frac{1}{8} B^\dagger B B^\dagger B |\Phi\rangle \end{aligned}$$

where in (a) we use (C31) which tells us that $\sqrt{\lambda_1}|\phi_1\rangle = (I_{AC} - \frac{1}{2}B^\dagger B)|\Phi\rangle$, and in (b) we use (C30) and $(\lambda_1^{(2)})^{-1/2} = 1 + O(\gamma^2)$. Hence

$$\begin{aligned} \||\phi_1^{(2)}\rangle - |\phi_1\rangle\| &= \|O(\gamma^4)|\phi_1\rangle + (1 + O(\gamma^2)) \frac{1}{8} B^\dagger B B^\dagger B |\Phi\rangle\| \\ &= O(\gamma^4). \end{aligned}$$

Step 4: The aim of this step is to derive (C23) by using (C26) and (C29).

First, we notice the spectral decomposition

$$\rho_\gamma^{(2)} = \lambda_1^{(2)} |\phi_1^{(2)}\rangle \langle \phi_1^{(2)}| + \lambda_2^{(2)} |\phi_2^{(2)}\rangle \langle \phi_2^{(2)}|.$$

Since the smallest non-zero eigenvalue of $\rho_\gamma^{(2)} = c\gamma^2 + O(\gamma^4)$ for some positive c , and (C26), we can use (C25) to find that

$$\begin{aligned} \||\phi_1^{(2)}\rangle - |\phi_1'\rangle\| &\leq 2\sqrt{2} \|\rho_\gamma - \bar{\rho}_\gamma^{(2)}\|_2 / g(\rho_\gamma^{(2)}) \\ &= O(\gamma^6) / (c\gamma^2 + O(\gamma^4)) = O(\gamma^4). \end{aligned} \quad (C32)$$

Hence combining (C29) with (C32) with the triangle inequality gives

$$\begin{aligned} \||\phi_1\rangle \langle \phi_1| - |\phi_1'\rangle \langle \phi_1'|\|_1 &\leq \||\phi_1\rangle - |\phi_1'\rangle\| \\ &\leq \||\phi_1^{(2)}\rangle - |\phi_1'\rangle\| + \||\phi_1^{(2)}\rangle - |\phi_1\rangle\| = O(\gamma^4). \end{aligned}$$

□

Next, using Lemmas 20 and 22, and defining $Q := F^\nu F^u - F^u F^\nu$, we prove the following lemma.

Lemma 24. Let $u, \nu = 1, 2, 3$. Then we have

$$\begin{aligned} &\text{Tr}(L_u \rho_\gamma L_\nu) - \text{Tr}(L_\nu \rho_\gamma L_u) \\ &= 4 \sum_{\substack{1 \leq k, l \leq p' \\ k \neq l}} \lambda_l' \left(\frac{(\lambda_k' - \lambda_l')^2}{(\lambda_k' + \lambda_l')^2} - 1 \right) \langle \phi_k' | F^u | \phi_l' \rangle \langle \phi_l' | F^\nu | \phi_k' \rangle \\ &\quad + 4 \sum_{l=1}^{p'} \lambda_l' \langle \phi_l' | Q | \phi_l' \rangle, \end{aligned} \quad (C33)$$

and

$$|\langle \phi_1' | Q | \phi_1' \rangle - \langle \phi_1 | Q | \phi_1 \rangle| \leq O(\gamma^4), \quad (C34)$$

$$|\langle \phi_2' | Q | \phi_2' \rangle - \langle \phi_2 | Q | \phi_2 \rangle| \leq O(\gamma^2), \quad (C35)$$

$$4 \sum_{l=1}^{p'} \lambda_l' \langle \phi_l' | Q | \phi_l' \rangle - 4 \sum_{l=1}^2 \lambda_l \langle \phi_l | Q | \phi_l \rangle = O(\gamma^4). \quad (C36)$$

For $(k, l) \in \{(1, 2), (2, 1)\}$, we have

$$|\langle \phi_k' | F^u | \phi_l' \rangle \langle \phi_l' | F^\nu | \phi_k' \rangle - \langle \phi_k | F^u | \phi_l \rangle \langle \phi_l | F^\nu | \phi_k \rangle| \leq O(\gamma^2). \quad (C37)$$

Proof. Now from Lemma 20 it follows that

$$\begin{aligned} &\text{Tr}(L_u \rho_\gamma L_\nu) \\ &= 4 \sum_{k=1}^{p'} \sum_{l=1}^{p'} \lambda_l' \left(\frac{(\lambda_k' - \lambda_l')^2}{(\lambda_k' + \lambda_l')^2} - 1 \right) \langle \phi_k' | F^u | \phi_l' \rangle \langle \phi_l' | F^\nu | \phi_k' \rangle \\ &\quad + 4 \sum_{l=1}^{p'} \lambda_l' \langle \phi_l' | F^\nu F^u | \phi_l' \rangle. \end{aligned} \quad (C38)$$

where p' is the number of non-zero eigenvalues of ρ_γ . Then from (C38), we get (C33) as

$$\begin{aligned} &\text{Tr}(L_u \rho_\gamma L_\nu) - \text{Tr}(L_\nu \rho_\gamma L_u) \\ &= 4 \sum_{\substack{1 \leq k, l \leq p' \\ k \neq l}} \lambda_l' \left(\frac{(\lambda_k' - \lambda_l')^2}{(\lambda_k' + \lambda_l')^2} - 1 \right) \langle \phi_k' | F^u | \phi_l' \rangle \langle \phi_l' | F^\nu | \phi_k' \rangle \\ &\quad + 4 \sum_{l=1}^{p'} \lambda_l' \langle \phi_l' | (F^\nu F^u - F^u F^\nu) | \phi_l' \rangle. \end{aligned}$$

Lemma 23 implies (C34) and (C35) as

$$\begin{aligned} |\langle \phi_1' | Q | \phi_1' \rangle - \langle \phi_1 | Q | \phi_1 \rangle| &\leq \|Q\| \cdot \||\phi_1'\rangle \langle \phi_1'| - |\phi_1\rangle \langle \phi_1|\|_1 \\ &= \|Q\| O(\gamma^4) = O(\gamma^4) \\ |\langle \phi_2' | Q | \phi_2' \rangle - \langle \phi_2 | Q | \phi_2 \rangle| &\leq \|Q\| \cdot \||\phi_2'\rangle \langle \phi_2'| - |\phi_2\rangle \langle \phi_2|\|_1 \\ &= \|Q\| O(\gamma^2) = O(\gamma^2). \end{aligned} \quad (C39)$$

From (C16) and it follows that

$$\begin{aligned} &4 \sum_{l=1}^{p'} \lambda_l' \langle \phi_l' | Q | \phi_l' \rangle - 4 \sum_{l=1}^2 \lambda_l \langle \phi_l | Q | \phi_l \rangle \\ &\stackrel{(a)}{=} 4 \lambda_1' \langle \phi_1' | Q | \phi_1' \rangle - 4 \lambda_1 \langle \phi_1 | Q | \phi_1 \rangle \\ &\quad + 4 \sum_{l=2}^{p'} \lambda_l' \langle \phi_l' | Q | \phi_l' \rangle - 4 \lambda_2 \langle \phi_2 | Q | \phi_2 \rangle \\ &\stackrel{(b)}{=} 4 \lambda_1' \langle \phi_1' | Q | \phi_1' \rangle - 4 \lambda_1 \langle \phi_1 | Q | \phi_1 \rangle \\ &\quad + 4 \lambda_2' \langle \phi_2' | Q | \phi_2' \rangle - 4 \lambda_2 \langle \phi_2 | Q | \phi_2 \rangle + O(\gamma^4) \\ &\stackrel{(c)}{=} 4 \lambda_1' \langle \phi_1 | Q | \phi_1 \rangle + \lambda_1' O(\gamma^4) - 4 \lambda_1 \langle \phi_1 | Q | \phi_1 \rangle \\ &\quad + 4 \lambda_2' \langle \phi_2 | Q | \phi_2 \rangle + \lambda_2' O(\gamma^2) - 4 \lambda_2 \langle \phi_2 | Q | \phi_2 \rangle \\ &\stackrel{(d)}{=} 4(\lambda_1 + O(\gamma^4)) \langle \phi_1 | Q | \phi_1 \rangle + O(\gamma^4) - 4 \lambda_1 \langle \phi_1 | Q | \phi_1 \rangle \\ &\quad + 4(\lambda_2 + O(\gamma^4)) \langle \phi_2 | Q | \phi_2 \rangle + O(\gamma^2) - 4 \lambda_2 \langle \phi_2 | Q | \phi_2 \rangle \end{aligned}$$

where in (a) we separate terms in the summation, in (b) we use (C16) which means that $\lambda'_l = O(\gamma^4)$ when $l \geq 3$, in (c) we use (C34) and (C35), in (d) we use (C16). Simplifying the above, we find (C36)

Next note that for $(k, l) \in \{(1, 2), (2, 1)\}$, we have

$$\begin{aligned}
& |\langle \phi'_k | F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle - \langle \phi_k | F^u | \phi_l \rangle \langle \phi_l | F^v | \phi_k \rangle| \\
& \stackrel{(a)}{=} |\text{Tr}(F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \langle \phi'_k |) - \text{Tr}(F^u | \phi_l \rangle \langle \phi_l | F^v | \phi_k \rangle \langle \phi_k |)| \\
& \stackrel{(b)}{=} |\text{Tr}(F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \langle \phi'_k |) - \text{Tr}(F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi_k \rangle \langle \phi_k |) \\
& \quad + |\text{Tr}(F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi_k \rangle \langle \phi_k |) - \text{Tr}(F^u | \phi_l \rangle \langle \phi_l | F^v | \phi_k \rangle \langle \phi_k |)| \\
& \stackrel{(c)}{=} O(\| |\phi'_k \rangle \langle \phi'_k | - |\phi_k \rangle \langle \phi_k | \|_1) + O(\| |\phi'_l \rangle \langle \phi'_l | - |\phi_l \rangle \langle \phi_l | \|_1) \\
& \stackrel{(d)}{=} O(\gamma^2).
\end{aligned}$$

where (a) follows by the cyclicity of the trace, (b) uses a telescoping sum with the triangle inequality, (c) follows because

$$\begin{aligned}
& |\text{Tr}(F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle (|\phi'_k \rangle \langle \phi'_k | - |\phi_k \rangle \langle \phi_k |))| \\
& \leq \|F^u | \phi'_l \rangle \langle \phi'_l | F^v \| \cdot \| |\phi'_k \rangle \langle \phi'_k | - |\phi_k \rangle \langle \phi_k | \|_1 \\
& \leq \|F^u \| \|F^v \| \cdot \| |\phi'_k \rangle \langle \phi'_k | - |\phi_k \rangle \langle \phi_k | \|_1,
\end{aligned}$$

and

$$\begin{aligned}
& |\text{Tr}(|\phi'_l \rangle \langle \phi'_l | - |\phi_l \rangle \langle \phi_l |)(F^v | \phi_k \rangle \langle \phi_k | F^u)| \\
& \leq \| |\phi'_l \rangle \langle \phi'_l | - |\phi_l \rangle \langle \phi_l | \|_1 \cdot \|F^v | \phi_k \rangle \langle \phi_k | F^u \| \\
& \leq \|F^u \| \|F^v \| \cdot \| |\phi'_l \rangle \langle \phi'_l | - |\phi_l \rangle \langle \phi_l | \|_1,
\end{aligned}$$

and (d) follows from Lemma 23. Hence, we obtain (C37). \square

Finally, using Lemmas 24 and 21, we derive (74). We can

get

$$\begin{aligned}
& (\text{Tr}(L_u \rho_\gamma L_v) - \text{Tr}(L_v \rho_\gamma L_u)) - (\text{Tr}(\bar{L}_u \bar{\rho}_\gamma \bar{L}_v) - \text{Tr}(\bar{L}_v \bar{\rho}_\gamma \bar{L}_u)) \\
& \stackrel{(a)}{=} 4 \sum_{\substack{1 \leq k, l \leq p' \\ k \neq l}} \lambda'_l \left(\frac{(\lambda'_k - \lambda'_l)^2}{(\lambda'_k + \lambda'_l)^2} - 1 \right) \langle \phi'_k | F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \\
& \quad - 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda_l \left(\frac{(\lambda_k - \lambda_l)^2}{(\lambda_k + \lambda_l)^2} - 1 \right) \langle \phi_k | F^u | \phi_l \rangle \langle \phi_l | F^v | \phi_k \rangle \\
& \quad + 4 \sum_{l=1}^{p'} \lambda'_l \langle \phi'_l | (F^v F^u - F^u F^v) | \phi'_l \rangle \\
& \quad - 4 \sum_{l=1}^2 \lambda_l \langle \phi_l | (F^v F^u - F^u F^v) | \phi_l \rangle \\
& \stackrel{(b)}{=} 4 \sum_{\substack{1 \leq k, l \leq p' \\ k \neq l}} \lambda'_l \left(\frac{(\lambda'_k - \lambda'_l)^2}{(\lambda'_k + \lambda'_l)^2} - 1 \right) \langle \phi'_k | F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \\
& \quad - 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda_l \left(\frac{(\lambda_k - \lambda_l)^2}{(\lambda_k + \lambda_l)^2} - 1 \right) \langle \phi_k | F^u | \phi_l \rangle \langle \phi_l | F^v | \phi_k \rangle \\
& \quad + O(\gamma^4) \\
& \stackrel{(c)}{=} 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda'_l \left(\frac{(\lambda'_k - \lambda'_l)^2}{(\lambda'_k + \lambda'_l)^2} - 1 \right) \langle \phi'_k | F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \\
& \quad - 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda_l \left(\frac{(\lambda_k - \lambda_l)^2}{(\lambda_k + \lambda_l)^2} - 1 \right) \langle \phi_k | F^u | \phi_l \rangle \langle \phi_l | F^v | \phi_k \rangle \\
& \quad + 4 \sum_{\substack{l=1, k > 2 \\ k \neq l}} \lambda'_l \left(\frac{(\lambda'_k - \lambda'_l)^2}{(\lambda'_k + \lambda'_l)^2} - 1 \right) \langle \phi'_k | F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \\
& \quad + 4 \sum_{\substack{l=2, k > 2 \\ k \neq l}} \lambda'_l \left(\frac{(\lambda'_k - \lambda'_l)^2}{(\lambda'_k + \lambda'_l)^2} - 1 \right) \langle \phi'_k | F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \\
& \quad + O(\gamma^4) \\
& \stackrel{(d)}{=} 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda'_l \left(\frac{(\lambda'_k - \lambda'_l)^2}{(\lambda'_k + \lambda'_l)^2} - 1 \right) \langle \phi'_k | F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \\
& \quad - 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda_l \left(\frac{(\lambda_k - \lambda_l)^2}{(\lambda_k + \lambda_l)^2} - 1 \right) \langle \phi_k | F^u | \phi_l \rangle \langle \phi_l | F^v | \phi_k \rangle \\
& \quad + 4 \sum_{\substack{k > 2 \\ k \neq l}} \lambda'_1 O(\gamma^4) \langle \phi'_k | F^u | \phi'_1 \rangle \langle \phi'_1 | F^v | \phi'_k \rangle \\
& \quad + 4 \sum_{\substack{k > 2 \\ k \neq l}} \lambda'_2 O(\gamma^2) \langle \phi'_k | F^u | \phi'_2 \rangle \langle \phi'_2 | F^v | \phi'_k \rangle + O(\gamma^4), \quad (\text{C40})
\end{aligned}$$

where (a) follows from (C33) and Lemma 12, (b) follows from (C36) (c) follows from expanding the first summation, and using $\lambda'_l = O(\gamma^4)$ for $l > 2$ from (C16), (d) follows from (C19) and (C20).

Hence we get

$$\begin{aligned}
& (\text{Tr}(L_u \rho_\gamma L_v) - \text{Tr}(L_v \rho_\gamma L_u)) - (\text{Tr}(\bar{L}_u \bar{\rho}_\gamma \bar{L}_v) - \text{Tr}(\bar{L}_v \bar{\rho}_\gamma \bar{L}_u)) \\
& \stackrel{(a)}{=} 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda'_l \left(\frac{(\lambda'_k - \lambda'_l)^2}{(\lambda'_k + \lambda'_l)^2} - 1 \right) \langle \phi'_k | F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \\
& \quad - 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda_l \left(\frac{(\lambda_k - \lambda_l)^2}{(\lambda_k + \lambda_l)^2} - 1 \right) \langle \phi_k | F^u | \phi_l \rangle \langle \phi_l | F^v | \phi_k \rangle \\
& \quad + O(\gamma^4) \\
& \stackrel{(b)}{=} 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda'_l (c_2 \gamma^2 + O(\gamma^4)) \langle \phi'_k | F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \\
& \quad - 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda_l (c_2 \gamma^2 + O(\gamma^4)) \langle \phi_k | F^u | \phi_l \rangle \langle \phi_l | F^v | \phi_k \rangle + O(\gamma^4) \\
& \stackrel{(c)}{=} 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda'_l (c_2 \gamma^2 + O(\gamma^4)) \langle \phi'_k | F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \\
& \quad - 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda'_l (c_2 \gamma^2 + O(\gamma^4)) \langle \phi_k | F^u | \phi_l \rangle \langle \phi_l | F^v | \phi_k \rangle \\
& \quad + 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda'_l (c_2 \gamma^2 + O(\gamma^4)) \langle \phi'_k | F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \\
& \quad - 4 \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda_l (c_2 \gamma^2 + O(\gamma^4)) \langle \phi_k | F^u | \phi_l \rangle \langle \phi_l | F^v | \phi_k \rangle + O(\gamma^4) \\
& \stackrel{(d)}{=} \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} \lambda'_l \gamma^2 O(\gamma^2) + \sum_{\substack{1 \leq k, l \leq 2 \\ k \neq l}} O(\gamma^2) \gamma^2 \langle \phi'_k | F^u | \phi'_l \rangle \langle \phi'_l | F^v | \phi'_k \rangle \\
& \stackrel{(e)}{=} O(\gamma^4),
\end{aligned}$$

where (a) is because the last two summations in (C40) are $O(\gamma^4)$, (b) is because of (C18) and (C17), (c) is because of a telescoping sum, (d) follows from collecting terms in the telescoping sum, and using (C37) in the first pair of summations and (C16) in the second pair of summations and collecting constants into the big-O notation, and (e) follows from addition in the big-O notation. Hence, the result follows.

Appendix D: Semidefinite program formulation with CVX

Now, we will formulate the mathematical optimization problems that correspond to J_2, J_3, J_4, J_5 as semidefinite programs to be used with the CVX package and provide the corresponding MatLab code. These formulations allow us to numerically evaluate the optimal values and solutions of J_2, J_3, J_4, J_5 .

For the numerical formulations, we consider probe states of finite dimension d_A , and quantum channels that map finite dimensional probe states of dimension d_A to states of finite dimension d_B . Then, for these optimization programs that correspond to J_2, J_3, J_4, J_5 , the optimization

variable is Y , which is a complex semidefinite matrix of size $(d+1)d_B d_A$. Hence, in our MatLab script, we define the dimension of Y as

$$dY = (d+1)*dB*dA;$$

The matrix Y has support on the tensor product space $\mathcal{R} \otimes \mathcal{H}_B \otimes \mathcal{H}_A$, and we like to interpret Y as a block matrix with blocks $Y_{j,k}$, where $j, k = 0, \dots, d$, and $Y_{j,k}$ are matrices on $\mathcal{H}_B \otimes \mathcal{H}_A$.

The optimization programs J_k take as input the size d weight matrix G , the channel's Choi matrix T , and the partial derivatives of the Choi Matrix. It is also convenient to specify the input and output dimensions of the quantum channel. Hence we can write the first line of the optimization programs J_2, \dots, J_5 respectively as

```
function [opt_val, opt_Y]=J2(G,T,DT,dA,dB)
function [opt_val, opt_Y]=J3(G,T,DT,dA,dB)
function [opt_val, opt_Y]=J4(G,T,DT,dA,dB)
function [opt_val, opt_Y]=J5(G,T,DT,dA,dB).
```

Here, DT is a cell of derivatives of T , and we can access the j th derivative of T using $DT\{j\}$. The channel's Choi matrix T is a matrix on the space $\mathcal{H}_B \otimes \mathcal{H}_A$. Each of the functions J_2, J_3, J_4 , and J_5 will find the optimal value and optimal solutions of the optimization problem that corresponds to J_k , given by opt_val and opt_Y respectively.

We can access the submatrices $Y_{j,k}$ of Y using the following Matlab function:

```
function Yjk = jk_blk(Y,j,k,d)
dT=size(Y,1)/d;
idx_row_start = j*dT+1;
idx_row_end = j*dT+dT;
idx_col_start = k*dT+1;
idx_col_end = k*dT+dT;
rows=idx_row_start:idx_row_end;
cols=idx_col_start:idx_col_end;
Yjk = Y(rows,cols);
end
```

The objective function to be minimized is $\text{Tr}Y(G \otimes T)$. We can write this objective function in MatLab code by first defining the variables

```
Gprime = [zeros(1,d+1); zeros(d,1),G];
GT = kron(Gprime, T);
```

and minimizing $\text{trace}(Y*GT)$.

The constraint that Y belongs to the cone \mathcal{B}'' , which corresponds to $Y_{k,0}$ being Hermitian for all $k = 0, \dots, d$ and

$Y_{k,j} = Y_{j,k}^\dagger$ for all $j, k = 0, \dots, d$ can be written as

```

for j=0:d
    jk_blk(Y,j,0,d) == jk_blk(Y,j,0,d)'
    for k=0:d
        jk_blk(Y,j,k,d) == jk_blk(Y,k,j,d)'
    end
end
(D1)

```

Now we will use the variable `vecb` to represent the matrix $|0\rangle \otimes |b\rangle \otimes I_A$. We use $IR = \text{eye}(d+1)$ to represent I_R , $zket_R = IR(:, 1)$ to represent $|0\rangle$, $IA = \text{eye}(dA)$ to represent I_A , $IB = \text{eye}(dB)$ to represent I_B , and $IB(:, b)$ to represent $|b\rangle$. Then the constraint for (i') can be written as

```

for b = 1:dB
    vecb=kron(kron(zket_R, IB(:,b)), IA);
    for bp = 1:dB
        vecbp=kron(kron(zket_R, IB(:,bp)), IA );
        if b == bp
            trace(vecb'*Y*vecb) == 1
        else
            vecbp'*Y*vecb == zeros(dA)
            vecb'*Y*vecb == vecbp'*Y*vecbp
        end
    end
end
(D2)

```

The constraint for (ii), which is the channel analogue of the locally unbiased constraint, can be written as

```

for j = 1:d
    for jp = 1:d
        jp_ket_R = IR(:,jp+1);
        zjp = zket_R*jp_ket_R';
        if j == jp
            0.5*trace(Y*kron(zjp + zjp', DT{j})) == 1
        else
            0.5*trace(Y*kron(zjp + zjp', DT{j})) == 0
        end
    end
end
(D3)

```

Now we will proceed to write J_5 as a semidefinite program that can be run using the CVX package in MatLab. For J_5 , the cone S_{AB}^5 requires the constraints $Y \geq 0$, $\text{Tr}Y(|j\rangle\langle i| \otimes$

$T) = 0$ and that $Y \in \mathcal{B}''$. Hence we can write J_5 as

```

cvx_begin sdp %J5
    variable Y(dY,dY) complex semidefinite
    minimize(trace(Y*GT))
    subject to
    Insert code for (D2)
    Insert code for (D3)
    Insert code for (D1)
    for j = 1:d
        j_ket_R = IR(:,j+1);
        for jp = 1:d
            jp_ket_R = IR(:,jp+1);
            j_jp = j_ket_R*jp_ket_R';
            imag(trace(Y*kron(j_jp,T))) == 0
        end
    end
cvx_end
(D4)

```

For J_4 , we can write the optimization as

```

cvx_begin sdp %J4
    variable Y(dY,dY) complex semidefinite
    minimize(trace(Y*GT))
    subject to
    Insert code for (D2)
    Insert code for (D3)
    Insert code for (D1)
cvx_end
(D5)

```

For J_3 , we will need a positive partial transpose constraint on Y . We can write the partial transpose of Y as Y_PT where

```

Y_PT = zeros((d+1)*dA*dB)
for j = 0:d
    idx_row_start = j*dT+1;
    idx_row_end = j*dT+dT;
    rows=idx_row_start:idx_row_end
    for k = 0:d
        idx_col_start = k*dT+1;
        idx_col_end = k*dT+dT;
        cols=idx_col_start:idx_col_end
        Y_PT(rows,cols)=jk_blk(Y,j,k,d);
    end
end
(D6)

```

and hence we can write J_3 as

```
cvx_begin sdp %J3
    variable Y(dY,dY) complex semidefinite
    minimize(trace(Y*GT))
    subject to
    Insert code for (D2)
    Insert code for (D3)
    Insert code for (D1)
    Insert code for (D6)
    Y_PT>=0
cvx_end
```

For J_2 , we need to optimize over the cone \mathcal{B} instead of \mathcal{B}'' .

The constraints for the cone \mathcal{B} can be written as

```
for j=0:d
    for k=0:d
        jk_blk(Y,j,k,d) == jk_blk(Y,k,j,d)
        jk_blk(Y,j,k,d) == jk_blk(Y,j,k,d)'
    end
end (D7)
```

Hence we can write J_2 as

```
cvx_begin sdp %J2
    variable Y(dY,dY) complex semidefinite
    minimize(trace(Y*GT))
    subject to
    Insert code for (D2)
    Insert code for (D3)
    Insert code for (D7)
cvx_end
```

For the optimization problems that correspond to J_2, \dots, J_5 , once we find the optimal Y^* , the corresponding optimal probe state ρ_A^* can be written as

$$\rho_A^* := \text{Tr}_{RB}[Y(|0\rangle\langle 0| \otimes I_{AB})]/d_B. \quad (\text{D8})$$