# The Cramér-Rao approach and global quantum estimation of bosonic states

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 AcademyFutian District, Shenzhen 518048, China
[3]Graduate School of Mathematics, Nagoya University, Nagoya, 464-8602, Japan
[4]School of Mathematical and Physical Sciences, University of Sheffield, Sheffield, S3 7RH, United Kingdom

Quantum state estimation is a fundamental task in quantum information theory, where one estimates real parameters continuously embedded in a family of quantum states. In the theory of quantum state estimation, the widely used Cramér Rao approach which considers local estimation gives the ultimate precision bound of quantum state estimation in terms of the quantum Fisher information. However practical scenarios need not offer much prior information about the parameters to be estimated, and the local estimation setting need not apply. In general, it is unclear whether the Cramér-Rao approach is applicable for global estimation instead of local estimation. In this paper, we find situations where the Cramér-Rao approach does and does not work for quantum state estimation problems involving a family of bosonic states in a non-IID setting, where we only use one copy of the bosonic quantum state in the large number of bosons setting. Our result highlights the importance of caution when using the results of the Cramér-Rao approach to extrapolate to the global estimation setting.

## I. INTRODUCTION

Quantum sensors promise to estimate parameters with unprecedented precision, and are based on a mathematical primitive known as quantum state estimation. In quantum state estimation, the task is to estimate physical parameters embedded within quantum states with minimal error. The Cramér-Rao approach [1–6], a prevalent technique in quantum state estimation which provides lower bounds on the minimum mean square error (MSE) of the estimate. Such lower bounds, known as Cramér-Rao bounds, use Fisher information obtained from quantum measurements. Since the Fisher information captures only the local structure of a statistical model, Cramér-Rao bounds are best suited for local estimation problems, where one assumes that the unknown parameter is within a small neighborhood of a known value. In multiparameter quantum state estimation, the Cramér-Rao approach is more complicated than in the single parameter case; unlike the single-parameter case, where the Cramér-Rao bound is simply the inverse of the quantum Fisher information [1], the multiparameter situation requires additional nontrivial techniques [6].

Despite the prevalence of Cramér-Rao bounds in quantum state estimation theory, one should take note that they are designed for the local estimation setting, and not the global estimation setting where neither the location nor the size of the parameters' neighborhood are known [7–13]. Even with quantum state estimation problems use independent and identically distributed (IID) parametrized quantum states, the MSE for global estimation settings and the MSE for local estimation settings can differ. For example, in Ref. [7] which investigates the canonical phase estimation problem for qubit systems, it was shown that the MSE of the global estimation problem is $\pi^2$ times larger than the MSE of the local estimation problem. Similar observations were made in [12, Section 4], [13, Section 5]. Ref. [14] investigates the phase estimation problem using two bosonic modes, and similarly finds that the NOON state fails to saturate the Cramér-Rao in the global estimation setting while several studies [15–18] implemented the NOON state for this aim according to the proposal [19–21].

An even larger disparity between the MSE of global estimation and the MSE of local estimation problems has been observed [22–26], even in some classical parameter estimation problems that pertain to estimating entropy and Rényi entropy in classical systems in the IID setting. In the IID setting of estimating quantum entropy and quantum Renyi entropies of quantum states, [25, 26] similarly found that quantum Cramér-Rao bounds are not accurate for a global estimation strategy [1].

Given the possible disparity between the MSE between global and local estimation problems in the IID setting, one might not expect Cramér-Rao bounds to be accurate in a more complicated non-IID setting. However, it was surprisingly found that in the context of classical estimation theory that the non-IID situation globally estimating parametrised classical Markovian processes and classical hidden Markovian processes with a finite state system have MSE accurately described by Cramér-Rao bounds [30, 31].

Hence, it is non-trivial to determine whether a Cramér-Rao bound is accurate for global estimation problems, even in the IID setting. While the attainability of the Cramér-Rao

* masahito@math.nagoya-u.ac.jp
† y.ouyang@sheffield.ac.uk

[1] This is because the Cramér-Rao bound is equal to the varentropy [27, 28], whose maximum value is upper bounded by $(\log d)^2$ [29, Lemma 8]. If the Cramér-Rao approach were to accurately describe the optimal performance of global estimation strategies with a constant error, then the number of copies of states needed (sample complexity) is $O(\log d)^2$ for a $d$-dimensional system. However, estimating both the classical and quantum entropy to a constant error requires much larger complexity, even in the classical setting [22–26].

bound has been studied in the one-copy setting [32], the attainability of Cramér-Rao bound for other non-IID settings has been less studied.

In this paper, we study the accuracy of Cramér-Rao bounds for the global estimation of bosonic quantum states in a very non-IID setting, where we only have one copy of a parametrised bosonic quantum state, and consider the limit where the number of bosons becomes very large. Bosonic quantum states are ubiquitous, because any fundamental particle in the universe is either a boson or a fermion. Furthermore, we can realize a boson as a composite particle, comprising of an even number of fermions and number of bosons. In the mathematical framework of second quantization, bosons are naturally represented in the Fock basis, where basis elements count the occupancy of bosons in the available modes. For indistinguishable bosons, the corresponding quantum state resides within a space spanned that is invariant under any permutation of the underlying bosons, and are hence symmetric. We can realize such bosons in various physical systems, such as Bose-Einstein condensates (BECs) in cold atomic systems [33].

Mathematically, bosonic states have the same structure as symmetric states. In the context of local estimation theory, symmetric states promise a quantum advantage in certain quantum metrology problems in the noiseless setting. The additional simplicity of the preparation and control of symmetric states [34, 35] makes symmetric states attractive candidates to demonstrate the near-term advantage of quantum technologies. However in a practical setting, we may not have the requisite prior information about the parameters embedded within these symmetric states that are to be estimated, and require a global estimation strategy. Given this, it is pertinent to understand the applicability of the Cramér-Rao approach, particularly for the symmetric states for which it is purported that a quantum advantage might be available.

First, we focus on families of bosonic states that are diagonal in the number basis. Namely, we consider bosonic states that are probabilistic mixtures of states with a fixed number of bosons with the following probability distributions; (1) a geometric distribution, (2) a binomial distribution, and (3) a delta distribution. The embedded parameter describes the probability distribution. In this case, we show that the MSE from the Cramér-Rao approach is equal to the MSE in the global estimation setting.

Second, we proceed to families generated by unitary evolutions of SU(2) over probe states that begin in the number basis. Unlike many previous studies [7–11, 36–46] that focus on estimating mutiple phases in a similar setting, we specifically address the less studied scenario of the estimation of the state family generated by unitary evolutions of SU(2). The probe states can be in a (1) binomial distribution, (2) geometric distribution, and a (3) delta distribution in the number basis. For this, we consider the problem of estimating parameters embedded in a unitary model, where the unitary channel acts on the probe state. We analyze the global estimation of this problem by drawing an analogue between the bosonic system and the SU(2) system with a spin-$j$ system, and we employ the covariant approach initiated by Holevo [47], [2, Chapter 3]. This covariant approach allows us to solve a minimax problem with group symmetry [47], [2, Chapter 4], [48, Chapter 4]. Since we may describe global estimation using an appropriate minimax problem, the covariant approach works for the global estimation. By calculating both local and global estimation bounds, we are able to determine if the Cramér-Rao approach is accurate for the global estimation of our unitary model.

The remaining part of this paper is organized as follows. First in Section II, we explain how bosonic states may arise in practice, and one may prepare the families of bosonic states that we consider in our paper. Second in Section III, we review the general formulation of quantum state estimation based on the Cramér-Rao approach. Third in Section IV, we discuss the attainability of the Cramér-Rao bound in the global estimation setting with respect to several quantum state estimation problems. Fourth in Section V, we discuss the attainability and the unattainability of the Cramér-Rao bound in the global estimation setting with respect to several quantum state estimation problems under a unitary model. Finally in Section VI, we have a final discussion of the results that we obtain.

## II. BOSONIC STATES AND THEIR PREPARATION

### A. Boson Fock space and geometric distribution

There are physical systems where we may realize indistinguishable identical bosons. For instance, ultracold neutral atoms, when sufficiently cooled and confined, can become indistinguishable, and hence are fundamentally bosonic states. We can interpret neutral atoms using their total spin or electronic states as internal degrees of freedom as bosonic states. Similarly, we can interpret photons that are indistinguishable in all aspects except for their polarizations as bosonic states.

An example of a bosonic system that is controllable in the near term with a large number of bosons is a system of ultracold neutral atoms. Neutral atoms can be realised as bosons if we interpret each neutral atom as a composite particle with an equal number of protons and electrons and an even number of neutrons. Almost every neutral atom has an isotope that is a boson. Examples of neutral atoms that are bosons include group I elements such as Li-7, Na-23 and Rb-87. BECs of such indistinguishable identical neutral atoms are now routinely realized in experiments, with the number of bosons being as large as $10^{10}$ [33]. For ultracold neutral atoms, the internal degrees of freedom can for instance correspond to the total spin of each atom, which can take on two accessible values. For photons, the internal degrees of freedom can correspond to their horizontal and vertical polarizations.

Our paper considers the quantum state estimation problem for a system of $n$ identical and indistinguishable bosons. We model the bosonic system with $d$ kinds of distinguishable modes as the $d$-mode bosonic Fock space $\mathcal{H}_{B,d}$, which is spanned by the basis $\{|n_1,n_2,\ldots,n_d\rangle_B : n_1,\ldots,n_d \geq 0\}$. The space $\mathcal{H}_{B,d}$ is written as the tensor product space $\mathcal{H}_B^{\otimes d}$, where $\mathcal{H}_B$ is the one-mode bosonic Fock space spanned by the basis $\{|n\rangle_B\}_{n=0}^{\infty}$. The $d$-mode Fock space can also decomposed as

$$\mathcal{H}_{B,d} = \bigoplus_{n=0}^{\infty} \mathcal{H}_{B,d,n}, \tag{1}$$

where $\mathcal{H}_{B,d,n}$ are Fock spaces with a total of $n$ bosons in $d$ modes spanned by

$$\mathcal{B}_n := \{|n_1,n_2,\ldots,n_d\rangle_B : \sum_{k=1}^{d} n_k = n, n_k \geq 0\}. \tag{2}$$

We may interpret the spaces $\mathcal{H}_{B,d,n}$ as constant excitation spaces with $n$ excitations [49], which are eigenspaces of Hamiltonians that are sums of independent and identical single-mode operators diagonal in the Fock basis. The space $\mathcal{H}_{B,d,n}$ is also isomorphic to the symmetric subspace in $n$-fold tensor product space of the $d$-dimensional space. Since automorphisms on symmetric space can be described using the group $\mathrm{SU}(d)$, we can also use $\mathrm{SU}(d)$ to describe automorphisms on $\mathcal{H}_{B,d,n}$.

Hereafter, we focus on the case with $d = 2$, which corresponds to bosons with two internal degrees of freedom. Denoting the spin-$\frac{n}{2}$ space as $\mathcal{H}_{\frac{n}{2}}$, we may decompose the space $\mathcal{H}_{B,d,n}$ as a direct sum of spin-spaces given by

$$\mathcal{H}_{B,2} = \bigoplus_{n=0}^{\infty} \mathcal{H}_{\frac{n}{2}}. \tag{3}$$

The spin-$j$ space is spanned by $\{|j;m\rangle\}_{m=-j}^{j}$, and its automorphisms have the symmetry of the group $\mathrm{SU}(2)$, where the operators $J_1, J_2$ and $J_3$ satisfy the commutation relations $[J_i, J_j] = \sqrt{-1}\epsilon_{i,j,k}J_k$ and form the Lie algebra of $\mathrm{SU}(2)$, with $\epsilon_{i,j,k}$ denoting the Levi-Civita symbol, and $J_3$ being a diagonal operator in the Fock basis. Using this idea, we can identify the vector $|n-k,k\rangle_B$ in the boson Fock space with the vector $|\frac{n}{2}; k-\frac{n}{2}\rangle$ in the spin-$\frac{n}{2}$ space.

We may realize the geometric distribution on two-mode Fock states in the number basis with a total of $n$ bosons by starting from thermal states of a two-mode Hamiltonian given by $G_{\alpha_1,\alpha_2} := \alpha_1 N_1 + \alpha_2 N_2$, where $N_j$ is the number operator on the $j$-th mode. This Hamiltonian $G_{\alpha_1,\alpha_2}$ represents the sum of two independent single-mode Hamiltonians in the Fock basis where the energy properties of the two modes can be different. The thermal state that corresponds to the Hamiltonian $G_{\alpha_1,\alpha_2}$ at the inverse temperature $\beta$ is given as $c\exp(-\beta G_{\alpha_1,\alpha_2})$ for some normalizing constant $c$, which we can write as

$$\rho_{\mathrm{G},\alpha_1,\alpha_2,\beta} = c \sum_{n_1,n_2\geq 0} e^{-\beta(\alpha_1 n_1 + \alpha_2 n_2)} |n_1,n_2\rangle_{\mathcal{B}}\langle n_1,n_2|_{\mathcal{B}}. \tag{4}$$

After we measure the total number of bosons and observe $n$ bosons, the state becomes

$$\begin{aligned}\rho_{\mathrm{G},r}^{(n)} &= c' \sum_{k=0}^{n} e^{-\beta(\alpha_1(n-k)+\alpha_2 k)} |n-k,k\rangle_{\mathcal{B}}\langle n-k,k|_{\mathcal{B}} \\ &= c'' \sum_{k=0}^{n} r^k |n-k,k\rangle_{\mathcal{B}}\langle n-k,k|_{\mathcal{B}}\end{aligned} \tag{5}$$

for other some normalization constants $c'$ and $c''$, where $r = e^{-\beta(\alpha_1-\alpha_2)}$. The state $\rho_{\mathrm{G},r}^{(n)}$ is a geometric distribution in the number of bosons in the second mode, with geometric ratio given by $r$.

Next, we consider the case when a beam splitter operator applies across the two modes. Since the beam splitter operator corresponds to an element of $g \in \mathrm{SU}(2)$, we can consider the state estimation for the state family $\{U_{\frac{n}{2},g}\rho_{\mathrm{G},r}^{(n)}U_{\frac{n}{2},g}^{\dagger} : g \in SU(2)\}$ on the spin-$\frac{n}{2}$ space $\mathcal{H}_{\frac{n}{2}}$, where $U_{\frac{n}{2},g}$ denotes a unitary representation of $g$ on the space $\mathcal{H}_{\frac{n}{2}}$.

Since $J_3$ is a diagonal operator in the Fock basis, it leaves the state $\rho_{\mathrm{G},r}^{(n)}$ invariant. Then we may identify the parameter space as the homogeneous space $\mathrm{SU}(2)/\mathrm{U}(1)$, where $\mathrm{U}(1)$ is the one-parameter group generated by $J_3$. We consider estimating the group parameter $[g] \in \mathrm{SU}(2)/\mathrm{U}(1)$ under the state family $\{U_{\frac{n}{2},g}\rho_{\mathrm{G},r}^{(n)}U_{\frac{n}{2},g}^{\dagger} : [g] \in \mathrm{SU}(2)/\mathrm{U}(1)\}$, which amounts to estimating two real-valued parameters.

### B. Symmetric space and binomial distribution

Since a spin-$\frac{n}{2}$ system is mathematically equivalent to the symmetric subspace of an $n$-qubit system, we can represent a bosonic state mathematically as a symmetric state on $n$ qubits. Manipulating symmetric states is achievable in the near-term, because the requisite quantum control techniques do not require the individual addressability of individual qubits [34]. By leveraging on existing experimental know-how both in creating BECs and controlling large numbers of identical indistinguishable neutral atoms [33] and controlling photonic systems [50], conducting actual quantum sensing experiments on such symmetric states is a near-term possibility.

In this scenario, we can consider another distribution instead of the geometric distribution. That is, we discuss how to prepare the following state over the spin-$j$ system

$$\rho = \sum_{m=-j}^{j} p_m |j;m\rangle\langle j;m|, \tag{6}$$

where (1) $p_k$ follows a binomial distribution and (2) $p_k$ follows a delta distribution.

On the symmetric space, the operators $J_1$, $J_2$ and $J_3$ are angular momentum operators that map symmetric states to symmetric states. In terms of the Pauli operators $\sigma_1, \sigma_2$ and $\sigma_3$, we can write the angular momentum operator $J_j$ as

$$J_j = \frac{1}{2}(\sigma_j^{(1)} + \cdots + \sigma_j^{(n)}), \tag{7}$$

where $\sigma_j^{(k)}$ denotes Pauli operator $\sigma_j$ on the $k$-th particle.

One may prepare a quantum state $\rho$ with a binomial distribution of states in the basis $\{|j;m\rangle\}_{m=-j}^{j}$, that is where $p_m = \binom{n}{\frac{n}{2}+m} p^{\frac{n}{2}-m}(1-p)^{\frac{n}{2}+m}$, according to the following procedure. First, one prepares the initial separable state

$$|\psi_p\rangle = \left(\sqrt{1-p}|\frac{1}{2};-\frac{1}{2}\rangle + \sqrt{p}|\frac{1}{2};\frac{1}{2}\rangle\right)^{\otimes n}. \tag{8}$$

Second, one dephases the pure state $|\psi_p\rangle$ in the basis $\mathcal{B}_n$ using the master equation $d\tau/dt = \mathcal{D}(\tau)$, where

$$\mathcal{D}(\rho) = \gamma(J_3\rho J_3^\dagger - \frac{1}{2}J_3^\dagger J_3\rho - \frac{1}{2}\rho J_3^\dagger J_3). \tag{9}$$

Since we have

$$|\psi_p\rangle = \sum_{m=-\frac{n}{2}}^{\frac{n}{2}} \sqrt{p^{\frac{n}{2}-m}(1-p)^{\frac{n}{2}-m}}\sqrt{\binom{n}{\frac{n}{2}+m}}|\frac{n}{2};m\rangle, \tag{10}$$

complete dephasing of the state $|\psi_p\rangle$ in the eigenbasis of $J_3$ will yield a binomial distribution of states in the basis $\{|j;m\rangle\}_{m=-j}^{j}$. In particular, we have $\lim_{t\to\infty} e^{\mathcal{D}t}(|\psi_p\rangle\langle\psi_p|) = \rho_{\mathrm{B},p}^{(n)}$ where

$$\rho_{\mathrm{B},p}^{(n)} := \sum_{m=-\frac{n}{2}}^{\frac{n}{2}} \binom{n}{\frac{n}{2}+m} p^{\frac{n}{2}-m}(1-p)^{\frac{n}{2}+m}|\frac{n}{2};m\rangle\langle\frac{n}{2};m|, \tag{11}$$

and $e^{\mathcal{D}t} = \mathcal{I} + \sum_{k=1}^{\infty}\frac{t^k}{k!}\mathcal{D}^k$, and $\mathcal{I}$ denotes the identity operator. Hence, one can apply $e^{t\mathcal{D}}$ on $|\psi_p\rangle$ for large $t$ to approximately obtain a binomial distribution on states in the basis $\{|j;m\rangle\}_{m=-j}^{j}$.

Once an unknown unitary $U_{\frac{n}{2},g}$ with $g \in \mathrm{SU}(2)$ is applied, in the same way as with the geometric distribution, we can consider the state family $\{U_{\frac{n}{2},g}\rho_{\mathrm{B},p}^{(n)}U_{\frac{n}{2},g}^\dagger : [g] \in \mathrm{SU}(2)/U(1)\}$.

### C. Delta distribution

A state with the delta distribution in (6) can be prepared as follows. There are probabilistic approaches to prepare a specific state in the number basis. For the probabilistic approach, one can prepare a binomial or geometric distribution of states in the basis $\{|j;m\rangle\}_{m=-j}^{j}$, and subsequently measure in the basis $\{|j;m\rangle\}_{m=-j}^{j}$. For the deterministic approach, one can use an ancillary bosonic mode along with a dispersive interaction Hamiltonian that is proportional to $a^\dagger a \otimes J_3$ to implement unitary operations in the spin-$j$ system using geometric phase gates [34, 51, 52].

## III. GENERAL FORMULATION OF CRAMÉR-RAO APPROACH

In quantum state estimation, we are given copies of an unknown state $\rho_{\theta_0}$ from the set of quantum states $\{\rho_\theta : \theta = (\theta^1,\dots,\theta^d) \in \Theta\}$ where $\Theta$ is a continuous set in $\mathbb{R}^d$. We assume that the quantum states $\rho_\theta$ are differentiable with respect to parameter $\theta$ for all $\theta \in \Theta$. Our objective is to find the minimum MSE of a locally unbiased estimator $\hat{\theta}$ that estimates the true parameter $\theta_0$.

We describe a measurement using a set of positive operators $\Pi = \{\Pi_x : x \in \mathcal{X}\}$ labeled by a set $\mathcal{X}$, where the completeness condition $\sum_{x\in\mathcal{X}}\Pi_x = I$ holds. By Born's rule, a measurement $\Pi$ on a quantum state $\rho_\theta$ gives the classical label $x$ and the state $\Pi_x\rho_\theta/\mathrm{Tr}(\Pi_x\rho_\theta)$ with probability $p_\theta(x) = \mathrm{Tr}(\Pi_x\rho_\theta)$. Given a function $f$ of the classical label $x$, we denote $\mathbb{E}[f(x)|\Pi]$ as the expectation of $f(x)$, with probability distribution obtained according to Born's rule.

Given a measurement $\Pi$ and an estimator $\hat{\theta}$ that depends on the classical label $x$, we denote $\hat{\Pi} = (\Pi, \hat{\theta})$ as an *estimator*. When the true parameter $\theta_0$ is equal to $\theta$, we define the mean-square error (MSE) matrix for the estimator $\hat{\Pi}$ as

$$V_\theta[\hat{\Pi}] = \sum_{i,j=1}^{d} |i\rangle\langle j|\mathbb{E}_\theta\big[(\hat{\theta}^i(x)-\theta^i)(\hat{\theta}^j(x)-\theta^j)|\Pi\big].$$

In multiparameter quantum metrology, the objective is to find an optimal estimator $\hat{\Pi} = (\Pi, \hat{\theta})$ that minimizes $\mathrm{Tr}GV_\theta[\hat{\Pi}]$, where a weight matrix $G$, a size $d$ positive semidefinite matrix, quantifies the relative importance of the different parameters.

Our estimator $\hat{\Pi}$ is unbiased at $\theta_0 = \theta$ if for all $i = 1,\dots,d$, the expectation of our estimator equals the true value of the parameter $\theta_0$ when $\theta_0 = \theta$, that is

$$\mathbb{E}_\theta\big[\hat{\theta}^i(x)|\Pi\big] = \sum_{x\in\mathcal{X}}\hat{\theta}^i(x)\mathrm{Tr}\big[\rho_\theta\Pi_x\big] = \theta^i. \tag{12}$$

Our estimator is globally unbiased if (12) holds for all $\theta \in \Theta$. We can also consider locally unbiased estimators, which are estimators that are unbiased in the neighborhood of the true parameter $\theta_0$. For this aim, we define $D_j := \frac{\partial\rho_\theta}{\partial\theta^j}|_{\theta=\theta_0}$, and $\rho := \rho_{\theta_0}$. Taking partial derivatives on both sides of (12), we get

$$\frac{\partial}{\partial\theta^j}\mathbb{E}_\theta\big[\hat{\theta}^i(x)|\Pi\big] = \sum_{x\in\mathcal{X}}\hat{\theta}^i(x)\mathrm{Tr}D_j\Pi_x = \delta_i^j. \tag{13}$$

The estimator $\hat{\Pi}$ is locally unbiased if (12) holds for all $i = 1,\dots,d$ for a fixed $\theta$ where $\theta_0 = \theta$, and when (13) holds for all $i,j = 1,\dots,d$.

For any weight matrix $G = \sum_{i,j=1}^{d} g_{i,j}|i\rangle\langle j|$, the tight Cramér-Rao (CR) type bound, i.e., the fundamental precision limit [6], is

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

where 'l.u. at $\theta$' indicates our minimization over all possible estimators under the locally unbiasedness condition. Since this minimum is attained by $\hat{\Pi}$ satisfying (12) when we impose only the condition (13), it suffices to consider $C_\theta[G]$ as a minimization with only the condition (13).

To evaluate $C_\theta[G]$, we often focus on the symmetric logarithm derivative (SLD) $L_j$, which is an operator that satisfies the equation

$$D_j = \frac{1}{2}(L_j\rho + \rho L_j). \tag{14}$$

The SLD Fisher information matrix $F = (F_{i,j})$ is given as

$$F_{i,j} := \frac{1}{2}\mathrm{Tr} L_i(L_j\rho + \rho L_j). \tag{15}$$

The tight CR bound $C_\theta[G]$ can be lower bounded as follows

$$C_\theta[G] \ge C_\theta^{\mathrm{S}}[G] := \mathrm{Tr} GF^{-1}. \tag{16}$$

The RHS of (16) is called the SLD bound.

In the one-parameter case, we do not need to handle the trade-off among various parameters. In this case, the equality in (16) holds. We can attain this bound using a projective measurement in the eigenbasis of the SLD $L$. Hence, in the multiple-parameter case, when the SLDs $L_j$ are noncommutative, their spectral decompositions cannot be measured simultaneously. However, it is possible to randomly choose one of the SLDs $L_j$ and measure it, as was studied in [53]. To discuss a simple case of this strategy, we assume that the SLD Fisher information matrix $J$ has no off-diagonal element. The tight CR bound $C_\theta[G]$ can be evaluated simply as follows [1].

$$C_\theta[G] \le d\,\mathrm{Tr} GF^{-1}. \tag{17}$$

We attain the SLD bound by measuring in the eigenbasis of the SLD $L_j$ with equal probability for $j = 1, \ldots, d$. Thus, when $d = 2$, the SLD bound decides $C_\theta[G]$ within twice the range.

To get a better lower bound, we often focus on the right logarithm derivative (RLD) $\tilde{L}_j$, which is an operator that solves the equation

$$D_j = \rho \tilde{L}_j. \tag{18}$$

The RLD Fisher information matrix $\tilde{F} = (\tilde{J}_{i,j})$ is given as

$$\tilde{F}_{i,j} := \mathrm{Tr}\tilde{L}_i \rho \tilde{L}_j. \tag{19}$$

The tight CR bound $C_\theta[G]$ can be lower bounded as follows.

$$C_\theta[G] \ge C_\theta^{\mathrm{R}}[G] := \mathrm{Tr}\mathrm{Re}\,\sqrt{G}\tilde{F}^{-1}\sqrt{G} + \mathrm{Tr}|\mathrm{Im}\,\sqrt{G}\tilde{F}^{-1}\sqrt{G}|. \tag{20}$$

The RHS of (20) is called the RLD bound [2, Chapter 6]. To consider this bound, we define the operator $D$ as

$$[\rho, X] = \frac{1}{2}(\rho D(X) + D(X)\rho). \tag{21}$$

We say the D-invariant condition holds if $D(L_j)$ is in the linear span of $L_1, \ldots, L_d$. We define matrix $D = (D_{j,k})$ as $D_{j,k} := \mathrm{Tr} D(L_j) D_k$. In this case, the $\tilde{F}^{-1}$ is calculated as [2, Chapter 6]

$$\tilde{F}^{-1} = F^{-1} + \frac{i}{2}F^{-1}DF^{-1}. \tag{22}$$

Then, the RLD bound is calculated as [2, Chapter 6]

$$C_\theta^{\mathrm{R}}[G] = \mathrm{Tr} GF^{-1} + \frac{1}{2}\mathrm{Tr}|\sqrt{G}F^{-1}DF^{-1}\sqrt{G}|. \tag{23}$$

and is a better lower bound than the SLD bound. Since (22) implies

$$\tilde{F}^{-1} \le 2F^{-1}, \tag{24}$$

we have

$$\mathrm{Tr} GF^{-1} \le C_\theta^{\mathrm{R}}[G] \le 2\mathrm{Tr} GF^{-1}. \tag{25}$$

That is, the RLD bound differs from the SLD bound by up to a factor of two for D-invariant models.

As a tighter lower bound, we employ Holevo-Nagaoka (HN) bound as follows [2, 5, 54]. Given a tuple of Hermitian matrices $\vec{X} = (X_1, \ldots, X_d)$, we define the matrix $Z(\vec{X}) = (Z_{j,k}(\vec{X}))$ as

$$Z_{j,k}(\vec{X}) := \mathrm{Tr}\rho X_j Z_k. \tag{26}$$

We impose the following condition to $\vec{X}$;

$$\mathrm{Tr} X_j D_k = \delta_{j,k}. \tag{27}$$

Then, we define

$$C_\theta^{\mathrm{HN}}[G] := \min_{\vec{X}} \mathrm{Tr} G \mathrm{Re}\, Z(\vec{X}) + \mathrm{Tr}|\sqrt{G}\mathrm{Im}\, Z(\vec{X})\sqrt{G}|, \tag{28}$$

where the minimum is taken under the condition (27). Then, we have

$$C_\theta^{\mathrm{HN}}[G] \le C_\theta[G]. \tag{29}$$

Furthermore, we have

$$C_\theta^{\mathrm{R}}[G] \le C_\theta^{\mathrm{HN}}[G] \tag{30}$$
$$C_\theta^{\mathrm{S}}[G] \le C_\theta^{\mathrm{HN}}[G]. \tag{31}$$

When the model is D-invariant, the equality in (30) holds [5].

For example, when we choose $\vec{X}$ as $\vec{X}_* = (X_{k,*})$ with $X_{k,*} := \sum_{j=1}^d (F^{-1})_{k,j} L_j$, $\vec{X}_*$ satisfies the condition (27). Also, we have $\mathrm{Re}\, Z(\vec{X}_*) = F^{-1}$. Since $Z(\vec{X}) \le 2\mathrm{Re}\, Z(\vec{X})$, we have

$$\begin{aligned} C_\theta^{\mathrm{HN}}[G] &\le \mathrm{Tr} G \mathrm{Re}\, Z(\vec{X}_*) + \mathrm{Tr}|\sqrt{G}\mathrm{Im}\, Z(\vec{X}_*)\sqrt{G}| \\ &\le 2\mathrm{Tr} G \mathrm{Re}\, Z(\vec{X}_*) = 2\mathrm{Tr} GF^{-1}. \end{aligned} \tag{32}$$

That is, the HN bound differs from the SLD bound by up to a factor of two for D-invariant models. Hence for D-invariant models, we have good upper and lower bounds on the tight CR-bound based on the easily computable SLD bound.

## IV. ATTAINABILITY OF THE CRAMÉR-RAO BOUND IN THE GLOBAL ESTIMATION SETTING

Here, we give examples where the tight Cramér-Rao bound equals to the minimum MSE for global estimation strategies.

First, we consider the task of estimating the parameter $p$ in the state $\rho^{(n)}_{\mathrm{B},p}$ that is a binomial distribution of states in the basis $\{|n-k,k\rangle_B\}_{k=0}^n$. By measuring in the basis $\{|n-k,k\rangle_B\}_{k=0}^n$, we obtain a binomial distribution, which has a Fisher information of $\frac{2j}{p(1-p)}$. Hence the tight Cramér-Rao bound $C_p[1]$ is $\frac{p(1-p)}{2j}$. We can attain this bound with a global estimator of the parameter $p$ according the following strategy. First we measure this density matrix in the basis $\{|n-k,k\rangle_B\}_{k=0}^n$. Second, if we observe the state $|n-k,k\rangle_B$, we set our estimate as $\frac{k}{n}$. This estimator is unbiased because it has expectation $p$. Moreover, it has MSE $\frac{p(1-p)}{n}$ which attains the tight Cramér-Rao bound.

Second, we consider estimating the parameter $r$ in the state $\rho^{(n)}_{\mathrm{G},r}$ which is a normalized geometric distribution on states in the basis $\mathcal{B}_n$. By measuring in the basis $\mathcal{B}_n$, the estimation problem reduces to estimating a geometric distribution. Now, let us see why the tight Cramér Rao bound in this case is equal to the minimum MSE for the global estimation of $r$.

We begin with the parametrization $P_\theta(k) := \frac{e^\theta-1}{e^{\theta(n+1)}-1}e^{\theta k}$, which is known as the *natural parameter* in the field of information geometry [55]. Since the geometric distribution is an exponential family, the tight CR bound is globally achieved under the *expectation parameter* $\eta(\theta)$ [55], which is defined as $\eta(\theta) := \sum_{k=0}^n kP_\theta(k)$ where we may calculate $\eta(\theta)$ as

$$\eta(\theta) = \frac{ne^{\theta(n+1)}+1}{e^{\theta(n+1)}-1} - \frac{1}{e^\theta-1} = \frac{nr^{n+1}+1}{r^{n+1}-1} - \frac{1}{r-1}. \quad (33)$$

Then, the Fisher information for $\theta$ is $F_\theta := \sum_{k=0}^n k^2 P_\theta(k) - \eta(\theta)^2$ which can be calculated as

$$\begin{aligned}F_\theta =& \frac{n(n-1)e^{\theta(n+1)}+2}{e^{\theta(n+1)}-1} + \frac{e^{\theta(n+1)}-3}{e^{\theta(n+1)}-1}\eta(\theta) - \eta(\theta)^2 \\ =& \frac{n(n-1)r^{n+1}+2}{r^{n+1}-1} + \frac{r^{n+1}-3}{r^{n+1}-1} - \eta(\theta)^2 \\ =& \frac{n(n-1)r^{n+1}(r^{n+1}-1)+2(r^{n+1}-1)}{(r^{n+1}-1)^2} \\ &+ \frac{(r^{n+1}-3)(nr^{n+1}+1)}{(r^{n+1}-1)^2}\eta(\theta) - \eta(\theta)^2. \quad (34)\end{aligned}$$

In this case, when the parameter to be estimated is set to $\eta(\theta)$, the estimator is given as $k$. This estimator satisfies the unbiasedness condition, and its variance is $F_\theta$, i.e., the Fisher information of the natural parameter. We can use this procedure to estimate $r$ globally with MSE that attains the tight CR bound $C_\theta[I]$.

## V. UNATTAINABILITY OF THE CRAMÉR-RAO BOUND IN THE GLOBAL ESTIMATION ESTIMATION SETTING

### A. Local estimation of a unitary channel

We consider the covariant model on symmetric states of $n$ qubits. Using the representation theory of SU(2), we interpret such symmetric states with a spin $j = \frac{n}{2}$ system, wherein it is natural to interpret the number state $|n-k,k\rangle_{\mathcal{B}}$ as a spin state $|j;-j+k\rangle$. We focus on a diagonal state $\rho$ for this basis given as

$$\rho := \sum_{m=-j}^{j} p_m |j;m\rangle\langle j;m|. \quad (35)$$

Then, given a parameter $\theta := (\theta_1,\theta_2)$, we consider the state family $\rho_\theta := U_\theta \rho U_\theta^\dagger$, where $U_\theta := \exp(i(\theta_1 J_1 + \theta_2 J_2))$. The two-parameter space $\Theta$ is given as $\{\theta | 0 \le |\theta| \le \pi\}$. This state family $\{\rho_\theta\}_\theta$ has two parameters, and is obtained from applying unitary operator $U_\theta$ on an initial probe state $\rho$. In this model, the SLD Fisher information is diagonal, in the sense that $F_{1,2} = F_{2,1} = 0$. This implies that $C^{\mathrm{S}}_\theta[I] = F_{1,1}^{-1} + F_{2,2}^{-1}$. Furthermore,

$$F_{1,1} = F_{2,2} = \sum_{m=-j}^{j-1} \frac{4(p_{m+1}-p_m)^2}{p_{m+1}+p_m}(j-m)(j+m+1). \quad (36)$$

From (36), we can apply the Cramér-Rao approach on probe states initialised as (1) a binomial distribution of number states, (2) a geometric distribution of number states, and (3) a delta distribution of number states.

1. **Binomial distribution:-** Now consider the case when $\rho = \rho^{(n)}_{\mathrm{B},p}$. When $p$ is fixed and $j = \frac{n}{2}$ increases, the diagonal element of the SLD Fisher information $F^{(n),\mathrm{B},p}$ can be calculated as

$$F^{(n),\mathrm{B},p}_{1,1} = F^{(n),\mathrm{B},p}_{2,2} \cong 2n. \quad (37)$$

Hence $C^{\mathrm{S}}_\theta[I] \cong 1/n$.

2. **Geometric distribution:-** Consider $\rho$ as $\rho^{(n)}_{\mathrm{G},r}$. From Appendix A, this model satisfies the D-invariant condition. Hence, the RLD bound gives a tighter lower bound than the SLD bound. As calculated in Appendix C, when $r$ is fixed and $j = \frac{n}{2}$ increases, the RLD bound is approximated as

$$C^{\mathrm{R}}_\theta[I] \cong \frac{4r}{n(r-1)}. \quad (38)$$

3. **Delta distribution:-** Consider $p_m = \delta_{m,a}$ for some integer $a \in [-j+1, j-1]$. Then we have

$$F_{1,1} = F_{2,2} = 2(j^2-a^2)+2j. \quad (39)$$

Hence when $a$ is proportional to $n$, both $F_{1,1}$ and $F_{2,2}$ are quadratic in $j$ and $n$. Then we have $C^{\mathrm{S}}_\theta[I] \cong (1/4-\alpha^2)^{-1}/n^2$ where $\alpha = a/n$.

### B. A group covariant approach for global estimation

We consider the state family $\{\rho_\theta\}_\theta$ as given in Section V. Since the state family $\{\rho_\theta\}_\theta$ has a group covariant structure, we can employ a group covariant approach [47], [2, Chapter 4], [48, Chapter 4], where we employ a group covariant error function.

Now we consider the spin $j$ system $\mathcal{H}_j$ spanned by $\{|j;m\rangle\}_{m=-j}^{j}$. For an unknown value of $\theta$, and our estimate $\hat{\theta}$, we can denote the fidelity between the states $U_\theta|\frac{1}{2};\frac{1}{2}\rangle$ and $U_{\hat{\theta}}|\frac{1}{2};\frac{1}{2}\rangle$ as

$$R(\theta,\hat{\theta}) := \mathrm{Tr}\Big(U_\theta|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|U_\theta^\dagger\Big)\Big(U_{\hat{\theta}}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|U_{\hat{\theta}}^\dagger\Big) = |\langle\frac{1}{2};\frac{1}{2}|U_\theta^\dagger U_{\hat{\theta}}|\frac{1}{2};\frac{1}{2}\rangle|^2. \tag{40}$$

When $\theta=0$, this fidelity simplifies to $R(0,\hat{\theta})=\cos^2(\hat{\theta}/2)$. We define the error function of our estimate to be $\eta(\hat{\theta}) := 4(1-R(0,\hat{\theta}))$, which to leading order in $\hat{\theta}$ is equivalent to $|\hat{\theta}|^2$. In this definition, the error function $\eta(\hat{\theta})$ is to leading order in $\hat{\theta}$ equivalent to the MSE of $\hat{\theta}$ in the local estimation scenario when $\theta=0$.

We identify the parameter space with the homogeneous space $\Theta=\mathrm{SU}(2)/\mathrm{U}(1)$, where $\mathrm{U}(1)$ is the one-parameter group generated by $J_3$. Then, we employ the invariant probability measure $\nu$ on our parameter space $\Theta$ under the above identification. The estimator is a POVM $M=\{M(d\hat{\theta}) : \hat{\theta}\in\Theta\}$ with outcomes parametrised according to elements in $\Theta$. Given an estimator $M$, we focus on the Bayesian average

$$R_\nu(M) := \int_\Theta\int_\Theta R(\theta,\hat{\theta})\mathrm{Tr}\rho_\theta M(d\hat{\theta})\nu(d\theta). \tag{41}$$

Also, we can calculate the performance of the global estimation strategy for the worst possible value of the true parameter using the expression

$$R(M) := \min_{\theta\in\Theta}\int_\Theta R(\theta,\hat{\theta})\mathrm{Tr}\rho_\theta M(d\hat{\theta}). \tag{42}$$

Namely, the expression $\eta(M) := 4(1-R(M))$ is our error function maximized over all values of the true parameter $\theta$. Minimizing $\eta(M)$ amounts to solving a minimax problem; we minimize over all POVMs and maximize over all possible true values of $\theta$.

As before, we denote the representation of $g\in\mathrm{SU}(2)$ on a spin-$j$ system $\mathcal{H}_j$ by $U_{j,g}$. We say that the POVM $M$ is covariant if

$$U_{j,g}\int_B M(d\hat{\theta})U_{j,g}^\dagger = \int_{gB} M(d\hat{\theta}) \tag{43}$$

for any subset $B\subset\Theta$ and $g\in\mathrm{SU}(2)$. When a state $T$ satisfies the condition $U_{j,g}TU_{j,g}^\dagger = T$ for $g\in\mathrm{U}(1)$, we define the covariant POVM $M_T$ as

$$M_T(B) := (2j+1)\int_B U_{j,g(\theta)}TU_{j,g(\theta)}^\dagger\nu(d\theta) \tag{44}$$

for $B\subset\Theta=\mathrm{SU}(2)/\mathrm{U}(1)$, where $g(\theta)$ is a representative element of $\theta\in\mathrm{SU}(2)/\mathrm{U}(1)$. Any covariant POVM can be written as the above form. For our situation, covariant POVMs $M_T$ have $T$ as an operator that is diagonal in the Fock basis. If $B$ represents an infinitesimal ball about some $b\in\Theta$ and if $T$ represents a pure state $|\phi\rangle\langle\phi|$, then $M_T(B)$ is proportional to the operator $U_{j,g(b)}TU_{j,g(b)}^\dagger$, which in corresponds to a projector onto the state $U_{j,g(b)}|\phi\rangle$. Physically, the measurement of a covariant POVM $M_{|\phi\rangle\langle\phi|}(B)$ corresponds to a projection onto the states $U_{j,g(\theta)}|\psi\rangle$ according to the measure $\nu$.

While the function $R(M)$ is more difficult to calculate than $R_\nu(M)$, the situation simplifies greatly when the optimal POVM $M$ is covariant. In this situaton, the Bayesian error function can be equal to the worst-case error function in the sense that $1-R_\nu(M)=1-R(M)$ [47], [2, Chapter 4], [48, Chapter 4]. This situation is possible when $R(g\theta,g\hat{\theta})=R(\theta,\hat{\theta})$ and when $\nu$ is invariant under any $g$. In our scenario the POVM that maximizes $R_\nu(M)$ ($R(M)$) is realized by a covariant POVM [47], [2, Chapter 4], [48, Chapter 4]. In such a situation, we can calculate $R_\nu(M)$ instead of $R_\nu(M)$.

To minimize the error function, we use the idea of the addition of a spin-1/2 particle to a spin-$j$ particle,

$$\mathcal{H}_{\frac{1}{2}}\otimes\mathcal{H}_j = \mathcal{H}_{j+\frac{1}{2}}\oplus\mathcal{H}_{j-\frac{1}{2}}, \tag{45}$$

where $\mathcal{H}_j$ denotes the space for spin $j$. We denote the projection to $\mathcal{H}_{j+\frac{1}{2}}$ and $\mathcal{H}_{j-\frac{1}{2}}$ by $P_{j+\frac{1}{2}}$ and $P_{j-\frac{1}{2}}$, respectively.

**Theorem 1** ([48, Theorem 4.6]). *When the relation*

$$\frac{1}{2j+2}\mathrm{Tr}P_{j+\frac{1}{2}}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho \ge \frac{1}{2j}\mathrm{Tr}P_{j-\frac{1}{2}}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho \tag{46}$$

*holds under the relation* (45)*, the maximum of $R(M)$ is*

$$\frac{2j+1}{2j+2}\mathrm{Tr}P_{j+\frac{1}{2}}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho, \tag{47}$$

*and the optimal measurement is given as $M_{|j;j\rangle\langle j;j|}$, which is defined in* (44)*, and comprises of a measure on pure states with maximum total angular momentum.*

For reader's convenience, we give a proof for Theorem 1 in Appendix D as a special case of [48, Theorem 4.6]. The proof of Theorem 1 uses the following ideas. First, we represent the error function using a spin-1/2 representation while considering the estimation of a quantum state in the spin-$j$ representation. Second, we apply the definition of $R(M)$ and use the fact that the product of traces is equal to the trace of the tensor products of the arguments, and use Schur's lemma appropriately.

### C. Global estimation of a unitary channel

Here we apply the theory reviewed in Section V B to calculate the minimum MSE for the global estimation of our unitary model.

### 1. Probe state as a binomial distribution in the number basis

First, we consider whether the condition (46) holds when $p_m$ is given as a binomial distribution and $2j = n$. In this case, the LHS of (46) is

$$\frac{1}{n+2}\sum_{m=-j}^{j}\frac{j+m+1}{2j+1}\binom{n}{j+m}p^{j+m}(1-p)^{j-m} \tag{48}$$

$$=\frac{1}{n+2}\sum_{k=0}^{n}\frac{k+1}{n+1}\binom{n}{k}p^{k}(1-p)^{n-k}=\frac{1}{n+2}\frac{np+1}{n+1}. \tag{49}$$

The RHS of (46) is

$$\frac{1}{n}\sum_{m=-j}^{j}\frac{j-m}{2j+1}\binom{n}{j+m}p^{j+m}(1-p)^{j-m} \tag{50}$$

$$=\frac{1}{n}\sum_{k=0}^{n}\frac{n-k}{n+1}\binom{n}{k}p^{k}(1-p)^{n-k}=\frac{1}{n}\frac{n(1-p)}{n+1}. \tag{51}$$

When $n$ goes to infinity, the limit of $n$ times of LHS of (46) equals $p$ and the limit of $n$ times of RHS of (46) equals $1-p$. When $p>1/2$, with sufficiently large $n$, the condition (46) holds.

Then, the maximum of $R(M)$ is $(n+1)\frac{1}{n+2}\frac{np}{n+1}=\frac{np}{n+2}$, which converges to $p$. Hence the error function $\eta(M)=4(1-R(M))$ converges to $4(1-p)$, which is strictly larger than 0. Thus, we cannot make a precise global estimate of $\theta$ even with sufficiently large $n$, and the Cramér-Rao approach does not work well for the global estimation problem in this case.

### 2. Probe state as a geometric distribution in the number basis

When $p_m$ is a geometric distribution $\frac{r-1}{r^{2j+1}-1}r^{j+m}$, the LHS of (46) is

$$\frac{1}{n+2}\sum_{m=-j}^{j}\frac{j+m+1}{2j+1}\frac{r-1}{r^{2j+1}-1}r^{j+m} \tag{52}$$

$$=\frac{1}{n+2}\Big(\frac{1}{n+1}+\frac{1}{n+1}\big(\frac{nr^{n+1}+1}{r^{n+1}-1}-\frac{1}{r-1}\big)\Big). \tag{53}$$

The RHS of (46) is

$$\frac{1}{n}\sum_{m=-j}^{j}\frac{j-m}{2j+1}\frac{r-1}{r^{2j+1}-1}r^{j+m}$$

$$=\frac{1}{n}\sum_{m=-j}^{j}\frac{2j-(j+m)}{2j+1}\frac{r-1}{r^{2j+1}-1}r^{j+m}$$

$$=\frac{1}{n}\Big(\frac{n}{n+1}-\frac{1}{n+1}\big(\frac{nr^{n+1}+1}{r^{n+1}-1}-\frac{1}{r-1}\big)\Big). \tag{54}$$

When $r>1$ and $n$ goes to infinity, LHS of (46) approaches 1 and and RHS of (46) approaches 0. With sufficiently large $n$, the condition (46) holds. Then, the maximum of $R(M)$ is $\frac{n+1}{n+2}\Big(\frac{1}{n+1}+\frac{1}{n+1}\big(\frac{nr^{n+1}+1}{r^{n+1}-1}-\frac{1}{r-1}\big)\Big)$, which converges to 1, where $R(M)$ is defined in (42). When $M_n$ is the optimal estimator, we show in Appendix E that the corresponding error function is

$$\eta(M)=\frac{4r}{n(r-1)}-\frac{8}{n^2(r-1)}+O(n^{-3})+O(r^{-n-1}). \tag{55}$$

Therefore, the minimum error for the global estimate coincides with the RLD bound (38). In this case, the Cramér-Rao approach works well for our global estimation problem.

### 3. Probe state as a delta distribution in the number basis

Next consider the case where $p_m=\delta_{a,m}$. Then the LHS of (46) is $\frac{1}{n+2}\frac{j+a+1}{2j+1}$. The RHS of (46) is given by $\frac{1}{n}\frac{j-a}{2j+1}$. The difference between the LHS and the RHS of (46) gives the expression

$$\frac{1}{n(n+2)}\big(n(j+a+1)-(n+2)(j-a)\big)$$

$$=\frac{1}{n(n+2)}\big(2(n+1)a+(n-2j)\big). \tag{56}$$

Since $n=2j$, the expression in (56) tells us that (46) is equivalent to the inequality

$$a\geq 0. \tag{57}$$

Hence whenever we have a state $|n/2;a\rangle$ with $a\geq 0$, the condition (46) holds, and the maximum of $R(M)$ is $\frac{n+1}{n+2}\frac{n/2+a+1}{n+1}$. In the limit of large $n$ becomes large, this maximum $R(M)$ becomes $\frac{1}{2}+\frac{a}{n}$. For positive $a$, this $R(M)$ is at least $\frac{1}{2}$, and is bounded away from zero. Hence the global estimation strategy for such states in the basis $\mathcal{B}_n$ has a constant error. In contrast, the local estimation strategy has MSE that scales as $O(1/n^2)$. Hence the Cramér-Rao approach does not work well for the global estimation problem in this case.

As an example, we may consider the probe state given by $|n/2;0\rangle$ which corresponds to using a half-Dicke state for distinguishable spins. The state $|n/2;0\rangle$, commonly discussed as a quantum probe state that we can use for a quantum advantage in quantum sensing, can be prepared for instance in the procedure described in Ref [56]. According to (36), $F_{1,1}=F_{2,2}=8j(j+1)=2n(n+2)$, and the tight CR bound scales as $O(1/n^2)$. However, for global estimation strategies, the minimum MSE is a constant, because $R(M)=1/2$. Hence under global estimation strategies, the half-Dicke state loses its quantum advantage in sensing.

## VI. DISCUSSION

We have shown that that there are situations where for the state estimation problem, the minimum MSE obtained from the CR approach is accurate for the error function obtained for global estimation strategies. We have also shown that the opposite can be true; namely, there are situations where for the state estimation problem, the minimum MSE obtained from the CR approach is very different from the error function obtained for global estimation strategies. The most striking difference between the minimum MSE obtained from the CR approach and the minimum error function from global estimation is the situation of estimating a unitary model with the probe state $|n;n/2\rangle_B = |n/2;0\rangle$ In the context of local estimation, there is a Heisenberg scaling in the minimum MSE if we use $|n/2;0\rangle$ as the probe state for this unitary model. However in the limit of large $n$, we show that this Heisenberg scaling vanishes for global estimation strategies. Our results recommends that caution must be exercised if we wish to use CR bounds in the context of global estimation.

In the case of unitary estimation of a single parameter, it is known that the optimal Cramér-Rao type bound cannot be attained with global estimation [12, 13, 57]. The papers [12, 13, 57] considers the optimization of the initial state for the estimation of the unitary. However, this paper considers the state estimation with a fixed initial state. Furthermore, the paper [12] showed that the optimal Cramér-Rao type bound cannot be attained even under the problem of local minimax estimation even under the setting of asymptotically many probe states used for global estimation of unitary channels. This phenomenon relates to the Heisenberg scaling under the unitary estimation. In unitary estimation, while the optimal Cramér-Rao type bound has the same order as the error function for optimum minimax estimation [12, 13, 57–59], they differ in the coefficients of their leading order terms. Moreover, it was shown that the input state of the optimal Cramér-Rao type bound and the optimum minimax estimation are different. For example, although the noon state realizes the optimal Cramér-Rao type bound, it does not work for global estimation [60, Section VI]. Our results add to the literature of examples where the behavior of Cramér-Rao type bounds differs from the optimum minimax estimation, particularly with regards to the quantum estimation of bosonic states both in a single-parameter and a multi-parameter setting.

Our work is also related to the question as to whether a family of bosonic states which embed parameters can have a Cramér-Rao type bound that can be attained in the single copy setting. One family of quantum states that we considered is the state family which is a geometric distribution in the Fock basis. In Ref. [5, Section IV] showed that this state family approximates a quantum Gaussian state family. Moreover, in the setting of multiple identical and independently distributed copies, this geometric state family converges to the quantum Gaussian state family [4, 61, 62]. Given that the RLD bound can be attained under the single copy setting for the quantum Gaussian state family, we can see that our problem relates to the question as to whether the state family of our interest also attains the Cramér-Rao type bound in the single copy setting. We leave this line of enquiry for future work.

**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).

## Appendix A: Local estimation under the general unitary model

To show several relations in Section V, we consider the spin $j$ system $\mathcal{H}_j$ spanned by $\{|j;m\rangle\}_{m=-j}^{j}$. For simplicity, if $j$ is clear from the context, we denote $|j;m\rangle$ as $|m\rangle$.

We have two operators as

$$J_+ := \sum_{m=-j}^{j-1} \sqrt{j(j+1)-m(m+1)}|m+1\rangle\langle m| \tag{A1}$$

$$= \sum_{m=-j}^{j-1} \sqrt{(j-m)(j+m+1)}|m+1\rangle\langle m| \tag{A2}$$

$$J_- := \sum_{m=-j}^{j-1} \sqrt{j(j+1)-m(m+1)}|m\rangle\langle m+1| \tag{A3}$$

$$= \sum_{m=-j}^{j-1} \sqrt{(j-m)(j+m+1)}|m\rangle\langle m+1|. \tag{A4}$$

We write the angular momentum operators as

$$J_1 := \frac{1}{2}(J_+ + J_-),\ J_2 := \frac{1}{2i}(J_+ - J_-) \tag{A5}$$

$$J_3 := \sum_{m=-j}^{j} m|m\rangle\langle m|. \tag{A6}$$

Then, the Casimir element $C$ is given as

$$C := \sum_{k=1}^{3} J_k^2 = \frac{1}{4}((J_+ + J_-)^2 - (J_+ - J_-)^2) + J_3^2 \tag{A7}$$

$$= \frac{1}{2}(J_+J_- + J_-J_+) + J_3^2 = j(j+1). \tag{A8}$$

Now, we consider the following case

$$\rho := \sum_{m=-j}^{j} p_m |m\rangle\langle m| \tag{A9}$$

$$D_1 := i[\rho, J_1],\ D_2 := i[\rho, J_2]. \tag{A10}$$

Then, we have

$$i[\rho, J_1] = \frac{i}{2}[\rho, (J_+ + J_-)] = \frac{i}{2}(-J_+\rho + \rho J_+ - J_-\rho + \rho J_-) \tag{A11}$$

$$i[\rho, J_2] = \frac{1}{2}[\rho, (J_+ - J_-)] = \frac{1}{2}(-J_+\rho + \rho J_+ + J_-\rho - \rho J_-). \tag{A12}$$

We define

$$K_+ := \sum_{m=-j}^{j-1} \frac{2(p_{m+1}-p_m)}{p_{m+1}+p_m}\sqrt{(j-m)(j+m+1)}|m+1\rangle\langle m| \tag{A13}$$

$$K_- := \sum_{m=-j}^{j-1} \frac{2(p_{m+1}-p_m)}{p_{m+1}+p_m}\sqrt{(j-m)(j+m+1)}|m\rangle\langle m+1|. \tag{A14}$$

Then, we have

$$\rho\circ(\frac{1}{2}(K_+ + K_-)) = i[\rho, J_2] = D_2 \tag{A15}$$
$$\rho\circ(-\frac{1}{2i}(K_+ - K_-)) = i[\rho, J_1] = D_1. \tag{A16}$$

Thus, the SLDs $L_1$ and $L_2$ of the first and second parameters are calculated as

$$L_2 := \frac{1}{2}(K_+ + K_-),\ L_1 := -\frac{1}{2i}(K_+ - K_-). \tag{A17}$$

Thus, we have

$$F_{1,1} = \mathrm{Tr}\rho L_1^2 \tag{A18}$$
$$= \sum_{m=-j}^{j-1}\Big(\frac{4(p_{m+1}-p_m)^2}{(p_{m+1}+p_m)^2}(j-m)(j+m+1)p_m \tag{A19}$$
$$+ \frac{4(p_{m+1}-p_m)^2}{(p_{m+1}+p_m)^2}(j-m)(j+m+1)p_{m+1}\Big) \tag{A20}$$
$$= \sum_{m=-j}^{j-1}\frac{4(p_{m+1}-p_m)^2}{p_{m+1}+p_m}(j-m)(j+m+1) \tag{A21}$$
$$F_{2,2} = \mathrm{Tr}\rho L_2^2 \tag{A22}$$
$$= \sum_{m=-j}^{j-1}\frac{4(p_{m+1}-p_m)^2}{p_{m+1}+p_m}(j-m)(j+m+1) \tag{A23}$$
$$F_{1,2} = \mathrm{Tr}\rho \tilde{L}_1\circ\tilde{L}_2 = 0, \tag{A24}$$

which shows (36).

In particular, when the distribution $\{p_m\}$ is a geometric distribution $p_{G,m} = \frac{r-1}{r^{j+1}-r^{-j}}r^m$, we have

$$\frac{2(p_{m+1}-p_m)}{p_{m+1}+p_m} = \frac{2(r^{m+1}-r^m)}{r^{m+1}+r^m} = \frac{2(r-1)}{r+1}, \tag{A25}$$

which implies that

$$K_+ = \frac{2(r-1)}{r+1}J_+,\ K_- = \frac{2(r-1)}{r+1}J_-. \tag{A26}$$

Using (A15) and (A16), we have

$$\rho\circ L_2 = i[\rho, \frac{r+1}{2(r-1)}L_1],\ \rho\circ L_1 = i[\rho, \frac{r+1}{2(r-1)}L_2]. \tag{A27}$$

These relations guarantee the D-invariance.

**Appendix B: Local estimation with the binomial distribution: Proof of (37)**

We recall the Clebsch–Gordan formula, which for $2j = n$ gives

$$\langle j+\frac{1}{2}; m+\frac{1}{2}|\frac{1}{2},\frac{1}{2}; j, m\rangle^2 = \frac{(2j+2)(2j)!(j+m)!(j-m)!(j+m+1)!(j-m)!}{(2j+2)!((j+m)!(j-m)!)^2} \tag{B1}$$
$$= \frac{(2j+2)(j+m+1)}{(2j+2)(2j+1)} = \frac{(j+m+1)}{(2j+1)}. \tag{B2}$$

Now we also consider $p_m = \binom{n}{j+m}p^{j+m}(1-p)^{j-m}$.

Then, we have

$$F_{1,1} = \sum_{m=-j}^{j-1} \frac{4(p_{m+1} - p_m)^2}{p_{m+1} + p_m}(j-m)(j+m+1) \tag{B3}$$

$$= \sum_{m=-j}^{j-1} \frac{4(\binom{n}{j+m+1}p^{j+m+1}(1-p)^{j-m-1} - \binom{n}{j+m}p^{j+m}(1-p)^{j-m})^2}{\binom{n}{j+m+1}p^{j+m+1}(1-p)^{j-m-1} + \binom{n}{j+m}p^{j+m}(1-p)^{j-m}} \tag{B4}$$

$$\cdot (j-m)(j+m+1) \tag{B5}$$

$$= \sum_{m=-j}^{j-1} \frac{4(\frac{j-m}{j+m+1}\frac{p}{1-p} - 1)^2\binom{n}{j+m}p^{j+m}(1-p)^{j-m}}{\frac{j-m}{j+m+1}\frac{p}{1-p} + 1}(j-m)(j+m+1) \tag{B6}$$

$$= \sum_{m=-j}^{j-1} \frac{4((j-m)\frac{p}{1-p} + (j+m+1))^2\binom{n}{j+m}p^{j+m}(1-p)^{j-m}}{(j-m)\frac{p}{1-p} + (j+m+1)}(j-m) \tag{B7}$$

$$= \sum_{m=-j}^{j-1} \frac{4((j-m)p - (j+m+1)(1-p))^2\binom{n}{j+m}p^{j+m}(1-p)^{j-m-1}}{(j-m)p + (j+m+1)(1-p)}(j-m) \tag{B8}$$

$$= \sum_{m=-j}^{j-1} \frac{4(-(j+m+1) + (2j+1)p)^2\binom{n}{j+m}p^{j+m}(1-p)^{j-m-1}}{j+m+1-(2m+1)p}(j-m) \tag{B9}$$

$$= \sum_{k=0}^{n-1} \frac{4(-(k+1) + (n+1)p)^2\binom{n}{k}p^{k}(1-p)^{2n-k-1}}{k+1-(2k-n+1)p}(n-k) \tag{B10}$$

$$= n^2 \sum_{k=0}^{n-1} \frac{4(-\frac{k+1}{n} + (1+\frac{1}{n})p)^2\binom{n}{k}p^{k}(1-p)^{2n-k-1}}{\frac{k+1}{n} - (\frac{2k}{n} - 1 + \frac{1}{n})p}(1 - \frac{k}{n}). \tag{B11}$$

Now we interpret the index $k$ as a measurement outcome we obtain from measuring the state in the Fock basis, and $Y = k/n$ as the corresponding random variable. Then we can write $F_{1,1}$ in terms of $Y$ to get

$$F_{1,1} = n^2 \mathbb{E}_{p, \frac{p(1-p)}{n}}\Big[\frac{4(-Y - \frac{1}{n} + (1+\frac{1}{n})p)^2(1-p)^{-1}(1-Y)}{Y + \frac{1}{n} - (2Y - 1 + \frac{1}{n})p}\Big], \tag{B12}$$

where the first subscript on the expectation denotes the mean of $Y$, and the second subscript denotes the variance of $Y$. We then define the random variable $X = \sqrt{n}(Y - p)$, and get

$$F_{1,1} = n^2 \mathbb{E}_{0,p(1-p)}\Big[\frac{4(-p - \frac{X}{\sqrt{n}} + (1+\frac{1}{n})p)^2(1-p)^{-1}(1-p-\frac{X}{\sqrt{n}})}{p + \frac{X}{\sqrt{n}} + \frac{1}{n} - (2p + 2\frac{X}{\sqrt{n}} - 1 + \frac{1}{n})p}\Big] \tag{B13}$$

$$= n^2 \mathbb{E}_{0,p(1-p)}\Big[\frac{4(\frac{X}{\sqrt{n}} - \frac{p}{n})^2(1-p)^{-1}(1-p-\frac{X}{\sqrt{n}})}{-2p^2 + 2p + (1-2p)\frac{X}{\sqrt{n}} + \frac{1-p}{n}}\Big] \tag{B14}$$

$$\cong n^2 \mathbb{E}_{0,p(1-p)}\Big[\frac{2\frac{X^2}{n}(1-p)^{-1}(1-p)}{p(1-p)}\Big] = n\mathbb{E}_{0,p(1-p)}\Big[\frac{2X^2}{p(1-p)}\Big] = 2n, \tag{B15}$$

where the congruent symbol indicates an approximation in the limit of large $n$. Hence, we obtain (37).

### Appendix C: Local estimation with the geometric distribution

Assume that $p_m = \frac{r-1}{r^{2j+1}-1}r^{j+m}$. Since

$$\frac{2(r^{j+m+1} - r^{j+m})}{r^{j+m+1} + r^{j+m}} = \frac{2(r-1)}{r+1}, \tag{C1}$$

we have

$$K_+ = \sum_{m=-j}^{j-1} \frac{2(r-1)}{r+1}\sqrt{(j-m)(j+m+1)}|m+1\rangle\langle m| = \frac{2(r-1)}{r+1}J_+ \tag{C2}$$

$$K_- = \sum_{m=-j}^{j-1} \frac{2(r-1)}{r+1}\sqrt{(j-m)(j+m+1)}|m\rangle\langle m+1| = \frac{2(r-1)}{r+1}J_-. \tag{C3}$$

Thus, we have

$$L_2 = \frac{2(r-1)}{r+1}J_1,\ L_1 = -\frac{2(r-1)}{r+1}J_2. \tag{C4}$$

Thus, we have

$$F_{1,1} = F_{2,2} = \sum_{m=-j}^{j-1} \frac{4(p_{m+1}-p_m)^2}{p_{m+1}+p_m}(j-m)(j+m+1) \tag{C5}$$

$$= \sum_{m=-j}^{j-1} 4p_m(r-1)\frac{2(r-1)}{r+1}(j-m)(j+m+1) \cong \frac{n}{2}\frac{2(r-1)}{r+1}. \tag{C6}$$

Relation (A15) and (A16) are rewritten as

$$-\rho \circ \frac{2(r+1)}{r-1}J_2 = i[\rho, J_1] \tag{C7}$$

$$\rho \circ \frac{2(r+1)}{r-1}J_1 = i[\rho, J_2]. \tag{C8}$$

Thus, this model satisfies the D-invariant condition. The matrix $D = (D_{j,k})$ is approximated to

$$\frac{n}{2}\frac{2(r-1)}{r+1} \cdot \frac{2(r-1)}{r+1}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{C9}$$

Thus

$$C_\theta^{\mathrm{R}}[I] = \mathrm{Tr} F^{-1} + \frac{1}{2}\mathrm{Tr}|F^{-1}DF^{-1}| \tag{C10}$$

$$\cong \frac{2(r+1)}{n(r-1)} + \frac{2}{n} = \frac{2(r+1)+2(r-1)}{n(r-1)} = \frac{4r}{n(r-1)}, \tag{C11}$$

which implies (38).

### Appendix D: Proof of Theorem 1

For reader's convenience, we give a proof for Theorem 1 as a special case of [48, Theorem 4.6].

When our estimator $\hat{\theta}$ corresponds to $g \in \mathrm{SU}(2)$ and when the true parameter is 0, the fidelity function as given in (40) can be expressed as

$$R(0,\hat{\theta}) = \cos^2\frac{|\hat{\theta}|}{2} = \mathrm{Tr}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|U_{1/2,g}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|U_{1/2,g}^\dagger \tag{D1}$$

where $U_{1/2,g}$ is a spin-1/2 unitary representation of $g \in \mathrm{SU}(2)$.

Now, let us write the true state as $\rho$, and write the POVM element that corresponds to $g \in \mathrm{SU}(2)$ as $(2j+1)U_{j,g}\rho' U_{j,g}^\dagger$ for some state $\rho'$. We consider the case when our measurement is given as $M_{\rho'}$. Then, by using the Haar measure $\nu$ on

SU(2), the Bayesian average of $R(0,\hat{\theta})$ is calculated as

$$R(M_{\rho'}) = \int_{\mathrm{SU}(2)} \mathrm{Tr}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|U_{1/2,g}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|U_{1/2,g}^\dagger \cdot (2j+1)\mathrm{Tr}\rho U_{j,g}\rho' U_{j,g}^\dagger\, \nu(dg) \tag{D2}$$

$$= \int_{\mathrm{SU}(2)} (2j+1)\mathrm{Tr}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho(U_{1/2,g}\otimes U_{j,g})|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho'(U_{1/2,g}\otimes U_{j,g})^\dagger\, \nu(dg) \tag{D3}$$

$$=(2j+1)\mathrm{Tr}(P_{j+\frac{1}{2}}+P_{j-\frac{1}{2}})(|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho)(P_{j+\frac{1}{2}}+P_{j-\frac{1}{2}})\int_{\mathrm{SU}(2)}(U_{1/2,g}\otimes U_{j,g})(|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho')(U_{1/2,g}\otimes U_{j,g})^\dagger\, \nu(dg). \tag{D4}$$

Here, $P_j$ is a projector onto the space with total spin of $j$. We obtain the above by using the properties of the trace function, which allows us to rewrite the Bayesian average as the integral of the trace of operators on a tensor product space. In the next step we use properties of how a spin-$j$ space combines with a spin-1/2 space.

$$\begin{aligned}R(M_{\rho'}) \overset{(a)}{=}&(2j+1)\mathrm{Tr}P_{j+\frac{1}{2}}(|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho)P_{j+\frac{1}{2}}\int_{\mathrm{SU}(2)}(U_{1/2,g}\otimes U_{j,g})(|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho')(U_{1/2,g}\otimes U_{j,g})^\dagger\, \nu(dg)\\ &+(2j+1)\mathrm{Tr}P_{j-\frac{1}{2}}(|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho)P_{j-\frac{1}{2}}\int_{\mathrm{SU}(2)}(U_{1/2,g}\otimes U_{j,g})(|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho')(U_{1/2,g}\otimes U_{j,g})^\dagger\, \nu(dg)\end{aligned} \tag{D5}$$

To obtain this note that since the operator $\int_{\mathrm{SU}(2)}(U_{1/2,g}\otimes U_{j,g})(|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho')(U_{1/2,g}\otimes U_{j,g})^\dagger\, \nu(dg)$ is commutative with $(U_{1/2,g'}\otimes U_{j,g'})$, Schur's lemma guarantees that this operator is a sum of constant times of $P_{j+\frac{1}{2}}$ and $P_{j-\frac{1}{2}}$. Hence, this operator is commutative with $P_{j+\frac{1}{2}}$ and $P_{j-\frac{1}{2}}$. Hence, we obtain Step $(a)$.

Next, we proceed to evaluate the integrals and obtain

$$\begin{aligned}R(M_{\rho'}) \overset{(b)}{=}&(2j+1)\mathrm{Tr}P_{j+\frac{1}{2}}(|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho)\frac{1}{2j+2}(\mathrm{Tr}P_{j+\frac{1}{2}}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho')P_{j+\frac{1}{2}}\\ &+(2j+1)\mathrm{Tr}P_{j-\frac{1}{2}}(|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho)\frac{1}{2j}(\mathrm{Tr}P_{j-\frac{1}{2}}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho')P_{j-\frac{1}{2}}\end{aligned} \tag{D6}$$

Since the irreducibility of the spaces $\mathcal{H}_{j+\frac{1}{2}}$ and $\mathcal{H}_{j-\frac{1}{2}}$ guarantees that the first and second integral terms in Step (a) equal constant times of $P_{j+\frac{1}{2}}$ and $P_{j-\frac{1}{2}}$, respectively, we obtain Step $(b)$.

Finally, we simplify further and get

$$R(M_{\rho'}) = \frac{2j+1}{2j+2}\mathrm{Tr}P_{j+\frac{1}{2}}(|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho)(\mathrm{Tr}P_{j+\frac{1}{2}}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho') \tag{D7}$$

$$+\frac{2j+1}{2j}\mathrm{Tr}P_{j-\frac{1}{2}}(|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho)(\mathrm{Tr}P_{j-\frac{1}{2}}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho') \tag{D8}$$

$$\overset{(c)}{\le}\frac{2j+1}{2j+2}\mathrm{Tr}P_{j+\frac{1}{2}}(|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho). \tag{D9}$$

The inequality $(c)$ follows from (46) and the relation $(\mathrm{Tr}P_{j+\frac{1}{2}}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho')+(\mathrm{Tr}P_{j-\frac{1}{2}}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes\rho')=1$. Further, the equality holds when $\rho'=|j;j\rangle\langle j;j|$ because $(\mathrm{Tr}P_{j+\frac{1}{2}}|\frac{1}{2};\frac{1}{2}\rangle\langle\frac{1}{2};\frac{1}{2}|\otimes|j;j\rangle\langle j;j|)=1$.

### Appendix E: The error function for the global estimate on the unitary model with the geometric distribution

With $M_n$ as the optimal covariant estimator for the unitary model on the geometric distribution, we have from Theorem 1 that

$$1-R(M_n)=1-\frac{n+1}{n+2}\Big(\frac{1}{n+1}+\frac{1}{n+1}(\frac{nr^{n+1}+1}{r^{n+1}-1}-\frac{1}{r-1})\Big)=1-\Big(\frac{1}{n+2}+\frac{1}{n+2}(\frac{nr^{n+1}+1}{r^{n+1}-1}-\frac{1}{r-1})\Big) \tag{E1}$$

$$=1-\Big(\frac{1}{n+2}+\frac{1}{n+2}(\frac{n+r^{-n-1}}{1-r^{-n-1}}-\frac{1}{r-1})\Big) \tag{E2}$$

$$=1-\Big(\frac{1}{n+2}+\frac{1}{n+2}(n+O(r^{-n-1})-\frac{1}{r-1})\Big) \tag{E3}$$

$$=\frac{1}{n+2}\Big(n+2-1-(n+O(r^{-n-1})-\frac{1}{r-1})\Big) \tag{E4}$$

$$=\frac{1}{n+2}\Big(1-\frac{1}{r-1}+O(r^{-n-1})\Big) \tag{E5}$$

$$=\frac{r}{(n+2)(r-1)}+O(r^{-n-1}) \tag{E6}$$

$$=\frac{r}{n(r-1)}-\frac{2}{n^2(r-1)}+O(\frac{1}{n^3})+O(r^{-n-1}). \tag{E7}$$


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

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

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

[4] Y. Yang, G. Chiribella, and M. Hayashi, Attaining the ultimate precision limit in quantum state estimation, Communications in Mathematical Physics **368**, 223 (2019).

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

[6] M. Hayashi and Y. Ouyang, Tight Cramér-Rao type bounds for multiparameter quantum metrology through conic programming, Quantum **7**, 1094 (2023).

[7] W. Górecki, R. Demkowicz-Dobrzański, H. M. Wiseman, and D. W. Berry, $\pi$-corrected heisenberg limit, Physical review letters **124**, 030501 (2020).

[8] J. J. Meyer, S. Khatri, D. S. França, J. Eisert, and P. Faist, Quantum metrology in the finite-sample regime, arXiv preprint arXiv:2307.06370 10.48550/arXiv.2307.06370 (2023).

[9] W. Górecki and R. Demkowicz-Dobrzański, Multiparameter quantum metrology in the heisenberg limit regime: Many-repetition scenario versus full optimization, Phys. Rev. A **106**, 012424 (2022).

[10] G. Chesi, A. Riccardi, R. Rubboli, L. Maccone, and C. Macchiavello, Protocol for global multiphase estimation, Phys. Rev. A **108**, 012613 (2023).

[11] M. J. W. Hall and H. M. Wiseman, Heisenberg-style bounds for arbitrary estimates of shift parameters including prior information, New Journal of Physics **14**, 033040 (2012).

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

[13] M. Hayashi, Parallel treatment of estimation of SU(2) and phase estimation, Physics Letters A **354**, 183 (2006).

[14] M. Hayashi, S. Vinjanampathy, and L. Kwek, Resolving unattainable cramer–rao bounds for quantum sensors, Journal of Physics B: Atomic, Molecular and Optical Physics **52**, 015503 (

[15] N. Thomas-Peter, B. J. Smith, A. Datta, L. Zhang, U. Dorner, and I. A. Walmsley, Real-world quantum sensors: Evaluating resources for precision measurement, Phys. Rev. Lett. **107**, 113603 (2011).

[16] J. A. Jones, S. D. Karlen, J. Fitzsimons, A. Ardavan, S. C. Benjamin, G. A. D. Briggs, and J. J. L. Morton, Magnetic field sensing beyond the standard quantum limit using 10-spin noon states, Science **324**, 1166 (2009).

[17] R. Okamoto, H. F. Hofmann, T. Nagata, J. L. O'Brien, K. Sasaki, and S. Takeuchi, Beating the standard quantum limit: phase super-sensitivity of n-photon interferometers, New Journal of Physics **10**, 073033 (2008).

[18] T. Nagata, R. Okamoto, J. L. O'Brien, K. Sasaki, and S. Takeuchi, Beating the standard quantum limit with four-entangled photons, Science **316**, 726 (2007).

[19] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum-enhanced measurements: Beating the standard quantum limit, Science **306**, 1330 (2004).

[20] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum metrology, Phys. Rev. Lett. **96**, 010401 (2006).

[21] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nature Photonics **5**, 222–229 (2011).

[22] L. Paninski, Estimating entropy on m bins given fewer than m samples, IEEE Transactions on Information Theory **50**, 2200 (2004).

[23] G. Valiant and P. Valiant, Estimating the unseen: an n/log (n)-sample estimator for entropy and support size, shown optimal via new clts, in *Proceedings of the forty-third annual ACM symposium on Theory of comput* (2011) pp. 685–694.

[24] J. Acharya, A. Orlitsky, A. T. Suresh, and H. Tyagi, The complexity of estimating rényi entropy, in *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms* (SIAM, 2014) pp. 1855–1869.
[25] J. Acharya, I. Issa, N. V. Shende, and A. B. Wagner, Estimating quantum entropy, IEEE Journal on Selected Areas in Information Theory **1**, 454 (2020).
[26] Q. Wang and Z. Zhang, Time-efficient quantum entropy estimator via samplizer, arXiv preprint arXiv:2401.09947 10.48550/arXiv.2401.09947 (2024).
[27] M. Tsang, F. Albarelli, and A. Datta, Quantum semiparametric estimation, Phys. Rev. X **10**, 031023 (2020).
[28] M. Hayashi, Measuring quantum relative entropy with finite-size effect, arXiv preprint arXiv:2406.17299 10.48550/arXiv.2406.17299 (2024).
[29] M. Hayashi, Optimal sequence of quantum measurements in the sense of Stein's lemma in quantum hypothesis testing, Journal of Physics A: Mathematical and General **35**, 10759 (2002).
[30] M. Hayashi and S. Watanabe, Information geometry approach to parameter estimation in markov chains, The Annals of Statistics **44**, 1495 (2016).
[31] M. Hayashi, Information geometry approach to parameter estimation in hidden Markov model, Bernoulli **28**, 307 (2022).
[32] H. Nagaoka and A. Fujiwara, Autoparallelity of quantum statistical manifolds in the light of quantum estimation theory (2023), arXiv:arXiv:2307.03431 [quant-ph].
[33] J. R. Anglin and W. Ketterle, Bose-Einstein condensation of atomic gases, Nature **416**, 211 (2002).
[34] M. T. Johnsson, N. R. Mukty, D. Burgarth, T. Volz, and G. K. Brennen, Geometric pathway to scalable quantum sensing, Phys. Rev. Lett. **125**, 190403 (2020).
[35] Y. Ouyang and G. K. Brennen, Quantum error correction on symmetric quantum sensors, arXiv preprint arXiv:2212.06285 10.48550/arXiv.2212.06285 (2022).
[36] G. D'Ariano, C. Macchiavello, and M. Sacchi, On the general problem of quantum phase estimation, Physics Letters A **248**, 103 (1998).
[37] C. Macchiavello, Optimal estimation of multiple phases, Phys. Rev. A **67**, 062302 (2003).
[38] V. Gebhart, A. Smerzi, and L. Pezzè, Bayesian quantum multiphase estimation algorithm, Phys. Rev. Appl. **16**, 014035 (2021).
[39] W. Górecki and R. Demkowicz-Dobrzański, Multiple-phase quantum interferometry: Real and apparent gains of measuring all the phases simultaneously, Phys. Rev. Lett. **128**, 040504 (2022).
[40] M. Hayashi, Z.-W. Liu, and H. Yuan, Global heisenberg scaling in noisy and practical phase estimation, Quantum Science and Technology **7**, 025030 (2022).
[41] A. Luis and J. Peřina, Optimum phase-shift estimation and the quantum description of the phase difference, Phys. Rev. A **54**, 4564 (1996).
[42] V. Bužek, R. Derka, and S. Massar, Optimal quantum clocks, Phys. Rev. Lett. **82**, 2207 (1999).
[43] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Entanglement-free heisenberg-limited phase estimation, Nature **450**, 393–396 (2007).
[44] A. Y. Kitaev, Quantum measurements and the abelian stabilizer problem (1995), arXiv:quant-ph/9511026 [quant-ph].
[45] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Quantum algorithms revisited, Proc. R. Soc. Lond. A. **454**, 339–354 (1998).
[46] H. Imai and M. Hayashi, Fourier analytic approach to phase estimation in quantum systems, New Journal of Physics **11**, 043034 (2009).
[47] A. Holevo, Covariant measurements and uncertainty relations, Reports on Mathematical Physics **16**, 385 (1979).
[48] M. Hayashi, *Group Representation for Quantum Theory* (Springer, 2017).
[49] Y. Ouyang and R. Chao, Permutation-invariant constant-excitation quantum codes for amplitude damping, IEEE Transactions on Information Theory **66**, 2921 (2019).
[50] M. Erhard, M. Krenn, and A. Zeilinger, Advances in high-dimensional quantum entanglement, Nature Reviews Physics **2**, 365 (2020).
[51] A. Luis, Quantum mechanics as a geometric phase: phase-space interferometers, Journal of Physics A: Mathematical and General **34**, 7677 (2001).
[52] X. Wang and P. Zanardi, Simulation of many-body interactions by conditional geometric phases, Phys. Rev. A **65**, 032327 (2002).
[53] M. Hayashi, A linear programming approach to attainable cramér-rao type bounds and randomness condition (1997), arXiv:quant-ph/9704044 [quant-ph].
[54] H. Nagaoka, A new approach to Cramér-Rao bounds for quantum state estimation, IEICE Tech Report **IT 89-42**, 9 (1989), (Reprinted in [**?** ]).
[55] S. Amari and H. Nagaoka, *Methods of information geometry*, Vol. 191 (American Mathematical Soc., 2000).
[56] Y.-Q. Zou, L.-N. Wu, Q. Liu, X.-Y. Luo, S.-F. Guo, J.-H. Cao, M. K. Tey, and L. You, Beating the classical precision limit with spin-1 Dicke states of more than 10,000 atoms, Proceedings of the National Academy of Sciences **115**, 6381 (2018).
[57] M. Hayashi, Fourier analytic approach to quantum estimation of group action, Communications in Mathematical Physics **347**, 3 (2016).
[58] A. Fujiwara, Estimation of SU(2) operation and dense coding: An information geometric approach, Phys. Rev. A **65**, 012316 (2001).
[59] H. Imai and A. Fujiwara, Geometry of optimal estimation scheme for SU(D) channels, Journal of Physics A: Mathematical and Theoretical **40**, 4391 (2007).
[60] M. Hayashi, Phase estimation with photon number constraint, Progress in Informatics , 81 (2011).
[61] J. Kahn and M. Guţă, Local asymptotic normality for finite dimensional quantum systems, Communications in Mathematical Physics **289**, 597 (2009).
[62] J. Kahn, *Quantum local asymptotic normality and other questions of quantu* (Leiden University, 2008).