Finite-error metrological bounds on the multiparameter Hamiltonian estimation

Naoto Kura
Department of Physics, University of Tokyo, 7-3-1 Hongo, Bunkyou-ku, 113-0022, Japan

Masahito Ueda
Department of Physics, University of Tokyo, 7-3-1 Hongo, Bunkyou-ku, 113-0022, Japan and
RIKEN Center for Emergent Matter Science (CEMS), Wako, Saitama, 351-0198, Japan
(Dated: July 23, 2018)

Abstract

Estimation of multiple parameters in an unknown Hamiltonian is investigated. We present upper and lower bounds on the time required to complete the estimation within a prescribed error tolerance $\delta$. The lower bound is given on the basis of the Cramér-Rao inequality, where the quantum Fisher information is bounded by the squared evolution time. The upper bound is obtained by an explicit construction of estimation procedures. By comparing the cases with different numbers of Hamiltonian channels, we also find that the few-channel procedure with adaptive feedback and the many-channel procedure with entanglement are equivalent in the sense that they require the same amount of time resource up to a constant factor.
I. INTRODUCTION

Since the birth of quantum estimation due in large part to Holevo \cite{1} and Helstrom \cite{2}, the information-theoretic aspects of quantum mechanics have been studied in many subfields of physics \cite{3–6}. Quantum metrology, the field in which the estimation of quantum dynamics is studied, marks significant differences between quantum and classical informatics. In estimating a phase-shift operator, for instance, the asymptotic accuracy increases in proportion to the amount of resource \cite{7, 8}, which is quadratically better than the limitation set by classical statistics. This quantum-metrological advantage has been demonstrated in optomechanics \cite{9–12} and ultracold atomic gases \cite{13–16}. The quantum-specific enhancement is related to some characteristic features in quantum mechanics, such as entanglement, spin squeezing and Bose statistics, on which quantitative studies have been carried out \cite{11, 17, 18}. Furthermore, it has been pointed out that quantum computation, including the best-known quantum algorithms by Shor \cite{19} and Grover \cite{20}, takes advantage of quantum metrology \cite{21, 22}.

Quantum metrology originally targeted the one-parameter dynamics, which is essentially the estimation of a single phase. In more general situations, however, inference on the dynamics involves more than one parameter, in which case the problem becomes more involved. For example, we need to take into consideration the simultaneous measurement of noncommuting observables and an exponential increase in the number of candidates for the true parameters. As such, quantum metrology in multiparameter cases has attracted growing interest in recent years \cite{23, 24}, including the estimation of multiple phases \cite{25–29}, the Hamiltonian itself \cite{30–32} and multidimensional fields \cite{33–35}. It is also known that estimation of a large-sized Hamiltonian plays a crucial role in setting computational bounds on quantum algorithms \cite{22, 36, 37}. The multiparameter quantum metrology also exhibits quantum enhancement \cite{38} in that the resource can significantly be reduced by quantum mechanics. On the other hand, it remains unclear how the resource depends on the size of the Hilbert space and the number of parameters to be estimated.

Recently, Yuan et al. studied the Hamiltonian estimation in a $d$-dimensional Hilbert space \cite{32, 39}. By comparing sequential and parallel schemes for exploiting the quantum resource, they conclude that the latter is $O(d)$ times more efficient than the former in estimating the full Hamiltonian. The proof involves two assumptions. First, the vector
parameter \( \theta \) to be estimated is sufficiently close to a certain value \( \theta_0 \). In other words, there exists a “search radius” \( E \) such that \( \| \theta - \theta_0 \| \leq E \) is presupposed. Second, the Hamiltonian \( H \) is replaced by the unitary channel \( e^{-i\tau H} \), with the evolution time \( \tau \) fixed. Here the following problem arises: although the search radii for two schemes are both sufficiently small, their ratio is found to be nowhere near unity. In fact, to compare the two schemes with \( r \) unitary channels, the radius for the sequential scheme should be \( r \) times smaller than that for the parallel scheme, since the former undergoes \( r \) times longer evolution than the latter. Noting that a larger search radius implies a stronger procedure, the comparison between the two schemes made in Ref. [32] is generally not fair for large \( r \).

We address the multiparameter quantum metrology in the following setting: We fix the tolerated error \( \delta \), and suppose that the evolution time \( \tau \) can be arbitrary. We obtain upper and lower bounds on the time resource required for the estimation in terms of \( m, d \) and \( \delta \), where \( m \) is the number of parameters and \( d \) is the dimension of the Hilbert space. In particular, we find that the time resource scales with \( \delta^{-1} \), which explicitly shows the quantum-metrological limit. Furthermore, we find that the sequential and parallel schemes require the same amount of time resource up to some constant factor in order to achieve the same accuracy of estimation, contrary to Yuan’s result.

Let us explain how the difference arises. First, one needs to prepare \( N \) copies of probe states \( |q_\theta\rangle \) by using the Hamiltonian \( H_\theta \), from which the unknown vector \( \theta \) is estimated. Given an unbiased estimator of \( \theta \), which we denote by \( \theta^* \), the covariance matrix \( V(\theta) \) is bounded from below by the quantum Cramér-Rao (QCR) inequality:

\[
V(\theta) \geq N^{-1}J(\theta)^{-1},
\]

where \( J(\theta) \) is the quantum Fisher information (QFI) matrix of \( |q_\theta\rangle \) formulated as

\[
[J(\theta)]_{jk} = 4 \text{Re} \left\langle \frac{\partial}{\partial \theta^j} q_\theta \right| \left[ 1 - |q_\theta\rangle\langle q_\theta| \right] \left| \frac{\partial}{\partial \theta^k} q_\theta \right\rangle.
\]

Suppose that \( N \) copies of the quantum states are given. Noting that \( \text{Tr}[V(\theta)] \) is the expectation value of \( \| \theta^* - \theta \|^2 \), the accuracy \( \delta \) can be achieved when

\[
N^{-1} \text{Tr}[J(\theta)^{-1}] \leq \text{Tr}[V(\theta)] \sim \delta^2,
\]

or equivalently when

\[
N \gtrsim \delta^{-2} \text{Tr}[J(\theta)^{-1}].
\]
The equality in (4) can be asymptotically saturated for sufficiently large $N$ on the condition that an appropriate measurement exists, which is the case with Ref. [32]. As a result, the equality in (4) is satisfied in the limit of $\delta \to 0$. If we denote by $\tau$ the time it takes to prepare a probe state $|q_\theta\rangle$, the metrological bound can be written as

$$T = N\tau \sim \delta^{-2} f(\tau) \quad (\delta \to 0),$$

(5)

$$f(\tau) = \tau \inf_{|q_\theta\rangle} \text{Tr}[J(\theta)^{-1}],$$

(6)

where the infimum is taken over all probe states $|q_\theta\rangle$ that can be prepared during time $\tau$.

While (6) gives a rigorous relation for every fixed $\tau$, it is guaranteed only for sufficiently small $\delta$. Moreover, the extent to which $\delta$ should be small depends on the time $\tau$ in the general case. Therefore, the metrological bound for finite $\delta > 0$ cannot be determined from (6) alone in the situation where $\tau$ can be chosen arbitrarily.

In this article, we prove a robust metrological bound in the form of $T \sim C\delta^{-1}$, which does not postulate the small $\delta$ limit. The constant $C$ only depends on the Hamiltonian model $H_\theta$ that serves as an information resource in a closed quantum system. Although we focus on bounds on the estimation time, the result is indeed applicable to other types of metrological bounds for a given evolution time fixed. For example, the results can be used to bound the energy amplification of the Hamiltonian, since quantum evolution is based on the product of time and energy. Another important corollary is on the number of channels or photons. As we explain in Sec. II, a lower bound on the time also sets a lower bound on the number of channels. As the contraposition, an upper bound on the required time can be derived from the upper bound on the number of channels, which we show in Sec. IV.

The rest of this article is organized as follows: In Sec. II, we describe the basic setting of the Hamiltonian estimation problem. In Sec. III, we compute a lower bound on the total time required for the estimation. In Sec. IV, we present an upper bound on the total time by constructing two explicit procedures, one for the sequential and the other for the parallel scheme.
II. PRELIMINARIES

A. Hamiltonian Model

We consider a Hamiltonian $H_\theta$ in a $d$-level system $\mathcal{H}_D = \mathbb{C}^d$ that depends linearly on an unknown parameter $\theta \in \mathbb{R}^m$:

$$H_\theta = \sum_{j=1}^{m} \theta^j X_j, \quad \theta = (\theta^1, \ldots, \theta^m).$$

(7)

Here the operators $X_1, \ldots, X_m$ are Hermitian, traceless, and satisfy $\text{Tr} X_j X_k = \delta_{jk}$. In other words, \{X_1, \ldots, X_m\} forms an orthonormal basis of an $m$-dimensional subspace of $\text{su}(d)$ with respect to the Hilbert-Schmidt (HS) inner product. The Hamiltonian model is specified by this subspace, in which the Hamiltonian is assumed to exist. In particular, we have a $(d^2 - 1)$-dimensional model with all possible Hamiltonians, which we refer to as the full model. Another example is specified by $d - 1$ simultaneously diagonalized matrices, which we call the phase estimation model.

B. Estimation Procedure

In quantum metrology, one needs to generate a probe state $|q_\theta\rangle$ through an unknown Hamiltonian $H_\theta$. First, we consider a scheme depicted in Fig. 1 (a), in which $r$ channels are driven by the Hamiltonian $H_\theta$. Since the total system consists of $r$ copies of driven systems plus an ancilla, the total Hilbert space can be written as $\mathcal{H}_{\text{tot}} = \mathcal{H}_D^\otimes r \otimes \mathcal{H}_A$.

The probe state after an evolution time $\tau$ can be written as

$$|q_\theta\rangle = U_r(\tau)|q_{\text{ini}}\rangle, \quad U_r(t) = (e^{-itH_\theta})^\otimes r \otimes I_A,$$

(8)

where $I_A$ is the identity operator on $\mathcal{H}_A$. Finally, the parameter $\theta$ is estimated by measuring $N$ copies of probe states: $|q_\theta\rangle^\otimes N$. The total time resource for this procedure amounts to $T = Nr\tau$, which we want to minimize.

We may introduce feedback control in the estimation procedure. Though it may involve general Kraus measurements or nonunitary evolutions, any feedback control can be represented by a series of unitary transformations [40]. We denote these unitary operations by $V_1, V_2, \ldots, V_n$ on the total Hilbert space $\mathcal{H}_{\text{tot}} = \mathcal{H}_D \otimes \mathcal{H}_A$. These operations are performed according to a given procedure as shown in Fig. 1 (b).
FIG. 1: Schematic diagrams for Hamiltonian estimation (a) for an \( r \)-channel procedure without feedback control and (b) for a 1-channel procedure with feedback control \( V_1, \ldots, V_n \). The initial state \( |q_{\text{ini}}\rangle \) is possibly entangled between the driven systems (\( \mathcal{H}_D \)) and the ancillae (\( \mathcal{H}_A \)). The driven systems evolve according to the Hamiltonian \( H_\theta \) during a sequence of time intervals indicated below the diagram, after which a probe state is obtained. In the course of estimation, \( N \) copies of the quantum states are measured.

If we denote by \( t_0, t_1, \ldots, t_n \) the time intervals between the unitary operations, the probe state can be written as

\[
|q_\theta\rangle = U_1(t_n)V_n \cdots U_1(t_1)V_1U_1(t_0)|q_{\text{ini}}\rangle.
\]  

(9)

We note that an \( r \)-channel procedure with evolution time \( \tau \) can be simulated by a 1-channel feedback procedure with evolution time \( r\tau \). This can be done as follows: we consider an \( r \)-channel procedure with driven channels \( D_1, D_2, \ldots, D_r \). In the corresponding 1-channel procedure, we regard \( D_1 \) as the only driven channel and include the rest in the ancillary system. Let \( V \) be a unitary operator that permutes the channels such that \( D_1, D_2, \ldots, D_r \) are substituted by \( D_r, D_1, \ldots, D_{r-1} \), respectively. If one applies the operator \( V \) after every interval of time \( t \), each system \( D_1, \ldots, D_r \) will be driven by \( H_\theta \) after the \( r \)th interval. In this way, the time-\( t \) evolution over \( r \) channels can be reduced to the time-\( rt \) evolution over one channel.

C. Success Criterion

An estimation procedure ultimately ends by yielding an estimator \( \theta^* \) for the vector parameter \( \theta \) to be estimated. The estimation is successful when the Euclid norm \( \|\theta^* - \theta\| \) does
not exceed a small value $\delta$. We require that the probability of failure be sufficiently small, say,

$$
P\left[\|\theta^* - \theta\| > \delta\right] \leq p_{\text{crit}} := 0.05.
$$

(10)

We note that the actual value of $p_{\text{crit}}$ is not important as long as it lies between zero and $1/2$, since the dependence on $p_{\text{crit}}$ is known to be at most logarithmic [41].

Furthermore, we set the search radius $E > 0$ such that $\|\theta\| \leq E$ is presumed for the vector parameter $\theta \in \mathbb{R}^m$. The condition (10) must then be satisfied for all $\theta$ within that radius. We will see later that the metrological bound does not depend on the search radius $E$, as long as $E$ is finite and the ratio $\delta/E$ is kept below some constant (e.g. $\delta/E \leq 1/5$).

III. THE CRAMÉR-RAO BOUND ON HAMILTONIAN ESTIMATION

The primary concern of this article is the minimal time resource $T$ required for the successful estimation of the Hamiltonian $H_\theta$. First, we derive a lower bound on $T$ that must hold for any kind of estimation procedures. As discussed in Sec. II B, any multiple-channel procedure can be reduced to a 1-channel feedback procedure with the same time resource. Hence, we only consider the latter case in this section.

We recall that a procedure consists of two types of process: a continuous process governed by the $\theta$-dependent Hamiltonian $H_\theta$ and a discrete process governed by the $\theta$-independent unitary operator $U_k$. Moreover, the QFI is by definition invariant under the discrete process, which is natural since the unitary transformation does not convey any information about $\theta$. Hence the QFI at time $t$, which we denote by $J(\theta,t)$, can increase only in the continuous process, leading to the following theorem.

**Theorem 1.** Let us define an operator $X$ on $\mathcal{H}_D$ by

$$
X = \sum_{j=1}^{m} (X_j)^2.
$$

Then, the QFI $J(\theta,t)$ satisfies $\text{Tr} J(\theta,t) \leq 4c t^2$ with $c = \|X\|$ being the operator norm of $X$. In particular, the trace of the QFI for the final state is at most $4c \tau^2$.

**Proof.** We omit the identity operator on the ancillary system since it does not affect the claim of the theorem. Let us define a matrix $G(\theta,t)$ by

$$
[G(\theta,t)]_{jj'} = \langle q_\theta(t) | X_j X_{j'} | q_\theta(t) \rangle.
$$

(11)
In Appendix A, we show that the growth rate of the QFI is bounded as
\[ \frac{\partial}{\partial t} \sqrt{\text{Tr} J(\theta, t)} \leq \sqrt{4 \text{Tr} G(\theta, t)}. \] (12)

Since \( \text{Tr} G(\theta, t) = \langle q_\theta(t) | X | q_\theta(t) \rangle \leq c \), the growth rate of the QFI is bounded regardless of the procedure as
\[ \frac{\partial}{\partial t} \sqrt{\text{Tr} J(\theta, t)} \leq \sqrt{4c}. \] (13)

In addition, we have \( J(\theta, t=0) = 0 \) since the initial state does not depend on \( \theta \). Equation (13) can thus be integrated, giving \( \sqrt{\text{Tr} J(\theta, t)} \leq \sqrt{4ct} \), from which the theorem follows. \( \square \)

Let us say that the Hamiltonian model is spherical if an additional condition
\[ X = \sum_{j=1}^{m} (X_j)^2 \propto I \] (14)
is satisfied. We note that both the full model and the phase estimation model meet this condition. The proportionality constant is determined from \( X = (m/d) I \), which is confirmed by comparing the trace. This results in an upper bound for QFI:
\[ \text{Tr} J(\theta) \leq \frac{4m}{d} \tau^2. \] (15)

The spherical condition can be interpreted as follows: Given that the parameter \( \theta \) has a prior distribution with the spherical symmetry in \( \mathbb{R}^m \), the model is spherical if and only if the prior expectation value of \( (H_\theta)^2 \) is proportional to the identity operator.

Now, a lower bound on \( T \) can then be derived from (15); the QCR inequality implies
\[ \text{Tr}[V(\theta)] \geq \frac{\text{Tr}[J(\theta)^{-1}]}{N} \geq \frac{m^2}{N \text{Tr}[J(\theta)]} \geq \frac{md}{4N\tau^2}, \] (16)
where the second inequality follows from the Schwartz inequality \( \text{Tr}[J(\theta)^{-1}] \text{Tr}[J(\theta)] \geq (\text{Tr} I)^2 = m^2 \). Since the successful estimation requires \( \text{Tr}[V(\theta)] \sim \delta^2 \), one obtains a trade-off relation
\[ N\tau^2 \gtrsim md/\delta^2. \] (17)

We combine this relation with \( N \geq 1 \) to obtain a lower bound on \( T \):
\[ T = N\tau = N^{1/2}(N\tau^2)^{1/2} \]
\[ \gtrsim 1(md/\delta^2)^{1/2} = O((md)^{1/2}/\delta). \] (19)
Although some careful treatment is necessary for the general situation with a biased estimator, the lower bound \([19]\) is unchanged up to a constant factor. The detail is described in Appendix \([3]\).

A lower bound for a nonspherical model is similarly obtained from Theorem \([1]\) but depends on \(c = \|X\|:\)

\[
T \gtrsim O(c^{1/2} d/\delta^2).
\]  

For instance, we consider a model with \(m = d - 1\) parameters:

\[
X_j = \frac{1}{\sqrt{2}} (|e_j\rangle\langle e_d| + |e_d\rangle\langle e_j|) \quad (1 \leq j \leq d - 1),
\]  

where \(\{|e_1\rangle, \ldots, |e_d\rangle\}\) is the basis of \(\mathbb{C}^d\). With this model, we find \(c = \frac{d-1}{2}\), which becomes much larger than \(m/d = O(1)\) in the large \(d\) limit. Furthermore, if we let the initial state \(|q_{\text{ini}}\rangle = |d\rangle\) freely evolve by \(H_\theta\), the inequality \(\text{Tr} J(\theta, t) \leq 4c\tau^2\) in Theorem \([1]\) can be saturated. Since the Fisher information \(\text{Tr} J(\theta) = 4c\tau^2 = 2(d-1)\tau^2\) violates the inequality in \([15]\), the same lower bound as \([19]\) cannot be derived in the nonspherical case.

### IV. EFFICIENT PROCEDURES FOR HAMILTONIAN ESTIMATION

We need an explicit estimation protocol to establish an upper bound on the time resource \(T\). Noting that the QCR bound is not saturated in general, the reverse inequality \(T \leq O((md)^{1/2}/\delta)\) is not guaranteed. In fact, the saturation of \([19]\) requires that the QCR inequality be saturated [i.e. \(N\tau^2 = O(md/\delta^2)\)] and that the number of samples be small [i.e. \(N = O(1)\)]. We need to control the quantum state \(|q_\theta\rangle\) for all \(\theta\) to satisfy these two competing requirements simultaneously. Such a control is rather difficult because the dependency of \(|q_\theta\rangle\) on \(\theta\) becomes generally chaotic with the large evolution time \(\tau = O((md)^{1/2}/\delta)\). At present, we find the lower bound to be saturated in the simplest cases with \(m = O(1)\), as we discuss in Sec. \([\nabla]\).

For a generic situation, we obtain a looser but rigorous upper bound. First, we present an \(O(mdE/\delta^2)\) procedure in the 1-channel scheme. After that, we improve the procedure in two distinct ways in order to achieve \(O(md/\delta)\): by introducing the adaptive feedback and by increasing the number of channels.
A. The one-channel scheme

First, we consider a simplest scheme corresponding to Fig. 1 (a) with \( r = 1 \). As an input state, we consider the maximally entangled state (MES) \( \ket{\Phi} \) associated with the Hilbert space \( \mathcal{H}_D \). The MES involves an ancillary Hilbert space \( \mathcal{H}_A \) of the same dimension as \( \mathcal{H}_D \):

\[
\ket{\Phi} = \frac{1}{d^{1/2}} \sum_{j=1}^{d} \ket{e_j} \otimes \ket{e'_j} \in \mathcal{H}_D \otimes \mathcal{H}_A.
\] (22)

Here \( \{\ket{e_j}\} \) and \( \{\ket{e'_j}\} \) are orthonormal bases of \( \mathcal{H}_D \) and \( \mathcal{H}_A \), respectively. After the time-\( \tau \) evolution, we obtain the probe state \( \ket{q_\theta} = (U_\theta \otimes I_A)\ket{\Phi} \) with \( U_\theta = e^{-i\tau H_\theta} \).

The crucial point is that, for the full model, the manifold formed by the probe states \( \{\ket{q_\theta} | \theta \in \mathbb{R}^m\} \) is geometrically similar to the Lie group \( \text{SU}(d) \). The QFI is therefore in proportion to the invariant metric of the Lie group:

\[
[J(\theta)]_{jk} = \frac{4}{d} \text{Tr} \left[ \frac{\partial U_\theta^\dagger}{\partial \theta^j} \frac{\partial U_\theta}{\partial \theta^k} \right].
\] (23)

Especially, the QFI at \( \theta = 0 \) satisfies \( [J(0)]_{jk} = (4/d)\tau^2 \delta_{jk} \), and hence reaches the upper bound in (15) when the Hamiltonian model is spherical.

As long as the QFI is concerned, larger \( \tau \) seems to be better for the estimation. In general, however, this is not the case. For example, suppose that the Hamiltonian model contains a Hamiltonian of the form \( H = \ket{\psi}\bra{\psi} - \frac{1}{d} I \), where \( \ket{\psi} \) is a normalized vector. Then, two Hamiltonians \( H_{\pm \theta} = \pm (\pi/\tau) X \) cannot be distinguished from each other, since they yield the same probe states. Such a situation can occur when the evolution time \( \tau \) is larger than \( O(1/E) \).

When \( \tau \leq O(1/E) \), on the other hand, the QCR bound \( \delta^2 \lesssim O(md/N\tau^2) \) can be nonasymptotically saturated. To see this, one projects the probe state \( \ket{q_\theta} \) to the \( (m+1) \)-dimensional Hilbert space spanned by \( \ket{\Phi}, X_1 \ket{\Phi}, \ldots, X_m \ket{\Phi} \). Although this projection involves a certain postselection, the probability of failure is at most \( O(1) \) and contributes only to a constant factor. After the projection, one conducts quantum tomography on \( N \) copies of the postselected state \( \ket{\tilde{q}_\theta} \). The efficiency of the tomography can be computed from the following quantity:

\[
I_\delta = \inf \left\{ I(q_\theta', \tilde{q}_\theta) \mid \|\theta\|, \|\theta'\| \leq E, \|\theta' - \theta\| \geq \delta \right\},
\] (24)

where \( I(q', q) = \sqrt{1 - |\langle q | q'\rangle|^2} \) forms a distance between \( \ket{q} \) and \( \ket{q'} \), which is often referred to as the infidelity. In this article, we call the quantity \( I_\delta \) as the \( \delta \)-resolution.
We suppose that $N$ copies of postselected states are given. According to a study on the pure-state quantum tomography, the $(m+1)$-dimensional quantum state can be estimated such that the expected square infidelity is $\frac{m}{N+m}$ [42]. Hence $N$ needs to be so large that $I_2^2 = O(m/N)$ holds. The saturation of the trade-off relation (17) then reduces to

$$I_2^2 = O(\tau^2 \delta^2 / d),$$

which corresponds to the bound on the QFI in (15). As we show in Appendix C, this condition is satisfied when $\tau E$ is small but stays at $O(1)$ with respect to $m$ and $d$. Since one needs $N = O(mdE^2/\delta^2)$ copies of probe states for this case, the time resource consumed by this procedure amounts to

$$T = N\tau = O(mdE/\delta^2).$$

### B. The one-channel scheme with adaptive feedback

As long as the error tolerance $\delta$ is concerned, the last procedure is analogous to the classical statistics: $T \propto 1/\delta^2$. We would like to improve this procedure to $T \propto 1/\delta$ by means of quantum enhancement.

Here, we introduce the adaptive feedback control [39]. One simulates application of the external field $V = -H_{\theta^*}$, where $\theta^*$ is an estimated value of $\theta$ estimated from the preceding measurements. We can use the Suzuki-Trotter decomposition [43] to approximate this external field with sufficiently many unitary operations.

Since the system is driven with the Hamiltonian $H = H_{\theta} - H_{\theta^*} = H_{\theta - \theta^*}$, the parameter $\theta$ would be effectively replaced by $\theta - \theta^*$. Moreover, suppose that the estimator $\theta^*$ satisfies $||\theta - \theta^*|| \leq E'$ with high probability; then the search radius $E$ would also be replaced by an effective radius $E'$.

First, we fix a sufficiently large integer $n_0$ such that $E \leq 2^{n_0}\delta$. One computes a sequence of estimators $\theta_{n_0}^*, \theta_{n_0-1}^*, \ldots, \theta_1^*, \theta_0^*$ in this order, starting with $\theta_{n_0} = 0$. The estimators are required to satisfy the following condition:

$$\mathbb{P}[||\theta_n^* - \theta|| \leq 2^n \delta] \geq 1 - p_{\text{crit}}/2^n$$

(27)

for any $||\theta|| \leq E$. The condition is obviously met for $n = n_0$, and is recursively satisfied for $n = n_0, \ldots, 1$ if the conditional probability

$$\mathbb{P}[||\theta_{n-1}^* - \theta|| \leq 2^{n-1} \delta \bigg| ||\theta_n^* - \theta|| \leq 2^n \delta]$$

(28)
FIG. 2: Schematic diagrams of (a) the one-channel (sequential) procedure with adaptive feedback and (b) the many-channel (parallel) procedure. The preceding estimator $\theta^*_n$ is used in different places: in (a), it appears in the time evolution; in (b), it is used only in the measurement process. Due to this difference, (a) involves adaptive feedback and (b) does not.

is no less than $1 - p_{\text{crit}}/2^n$. In Fig. 2 (a) we show how the adaptive feedback works.

**Theorem 2.** The time resource required for the above procedure is $T = O(md/\delta)$.

**Proof.** The required time resource for the estimator $\theta^*_{n-1}$ can be computed in the same way as in Sec. [IVA] with $E$ and $\delta$ replaced by $E' = 2^n \delta$ and $E'/2 = 2^{n-1} \delta$, respectively. In addition, we need to take into account the fact that the critical rate $p_{\text{crit}}/2^n$ becomes exponentially small with increasing $n$. By the Chernoff bound, this requires $O(n)$ times more probe states than the case with the critical rate $p_{\text{crit}}$. As a result, we obtain $T_n = O(n)O(mdE'(E'/2)^{-2}) = O(mdE'/2^n\delta)$ for $n \geq 1$. Since both $T_1$ and $\sum_{n=1}^{\infty} T_n$ are of the order of $O(md/\delta)$, the total time resource is also $T = T_{n_0} + \cdots + T_1 = O(md/\delta)$. 

This time scale can be obtained from (26) by setting $E$ to be comparable with $\delta$. We see that this result is independent of the initial search radius $E$. In fact, the time resource is consumed mostly in the regime $E \sim \delta$, since the estimation proceeds fast when the energy scale is large.

C. The many-channel scheme

Next, we consider another procedure with a sufficiently large number of channels, but without feedback control.
The many-channel procedure that we present here can be regarded as a modified version of the one-channel procedure with adaptive feedback. Again, one takes a positive integer $n_0$ such that $E \leq 2^n \delta$ and computes a sequence of estimators $\theta_{n_0}^*, \ldots, \theta_1^*, \theta_0^*$ in a row. The schematic diagram is shown in Fig. 2(b). The probe state for the estimator $\theta_{n-1}^*$ is entangled between $r = 2^{n_0} - n d$ channels under the evolution during the time interval $\tau = O(1/2^n \delta)$.

We note that the evolution is not adaptive, and that the time $\tau$ is independent of $n$. Hence the time evolution for all $n$ can be conducted in parallel, which requires as many channels as $R = O(Ed/\delta)$. In contrast, the measurement of the probe states for $\theta_{n-1}^*$ depends on the preceding estimator $\theta_n^*$.

In the following, we will show that $O(mn)$ copies of probe states are required for the estimator $\theta_{n-1}^*$. Given this statement is true, the total time resource is $T_n = O(mnd/2^n \delta)$, leading to the same result as the one-channel adaptive procedure.

Following Imai and Fujiwara [38], we take the symmetric subspace $\mathcal{H}_{D,r} = \bigotimes_{sym} \mathcal{H}_D$ of the tensor product space $\mathcal{H}_D^\otimes r$. One begins with the MES $|\Phi_r\rangle$ associated with the Hilbert space $\mathcal{H}_{D,r}$. The probe state $|q_{\theta}\rangle$ can be written as

$$|q_{\theta}\rangle = (e^{-i \tau H_{\theta}})^{\otimes r}|\Phi_r\rangle,$$

where we have omitted the identity operator on the ancillary Hilbert space.

Now we assume $\|\theta - \theta_n^*\| \leq E' := 2^n \delta$, and attempt to construct the next estimator within error $E'/2$. On the measurement, the quantum state $|q_{\theta}\rangle$ is first transformed by the unitary matrix $(e^{i \tau q_{\theta}})^{\otimes r}$ with $H^* = H_{\theta_n^*}^*$. The resultant quantum state is

$$|q'_{\theta}\rangle = (e^{i \tau q_{\theta}})^{\otimes r}(e^{-i \tau H_{\theta}})^{\otimes r}|\Phi_r\rangle = (e^{i \tau H_{\theta}} e^{-i \tau H_{\theta}})^{\otimes r}|\Phi_r\rangle.$$

By the Magnus expansion, we have the operator $M_{\theta} = M_{\theta}^*(\tau) \in \text{su}(d)$ satisfying $e^{-i \tau M_{\theta}} = e^{i \tau H_{\theta}} e^{-i \tau H_{\theta}}$. The quantum state $|q'_{\theta}\rangle$ can then be written as

$$|q'_{\theta}\rangle = (e^{-i \tau M_{\theta}})^{\otimes r}|\Phi_r\rangle = e^{-i \tau \{M_{\theta}\}_r}|\Phi_r\rangle.$$

Here we denote by $\{A\}_r$ the $r$-fold collective operator for $A$; it can be defined as $\{A\}_r = P \sum_{j=1}^r A^{(j)}$, where $A^{(j)}$ is the operator $A$ acting on the $j$th Hilbert space and $P$ the projec-
tion onto $\mathcal{H}_{D,r}$. We approximate this state as

$$
|q'_\theta\rangle \approx (I - i\tau \{M_\theta\}_r)|\Phi_r\rangle \\
\approx (I - i\tau \{H_\theta - H^*_\theta\}_r)|\Phi_r\rangle,
$$

(32)

so that we may regard $|q'_\theta\rangle$ to be in an $(m + 1)$-dimensional subspace spanned by

$$
|\Phi_r\rangle, \{X_1\}_r|\Phi_r\rangle, \ldots, \{X_m\}_r|\Phi_r\rangle.
$$

(33)

Therefore, we obtain an $(m + 1)$-dimensional state $|\bar{q}_\theta\rangle$ after projecting $|q'_\theta\rangle$ onto this subspace. The $\delta$-resolution of $|\bar{q}_\theta\rangle$ defined in (24) is now given as

$$
I_{E'/2} = O(\tau E) + O(\tau^2 E^2),
$$

(34)

Therefore, when $\tau$ is sufficiently small but of the order of $O(1/E)$, $O(mn)$ copies of quantum states are sufficient.

In Appendix D, we present the proof of (34). We emphasize that the approximation in (32) is valid only when $\tau E' \leq O(d/r)$ holds, which essentially determines the number of necessary channels.

We thus find that, when there are as many channels as $R = O(Ed/\delta)$, we can achieve the upper bound $T = O(md/\delta)$ by entanglement and without feedback control. The procedure does not improve any longer by further adaptive feedback, since it already simulates adaptive feedback control by adaptive measurement. It can also be inferred that we need adaptive feedback control with fewer than $R$ channels, since the initial search radius $E$ is too large for this case. When more than $R$ channels are available, on the other hand, we can expand the search radius without changing the amount of time resource.

V. COMPARISON WITH PREVIOUS RESULTS

Thus far we showed that, in the case of a spherical Hamiltonian model, the bounds are given as

$$
O((md)^{1/2}/\delta) \leq T \leq O(md/\delta),
$$

(35)

where $\delta$ is the tolerated error in the estimation. In this section, we compare the bounds in (35) with the existing results in quantum metrology. We see that the presented metrological bounds are consistent with the previous results and that they give more general insight into the theory.
A. Finite-dimensional metrology

As long as a fixed system is concerned, we can set $m$ and $d$ to be of the order of unity. In this case, the bounds simply reduce to $T = O(1/\delta)$, which is the original Heisenberg limit.

B. Hamiltonian tomography

Estimation of an arbitrary Hamiltonian on $\mathbb{C}^n$, or the Hamiltonian tomography, is treated by the full model with $m = d^2 - 1$. The metrological bounds are therefore $O(d^{2/3}/\delta) \leq T \leq O(d^3/\delta)$.

Reference [32] gives the QCR bound $\delta^2 \geq O(d^3/\tau^2)$ for one probe state, which corresponds to the lower bound. If we consider the QCR bound $\delta^2 \geq O(d^3/N\tau^2)$ for $N$ probe states and regard the restriction on the evolution time $\tau \leq O(1/\delta)$, we obtain $T = N\tau \geq O(d^2/\delta)$, the saturation of which corresponds to the upper bound.

C. Multiple phase estimation

The phase estimation model in $\mathbb{C}^d$ is generated by $m = d-1$ diagonal matrices, from which the bounds are $O(d/\delta) \leq T \leq O(d^2/\delta)$. In a practical situation, however, the parameters are phase shifts of $d-1$ independent modes relative to a reference mode $|0\rangle$ [25, 26]. This assumes the generators of the form

$$Z_j = |j\rangle\langle j| - |0\rangle\langle 0| \quad (j = 1, \ldots, d - 1). \tag{36}$$

Since these generators are not orthonormal in $\text{su}(d)$, the error $\delta'$ in the $Z$ basis is generally different from the error $\delta$ in the orthonormal basis. They can be related as $\delta \leq \delta' \leq d\delta$, where the factor $d$ comes from the fact that the $Z$ basis uses the reference mode $O(d)$ times more than the others. As a result, the metrological bounds change into $O(d/\delta') \leq T \leq O(d^3/\delta')$. The QCR bound corresponding to this lower bound is also seen in Ref. [26].

D. Few-parameter estimation

When we consider a spherical model with $m = O(1)$ constant with respect to $d$, the lower bound $T = O(d^{1/2}/\delta)$ can be saturated. In fact, the operator norm of $H_0$ with $\|\theta\| \leq E$ is
at most $Em^{1/2}/d^{1/2}$ since

\[ \|H_\theta\|^2 = \|H_\theta^2\| \leq E^2\|X\| = E^2(m/d). \]  

(37)

Therefore, the evolution time in the one-channel scheme (Sec. IV A) can be set to be $\tau = O(d^{1/2}/m^{1/2}E)$, with which the number of probe states reduces to $N = O(mE/\delta^2)$. Therefore, the required time for the one-channel scheme is $T = N\tau = O(m^{3/2}d^{1/2}E/\delta^2)$, which reduces to $T = O(m^{3/2}d^{1/2}/\delta)$ by using adaptive feedback. This new upper bound is tighter than the general one, $T = O(md/\delta)$, when $m$ is smaller than $O(d)$. Especially in the case $m = O(1)$, this upper bound $T = O(\sqrt{d}/\delta)$ is found to coincide with the lower bound.

VI. CONCLUSION

In this paper, we have investigated the estimation of an $m$-parameter Hamiltonian in a $d$-level system, and derived rigorous upper and lower bounds (35) on the time resource $T$. We note that it is possible to calculate the model-independent constant factors.

The upper bound in (35) is consistent with Yuan’s result on the full model [32] and the multiple phase model [17, 26] where the evolution time $\tau$ is set to be $O(1)$. We present two procedures to achieve this upper bound: the one-channel procedure with adaptive feedback and the many-channel procedure without it. This result is different from the work by Yuan, where the former scheme is claimed to outperform the latter scheme by a factor of $O(d)$. The difference arises from the precondition that the search radius $E$ and the tolerated error $\delta$ are given independently of the number of channels.

The lower bound in (35) is by far smaller than the upper bound, and no concrete procedure corresponding to this lower bound has been found except for the case of $m = O(1)$. This bound is also related to Grover’s algorithm, which requires an $O(d^{1/2})$ time for the estimation of a discrete parameter in a $d$-level Hamiltonian. In fact, the derivation of this bound is closely related to that of the optimal time in Grover’s search problem [22, 36]. It remains an open question whether any actual procedure can achieve $T = O((md)^{1/2}/\delta)$ because no corresponding procedure has been constructed.
ACKNOWLEDGEMENT

We gratefully acknowledge Tomohiro Shitara for a number of critical comments. This work was supported by KAKENHI Grant No. 26287088 from the Japan Society for the Promotion of Science (JSPS), a Grant-in-Aid for Scientific Research on Innovative Areas “Topological Materials Science” (KAKENHI Grant No. 15H05855) from the JSPS, and the Photon Frontier Network Program from MEXT of Japan. N. K. acknowledges support from the Advanced Leading Graduate Course for Photon Science (ALPS) of JSPS.

Appendix A: The proof of Theorem 1

Let us simply write $|q_\theta\rangle$ for $|q_\theta(t)\rangle$, and define $F_j$ and $G_j$ as

$$F_j := \frac{1}{4}[J(\theta, t)]_{jj} = \langle \partial_j q_\theta [1 - |q_\theta\rangle\langle q_\theta|] |\partial_j q_\theta\rangle,$$

$$G_j := [G(\theta, t)]_{jj} = \langle q_\theta|X_j^2|q_\theta\rangle. \tag{A1}$$

The time evolutions of $|q_\theta\rangle$ and $|\partial_j q_\theta\rangle$ by the Hamiltonian $H_\theta$ are governed by

$$i\frac{\partial}{\partial t}|q_\theta\rangle = H_\theta|q_\theta\rangle, \tag{A2}$$

$$i\frac{\partial}{\partial t}|\partial_j q_\theta\rangle = \partial_j(H_\theta|q_\theta\rangle) = X_j|q_\theta\rangle + H_\theta|\partial_j q_\theta\rangle. \tag{A3}$$

Hence the term involving $H_\theta$ is canceled upon differentiation $\frac{\partial F_j}{\partial t}$:

$$\frac{\partial F_j}{\partial t} = 2\text{Im}\langle \partial_j q_\theta [1 - |q_\theta\rangle\langle q_\theta|] X_j|q_\theta\rangle \tag{A4}$$

$$\leq 2\| [1 - |q_\theta\rangle\langle q_\theta|] |\partial_j q_\theta\rangle \| \| X_j|q_\theta\rangle \| = 2(F_j G_j)^{1/2},$$

where the inequality follows from the Schwartz inequality. If we employ the Schwartz inequality again, we obtain

$$\frac{\partial}{\partial t} \sum_{j=1}^m F_j \leq 2 \sum_{j=1}^m (F_j G_j)^{1/2} \leq 2 \left( \sum_{j=1}^m F_j \sum_{j'=1}^m G_{j'} \right)^{1/2} \tag{A5}$$

$$\implies \frac{\partial}{\partial t} \left( \sum_{j=1}^m F_j \right)^{1/2} \leq \left( \sum_{j=1}^m G_j \right)^{1/2},$$

the latter of which is equivalent to the claim.
Appendix B: A uniform Cramer-Rao bound for a biased estimator

The bias of an estimator \( \theta^\ast \) is defined as \( b(\theta) = \mathbb{E}_\theta[\theta^\ast] - \theta \), where \( \mathbb{E}_\theta \) denotes the expectation value with respect to the true parameter \( \theta \). The Cramér-Rao inequality (1) assumes that the estimator is unbiased, namely \( b(\theta) = 0 \). In a general situation, the inequality must be modified as

\[
V(\theta) \geq N^{-1} [I + D(\theta)] J(\theta)^{-1} [I + D(\theta)]^T,
\]

where \( D(\theta) \) is a matrix defined as \( [D(\theta)]_{jk} = \partial_j b_k(\theta) \). By this inequality, the evaluation of the variance in (16) can be modified as

\[
\text{Tr}[V(\theta)] \geq \frac{(\text{Tr}[I + D(\theta)])^2}{N \text{Tr}[J(\theta)]}.
\]

Although a nonzero bias may decrease the variance \( \text{Tr}[V(\theta)] \), it may increase the total error as well: \( \delta^2 \sim \text{Tr}[V(\theta)] + \|b(\theta)\|^2 \). If we assume \( p_{\text{crit}} = 0.05 \) and \( \delta/E \leq 1/5 \), for simplicity, \( \text{Tr}[V(\theta)] + \|b(\theta)\|^2 \leq 6\delta^2 \) is necessary for a successful estimator. Therefore the bias is under constraint \( \|b(\theta)\| \leq \sqrt{6}\delta \) for all \( \theta \) within radius \( E \).

We would like to show that, under this constraint, \( \text{Tr}[I + D(\theta)] > \frac{m}{2} \) holds for some \( \theta \). This will lead to the conclusion that the inequality (16) is modified only by the constant factor of \( \frac{1}{4} \) by introducing a biased estimator.

Let us assume the contrary, that is, \( \text{Tr}[I + D(\theta)] \leq \frac{m}{2} \) for all \( \theta \). This assumption can be rewritten as \( \nabla \cdot b(\theta) = \text{Tr} D(\theta) \leq -\frac{m}{2} \), which upon integration becomes

\[
\int_{\|\theta\| \leq 5\delta} \nabla \cdot b(\theta) d\theta \leq -\frac{m}{2} \Pi_m (5\delta)^m.
\]

where we denote by \( \Pi_m \) the volume of a unit ball in \( \mathbb{R}^m \). On the other hand, by the constraint \( \|b(\theta)\| \leq \sqrt{6}\delta \), we have

\[
\left| \int_{\|\theta\| = 5\delta} b(\theta) \cdot dn(\theta) \right| \leq \sqrt{6}\delta m \Pi_m (5\delta)^{m-1}.
\]

Now, the integrals in (B3) and (B4) are equal by the divergence theorem. This leads to the contradiction since \( \frac{m}{2} > \sqrt{6} \).
Appendix C: A rigorous evaluation of the $\delta$-resolution

We show that the $\delta$-resolution defined in (21) satisfies $I_\delta^2 = O(\tau^2\delta^2/d)$. First, we examine the full model, where we do not perform the postselection. We take two candidates $\theta$ and $\theta'$, which satisfy the condition in (24).

It is straightforward that the MES defined by (22) satisfies $\langle \Phi|A|\Phi \rangle = \frac{1}{d}\text{Tr} A$ for an arbitrary operator $A$ on $\mathcal{H}_D$. Hence the infidelity between $|q\rangle$ and $|q'\rangle$ can be described as

$$1 - |\langle q|A|q' \rangle|^2 = 1 - \frac{1}{d^2}\left|\text{Tr} e^{i\tau H\theta'}e^{-i\tau H_0}\right|^2.$$  \hspace{1cm} (C1)

It is known in the context of Loschmidt echo that, for sufficiently small $\tau$, the right-hand side of (C1) can be approximated by $\frac{\tau^2}{2d} \text{Tr}(H_{\theta'} - H_0)^2 = \frac{\tau^2}{2d} \|\theta' - \theta\|^2$. Therefore, the estimation $I_\delta^2 = O(\tau^2\delta^2/d)$ is correct as long as this short-term approximation is valid for $\tau = O(1/E)$.

We recall that the postselection subspace is spanned by $|\Phi\rangle, X_1, |\Phi\rangle, \ldots, X_m|\Phi\rangle$. We denote by $P$ the projection operator onto this subspace. The infidelity between two postselected space $|\bar{q}\rangle, |\bar{q}'\rangle$ can be written as

$$1 - |\langle \bar{q}|P|\bar{q}' \rangle|^2 = \frac{\langle q_0|P|q_0\rangle \langle q_0|P|q_0\rangle - \langle q_0|P|q_0\rangle}{\langle q_0|P|q_0\rangle \langle q_0|P|q_0\rangle}$$

$$\geq |\langle q_0|P|q_0\rangle \langle q_0|P|q_0\rangle - \langle q_0|P|q_0\rangle|^2$$

$$= |\langle q_0|P|q_0\rangle \langle q_0|P|q_0\rangle - \langle q_0|P|q_0\rangle|^2,$$  \hspace{1cm} (C2)

where $|\Delta q\rangle = |q_0\rangle - |q_0\rangle$. Therefore, it suffices to show that $\langle \Delta q|P|\Delta q\rangle$ is at least $O(\tau^2\delta^2/d)$, while the last term in $\langle q_0|P|q_0\rangle$ is negligible.

First, we note that the equality $\langle \Phi|A|\Phi \rangle = \langle \text{Tr} A \rangle/d$ can be applied only when $A$ belongs to $L(\mathcal{H}_D)$, the operator space on $\mathcal{H}_D$. Since $P$ is a projection operator on $\mathcal{H}_D \otimes \mathcal{H}_A$, we need a special treatment with the postselection. A superoperator $S$ on $L(\mathcal{H}_D)$ is defined as

$$S(A) = \frac{1}{d}(\text{Tr} A)I + \sum_{j=1}^m (\text{Tr} AX_j)X_j.$$  \hspace{1cm} (C3)

This superoperator is a projection operator on $L(\mathcal{H}_D)$ equipped with the HS inner product. Then $PA|\Phi\rangle = S(A)|\Phi\rangle$ holds for an arbitrary $A \in L(\mathcal{H}_D)$, which is a great convenience.

Next, we define an operator $B = e^{-i\tau H'}e^{i\tau H} - I$ such that $|\Delta q\rangle = Be^{-i\tau H}|\Phi\rangle$. Then, it follows from the Dyson expansion that

$$B = -i \int_0^\tau ds e^{-i\tau H'}(H' - H)e^{i\tau H} + O(\tau^2\delta^2),$$  \hspace{1cm} (C4)
where the residual term is measured by the trace norm. The Taylor expansion along with $S$ and $B$ defined above yields

$$
\langle q_\theta | P | \Delta q \rangle = \langle \Phi | S(e^{i\tau H})B e^{-i\tau H} | \Phi \rangle = \frac{1}{d} \text{Tr}[e^{-i\tau H} S(e^{i\tau H}) B] \leq O(\tau^2 E \delta / d), \quad (C5)
$$

and

$$
\langle \Delta q | P | \Delta q \rangle = \langle \Phi | S(B e^{-i\tau H}) S(B e^{i\tau H}) | \Phi \rangle = \frac{1}{d} \| S(B e^{-i\tau H}) \|^2_{\text{HS}} \leq \frac{1}{d} \| B e^{-i\tau H} \|^2_{\text{HS}} \leq (\tau^2 \delta^2 / d) (1 + O(\tau E)), \quad (C6)
$$

which completes the evaluation of $I_2^2$. We note that $\| X \|_{\text{HS}} = \text{Tr} | X |^2$ denotes the HS norm.

**Appendix D: The derivation of (34)**

In the proof of (34), we need to evaluate the approximation (32). First, the operator $M_\theta$ in the Magnus expansion can be approximated by $H_\theta' := H_\theta - H^*\tau$, given that both $\tau H^*$ and $\tau H_\theta'$ are small relative to unity. In terms of the HS norm, this is expressed as

$$
\| M_\theta - H_\theta' \|_{\text{HS}} = E' O(\tau E), \quad (D1)
$$

where we take into account $\| \tau H^* \|_{\text{HS}} \leq \tau E = O(1)$ and $\| H_\theta' \|_{\text{HS}} \leq \| H_{\theta-\tau} \|_{\text{HS}} \leq E'$.

To compute the distance between $\{ M_\theta \}_r$ and $\{ H_\theta \}_r$, we introduce the following relation [38]: for an arbitrary $X \in \text{su}(d)$,

$$
\frac{1}{D} \text{Tr}(\{ X \}_r)^2 = F_2 \text{Tr} X^2, \quad F_2 = \frac{r(d+r)}{d(d+1)} \quad (D2)
$$

with $D = \text{dim} \mathcal{H}_{D,r} = \frac{(r+d)(r-d-1)}{d(d-1)}$. This means that the HS norm of $\{ X \}_r$ is $\sqrt{DF_2}$ times that of $X$, leading to the evaluation

$$
\| \tau \{ M_\theta \}_r - \tau \{ H_\theta \}_r \|_{\text{HS}} = \sqrt{DF_2} \tau E' O(\tau E) = \sqrt{D} O(\tau^2 E^2). \quad (D3)
$$

Note that $F_2 = O((r/d)^2)$ for $r \geq d$ and that $E' = (r/d) E$.

Next, we check the approximation of $e^{-i\tau \{ M_\theta \}_r}$. Since $| e^{-i\alpha} - 1 + i\alpha |^2 \leq \frac{1}{4} \alpha^4$ holds for any real number $\alpha$, we have

$$
\| e^{-i\tau \{ M_\theta \}_r} - I + i\tau \{ M_\theta \}_r \|_{\text{HS}}^2 \leq \frac{\tau^4}{4} \text{Tr}(\{ M_\theta \}_r)^4. \quad (D4)
$$
The right-hand side of this inequality can be evaluated similarly to (D2) as
\[
\frac{1}{D} \text{Tr}(\{X\}_r)^4 = F_4 \text{Tr} X^4 + F_{22}(\text{Tr} X^2)^2, \tag{D5}
\]
where the coefficients \( F_4 = \frac{r(r+d)(6r^2+6dr+d^2-d)}{d(d+1)(d+2)d(d+3)} \) and \( F_{22} = \frac{3r(r+d)(r-1)(d+r+1)}{d(d+1)(d+2)(d+3)} \) are both \( O\left(\left(\frac{r}{d}\right)^4\right) \) for \( r \geq d \). Since \( \text{Tr}(M_\theta)^2 \leq E'^2 \) and \( \text{Tr}(M_\theta)^4 \leq E'^4 \), we have
\[
\|e^{-i\tau\{M_\theta\}_r} - I + i\tau\{M_\theta\}_r\|^2_{\text{HS}} \leq D\tau^4 O\left((r/d)^4\right)
\]
\[
= DO(\tau^4 E'^4). \tag{D6}
\]
Finally, we compute the distance between the vectors \( |q'_\theta\rangle = e^{-i\tau\{M_\theta\}_r} |\Phi_r\rangle \) and \( |q''_\theta\rangle = (I - i\tau \{H'_\theta\}_r) |\Phi_r\rangle \) as
\[
\| |q''_\theta\rangle - |q'_\theta\rangle \|^2 = \frac{1}{D} \langle \Phi_r | (e^{-i\tau\{M_\theta\}_r} - I + i\tau \{H'_\theta\}_r)^2 |\Phi_r\rangle
\]
\[
= \frac{1}{D} \| e^{-i\tau\{M_\theta\}_r} - I + i\tau \{H'_\theta\}_r \|^2_{\text{HS}}
\]
\[
\leq O(\tau^4 E'^4), \tag{D7}
\]
where the inequality is obtained by combining (D3) with (D6). The probability of failure in the postselection of \( |q'_\theta\rangle \) is therefore at most \( O(\tau^2 E'^2) \), because \( |q''_\theta\rangle \) belongs to the target subspace. This implies that the postselected state \( |\bar{q}_\theta\rangle \) also satisfies \( \| |q_\theta\rangle - |q''_\theta\rangle \| \leq O(\tau^2 E'^2) \).

Finally, we consider the infidelity between \( |\bar{q}_\theta\rangle \) and \( |\bar{q}_\eta\rangle \) with \( \|\theta - \eta\| \geq E' / 2 \). For \( \phi = \arg \langle \bar{q}_\theta | \bar{q}_\eta \rangle \), we obtain
\[
I(\bar{q}_\theta, \bar{q}_\eta) \geq \frac{1}{\sqrt{2}} \left\| e^{i\phi} |\bar{q}_\theta\rangle - |\bar{q}_\eta\rangle \right\|
\]
\[
= \frac{1}{\sqrt{2}} \left\| e^{i\phi} |q''_\theta\rangle - |q''_\eta\rangle \right\| + O(\tau^2 E'^2). \tag{D8}
\]
We can compute the distance by \( \| e^{i\phi} |q''_\theta\rangle - |q''_\eta\rangle \| \) by using the relation (D2), which turns out to be at least \( O(\tau E) \). Thus the derivation is completed.

[1] A. S. Holevo, Journal of Multivariate Analysis 3, 337 (1973).
[2] C. W. Helstrom, Journal of Statistical Physics 1, 231 (1969).
[3] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Reviews of Modern Physics 74, 145 (2002).
[4] A. Galindo and M. A. Martín-Delgado, Reviews of Modern Physics 74, 347 (2002).
[5] A. Peres and D. R. Terno, Reviews of Modern Physics 76, 93 (2004).
[6] M. Herrero-Collantes and J. C. Garcia-Escartin, Reviews of Modern Physics 89, 015004 (2017).
[7] M. J. Holland and K. Burnett, Physical Review Letters 71, 1355 (1993).
[8] V. Giovannetti, S. Lloyd, and L. Maccone, Science 306, 1330 (2004).
[9] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Nature 450, 393 (2007).
[10] T. Nagata, R. Okamoto, J. L. O’Brien, K. Sasaki, and S. Takeuchi, Science 316, 726 (2007).
[11] B. L. Higgins, D. W. Berry, S. D. Bartlett, M. W. Mitchell, H. M. Wiseman, and G. J. Pryde, New Journal of Physics 11, 073023 (2009).
[12] H. Yonezawa, D. Nakane, T. A. Wheatley, K. Iwasawa, S. Takeda, H. Arao, K. Ohki, K. Tsumura, D. W. Berry, T. C. Ralph, et al., Science 337, 1514 (2012).
[13] C. F. Roos, M. Chwalla, K. Kim, M. Riebe, and R. Blatt, Nature 443, 316 (2006).
[14] M. Vengalattore, J. M. Higbie, S. R. Leslie, J. Guzman, L. E. Sadler, and D. M. Stamper-Kurn, Physical Review Letters 98, 200801 (2007).
[15] G. Tóth and I. Apellaniz, Journal of Physics A 47, 424006 (2014).
[16] M. Napolitano, M. Koschorreck, B. Dubost, N. Behbood, R. J. Sewell, and M. W. Mitchell, in Conference on Lasers and Electro-Optics Europe and 12th European Quantum Electronics Conference (IEEE, 2011).
[17] M. A. Ballester, Physical Review A 70, 032310 (2004).
[18] A. Datta and A. Shaji, Modern Physics Letters B 26, 1230010 (2012).
[19] P. W. Shor, in Proceedings of the 35th Annual Symposium on Fundamentals of Computer Science (IEEE, 1994), pp. 124–134.
[20] L. K. Grover, in Proceedings of the 28th annual ACM symposium on Theory of computing (ACM, 1996), pp. 212–219.
[21] A. Y. Kitaev, arXiv:9511026 (1995).
[22] R. Demkowicz-Dobrzański and M. Markiewicz, Physical Review A 91, 062322 (2015).
[23] M. Szczykulska, T. Baumgratz, and A. Datta, Advances in Physics: X 1, 621 (2016).
[24] N. Spagnolo, L. Aparo, C. Vitelli, A. Crespi, R. Ramponi, R. Osellame, P. Mataloni, and F. Sciarrino, Scientific Reports 2, 862 (2012).
[25] C. Macchiavello, Physical Review A 67, 062302 (2003).
[26] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, Physical Review Letters 111,
070403 (2013).

[27] M. D. Vidrighin, G. Donati, M. G. Genoni, X.-M. Jin, W. S. Kolthammer, M. Kim, A. Datta, M. Barbieri, and I. A. Walmsley, Nature Communications 5, 3532 (2014).

[28] Y. Yao, L. Ge, X. Xiao, X. Wang, and C. P. Sun, Physical Review A 90, 062113 (2014).

[29] D. W. Berry, M. Tsang, M. J. W. Hall, and H. M. Wiseman, Physical Review X 5, 031018 (2015).

[30] S. G. Schirmer and D. K. L. Oi, Physical Review A 80, 022333 (2009).

[31] M. P. Da Silva, O. Landon-Cardinal, and D. Poulin, Physical Review Letters 107, 210404 (2011).

[32] H. Yuan, Physical Review Letters 117, 160801 (2016).

[33] Y.-R. Zhang and H. Fan, Physical Review A 90, 043818 (2014).

[34] S. Pang and T. A. Brun, Physical Review A 90, 022117 (2014).

[35] T. Baumgratz and A. Datta, Physical Review Letters 116, 030801 (2016).

[36] E. Farhi and S. Gutmann, Physical Review A 57, 2403 (1998).

[37] N. Liu, J. Thompson, C. Weedbrook, S. Lloyd, V. Vedral, M. Gu, and K. Modi, Physical Review A 93, 052304 (2016).

[38] H. Imai and A. Fujiwara, Journal of Physics A 40, 4391 (2007).

[39] H. Yuan and C.-H. F. Fung, Physical Review Letters 115, 110401 (2015).

[40] W. F. Stinespring, Proceedings of the American Mathematical Society 2, 211 (1955).

[41] H. Chernoff, The Annals of Mathematical Statistics 23, 493 (1952).

[42] A. Hayashi, T. Hashimoto, and M. Horibe, Physical Review A 72, 032325 (2005).

[43] M. Suzuki, Communications in Mathematical Physics 51, 183 (1976).

[44] H. L. Van Trees, Detection, estimation, and modulation theory (John Wiley & Sons, 2004).

[45] D. A. Wisniacki, Physical Review E 67, 016205 (2003).