Precision-Guaranteed Quantum Metrology

Takanori Sugiyama

1 Institute for Theoretical Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, CH-8093 Zurich, Switzerland

(Dated: July 31, 2014)

Quantum metrology is a general term for methods to precisely estimate the value of an unknown parameter by actively using quantum resources like entanglement and squeezing [1, 2]. For example, when we use a separable state on an N-partite system in a Ramsey interferometer, an estimation error of phase, $\delta \phi$, scales as $O(1/\sqrt{N})$ (the standard quantum limit, SQL). On the other hand, when we use an entangled state like a Greenberger-Horne-Zeilinger (GHZ) state with the same number of particles, $\delta \phi$ scales as $O(1/N)$ (the Heisenberg limit, HL). Such quantum enhancement of precision has been experimentally achieved in several quantum systems like quantum optics [4], ions [5], and atoms [6].

One of the main goal of quantum metrology theory is to derive a fundamental lower bound on the estimation error. So far many different benchmarks for the estimation error have been proposed and analyzed [3]. The most popular benchmark is the root mean squared error (RMSE), and there are two standard approaches for analyzing RMSE. One is to apply a linear approximation of an estimation method to RMSE. The approximated RMSE is called a linearized uncertainty (LU). The other is to analyze the classical and quantum Cramér-Rao bounds (CRBs), which are lower bounds on RMSE, and they are not attainable by any estimation method. CRBs are lower bounds of RMSE, and they are not attainable when the amount of data is finite [5]. Unattainable lower bounds on an estimation error cannot be used to guarantee an estimation precision. Because the final goal of quantum metrology experiments is a highly precise estimation of an unknown parameter, it is best to rigorously guarantee an estimation precision, if possible. In order to do that, we need an upper bound on an estimation error satisfying two conditions: (1) be independent of the unknown parameter $\phi$, and (2) be valid for finite data.

In this paper, we derive an upper bound satisfying these two conditions for a general setting in quantum metrology. Our result makes it possible to rigorously guarantee the estimation precision in experiments, which is not possible by the standard approach of quantum metrology theory. We also prove that the upper bound shows the scaling the same as the LU, which means that the upper bound shows the HL scaling whenever the LU shows it. As an example, we apply our method to a Ramsey interferometer, and perform Monte Carlo simulations for $N = 1 \sim 100$. The numerical results indicate that the upper bound can exhibit the quantum enhancement of precision for finite data.

PACS numbers: 06.20.-f, 06.20.Dk, 03.67.-a, 02.50.Tt

High-precision measurement is one of the most important techniques for developing science and technology. Quantum metrology is a general term for methods to precisely estimate the value of an unknown parameter by actively using quantum resources like entanglement and squeezing [1, 2]. For example, when we use a separable state on an N-partite system in a Ramsey interferometer, an estimation error of phase, $\delta \phi$, scales as $O(1/\sqrt{N})$ (the standard quantum limit, SQL). On the other hand, when we use an entangled state like a Greenberger-Horne-Zeilinger (GHZ) state with the same number of particles, $\delta \phi$ scales as $O(1/N)$ (the Heisenberg limit, HL). Such quantum enhancement of precision has been experimentally achieved in several quantum systems like quantum optics [4], ions [5], and atoms [6].

One of the main goal of quantum metrology theory is to derive a fundamental lower bound on the estimation error. So far many different benchmarks for the estimation error have been proposed and analyzed [3]. The most popular benchmark is the root mean squared error (RMSE), and there are two standard approaches for analyzing RMSE. One is to apply a linear approximation of an estimation method to RMSE. The approximated RMSE is called a linearized uncertainty (LU). The other is to analyze the classical and quantum Cramér-Rao bounds (CRBs), which are lower bounds on RMSE, and they are not attainable by any estimation method. CRBs are lower bounds of RMSE, and they are not attainable when the amount of data is finite [5]. Unattainable lower bounds on an estimation error cannot be used to guarantee an estimation precision. Because the final goal of quantum metrology experiments is a highly precise estimation of an unknown parameter, it is best to rigorously guarantee an estimation precision, if possible. In order to do that, we need an upper bound on an estimation error satisfying two conditions: (1) be independent of the unknown parameter $\phi$, and (2) be valid for finite data.

In this paper, we derive an upper bound satisfying these two conditions for a general setting in quantum metrology. Our result makes it possible to rigorously guarantee the estimation precision in experiments, which is not possible by the standard approach of quantum metrology theory. We also prove that the upper bound shows the scaling the same as the LU, which means that the upper bound shows the HL scaling whenever the LU shows it. As an example, we apply our method to a Ramsey interferometer with N atoms, and perform Monte Carlo simulations for $N = 1 \sim 100$. The numerical results indicate that the upper bound can exhibit the quantum enhancement of precision for finite data.

PACS numbers: 06.20.-f, 06.20.Dk, 03.67.-a, 02.50.Tt

High-precision measurement is one of the most important techniques for developing science and technology. Quantum metrology is a general term for methods to precisely estimate the value of an unknown parameter by actively using quantum resources like entanglement and squeezing [1, 2]. For example, when we use a separable state on an N-partite system in a Ramsey interferometer, an estimation error of phase, $\delta \phi$, scales as $O(1/\sqrt{N})$ (the standard quantum limit, SQL). On the other hand, when we use an entangled state like a Greenberger-Horne-Zeilinger (GHZ) state with the same number of particles, $\delta \phi$ scales as $O(1/N)$ (the Heisenberg limit, HL). Such quantum enhancement of precision has been experimentally achieved in several quantum systems like quantum optics [4], ions [5], and atoms [6].

One of the main goal of quantum metrology theory is to derive a fundamental lower bound on the estimation error. So far many different benchmarks for the estimation error have been proposed and analyzed [3]. The most popular benchmark is the root mean squared error (RMSE), and there are two standard approaches for analyzing RMSE. One is to apply a linear approximation of an estimation method to RMSE. The approximated RMSE is called a linearized uncertainty (LU). The other is to analyze the classical and quantum Cramér-Rao bounds (CRBs), which are lower bounds on RMSE, and they are not attainable by any estimation method. CRBs are lower bounds of RMSE, and they are not attainable when the amount of data is finite [5]. Unattainable lower bounds on an estimation error cannot be used to guarantee an estimation precision. Because the final goal of quantum metrology experiments is a highly precise estimation of an unknown parameter, it is best to rigorously guarantee an estimation precision, if possible. In order to do that, we need an upper bound on an estimation error satisfying two conditions: (1) be independent of the unknown parameter $\phi$, and (2) be valid for finite data.

In this paper, we derive an upper bound satisfying these two conditions for a general setting in quantum metrology. Our result makes it possible to rigorously guarantee the estimation precision in experiments, which is not possible by the standard approach of quantum metrology theory. We also prove that the upper bound shows the scaling the same as the LU, which means that the upper bound shows the HL scaling whenever the LU shows it. As an example, we apply our method to a Ramsey interferometer with N atoms, and perform Monte Carlo simulations for $N = 1 \sim 100$. The numerical results indicate that the upper bound can exhibit the quantum enhancement of precision for finite data.

PACS numbers: 06.20.-f, 06.20.Dk, 03.67.-a, 02.50.Tt

High-precision measurement is one of the most important techniques for developing science and technology. Quantum metrology is a general term for methods to precisely estimate the value of an unknown parameter by actively using quantum resources like entanglement and squeezing [1, 2]. For example, when we use a separable state on an N-partite system in a Ramsey interferometer, an estimation error of phase, $\delta \phi$, scales as $O(1/\sqrt{N})$ (the standard quantum limit, SQL). On the other hand, when we use an entangled state like a Greenberger-Horne-Zeilinger (GHZ) state with the same number of particles, $\delta \phi$ scales as $O(1/N)$ (the Heisenberg limit, HL). Such quantum enhancement of precision has been experimentally achieved in several quantum systems like quantum optics [4], ions [5], and atoms [6].

One of the main goal of quantum metrology theory is to derive a fundamental lower bound on the estimation error. So far many different benchmarks for the estimation error have been proposed and analyzed [3]. The most popular benchmark is the root mean squared error (RMSE), and there are two standard approaches for analyzing RMSE. One is to apply a linear approximation of an estimation method to RMSE. The approximated RMSE is called a linearized uncertainty (LU). The other is to analyze the classical and quantum Cramér-Rao bounds (CRBs), which are lower bounds on RMSE, and they are not attainable by any estimation method. CRBs are lower bounds of RMSE, and they are not attainable when the amount of data is finite [5]. Unattainable lower bounds on an estimation error cannot be used to guarantee an estimation precision. Because the final goal of quantum metrology experiments is a highly precise estimation of an unknown parameter, it is best to rigorously guarantee an estimation precision, if possible. In order to do that, we need an upper bound on an estimation error satisfying two conditions: (1) be independent of the unknown parameter $\phi$, and (2) be valid for finite data.

In this paper, we derive an upper bound satisfying these two conditions for a general setting in quantum metrology. Our result makes it possible to rigorously guarantee the estimation precision in experiments, which is not possible by the standard approach of quantum metrology theory. We also prove that the upper bound shows the scaling the same as the LU, which means that the upper bound shows the HL scaling whenever the LU shows it. As an example, we apply our method to a Ramsey interferometer with N atoms, and perform Monte Carlo simulations for $N = 1 \sim 100$. The numerical results indicate that the upper bound can exhibit the quantum enhancement of precision for finite data.

PACS numbers: 06.20.-f, 06.20.Dk, 03.67.-a, 02.50.Tt
tum enhancement of precision for finite data.

**Preliminaries.**— We consider the following procedure of quantum metrology (Fig. 1): Step 1. Prepare a known quantum state $\rho$ on a probe system. Step 2. The state undergoes a dynamical process $\kappa_\phi$ with an unknown parameter $\phi$. Our aim is to estimate $\phi \in \Phi := [\phi_{\text{min}}, \phi_{\text{max}}]$, where $\phi_{\text{min}}$ and $\phi_{\text{max}}$ are assumed to be known. Step 3. After the dynamical process, the state changes to a known measurement on the state and obtain a measure- 

\[ \Pi \]

ment estimator, which is dependent of $\phi$. We perform a known measurement on the state and obtain a measure- 

\[ \text{out of } \Phi. \]

The asymptotic ($n \to \infty$) behavior of MLE is well known in classical statistics [10], but a rigorous analysis for finite $n$ is an open problem. Instead of MLE, we consider a different estimator relatively easier to be analyzed. 

\[ \phi_n^{\text{ML}}(\mathbf{x}^n) := \arg\min_{\phi'} |S_n - f(\phi')| \tag{1} \]

Same as MLE, $\phi_n^{\text{LS}}(\mathbf{x}^n)$ always exists and takes a value in $\Phi$. Let us define 

\[ S_n^{\text{LS}} := \arg\min_{r' \in \mathcal{R}_f} |r' - S_n|, \tag{2} \]

\[ V_{\text{max}} := (b - a)^2/4, \tag{3} \]

\[ V_n := \frac{1}{n - 1} \sum_{i=1}^{n} \left( x_i - \frac{1}{n} \sum_{j=1}^{n} x_j \right)^2, \tag{4} \]

\[ L := \max_{\phi \in \Phi} \left\{ \frac{df}{d\phi}(\phi') \right\}^{-3} \frac{d^2f}{d\phi^2}(\phi') \tag{5} \]

Note that $V_n$ is the sample variance of the data $\mathbf{x}^n$ and satisfies $\mathbb{E}[V_n] = \mathbb{V}[\Pi_\rho]$. Using the quantities introduced above, we define three functions of data $\mathbf{x}^n$ and an user-specified constant $\epsilon$.

\[ \delta_1(\mathbf{x}^n, \epsilon) := \frac{1}{\frac{df}{d\phi}((\phi_n^{\text{LS}}))} \sqrt{\frac{2}{n} V_{\text{max}} \ln \frac{2}{\epsilon} + \frac{L}{n} V_{\text{max}} \ln \frac{2}{\epsilon}}, \tag{6} \]

\[ \delta_2(\mathbf{x}^n, \epsilon) := \frac{1}{\frac{df}{d\phi}((\phi_n^{\text{LS}}))} \left\{ \frac{2}{n} V_{\text{max}} \ln \frac{4}{\epsilon} + \frac{8(b - a)}{3(n - 1)} \ln \frac{4}{\epsilon} \right\} \]

\[ + \frac{L}{2} \left\{ \frac{2}{n} V_{\text{max}} \ln \frac{4}{\epsilon} + \frac{8(b - a)}{3(n - 1)} \ln \frac{4}{\epsilon} \right\}^2 \tag{7} \]

\[ \delta(\mathbf{x}^n, \epsilon) := \min \{ \delta_1(\mathbf{x}^n, \epsilon), \delta_2(\mathbf{x}^n, \epsilon) \} \tag{8} \]

The following theorem guarantees that the deviation of the LS estimates from the true parameter is upper bounded by $\delta$ with high probability.

**Theorem 1** For any number of measurement trials $n \geq 2$, user-specified constant $\epsilon \in (0, 1)$, and true parameter $\phi \in \Phi$, 

\[ |\phi_n^{\text{LS}}(\mathbf{x}^n) - \phi| \leq \delta(\mathbf{x}^n, \epsilon) \tag{9} \]

holds with probability at least $1 - \epsilon$.

We sketch the proof, with the details shown in the Supplemental Material. The LS estimator is a nonlinear function of the sample mean, which is the origin of the main difficulty for the analysis. We use the Taylor expansion up to the second order with the reminder, and reduce the problem to an analysis on the deviation of the sample mean from the true expectation, $|S_n - f(\phi)|$. In the reduction, we use the contracitivity of the LS estimator,

\[ |S_n^{\text{LS}} - f(\phi)| \leq |S_n - f(\phi)|, \forall \mathbf{x}^n, \phi \in \Phi. \tag{10} \]

The contracitivity is one of the two main keys in this proof, and this is the reason why we choose the LS estimator. After the reduction, we use two inequalities for
evaluating $|S_n - f(\phi)|$. One is Hoeffding’s inequality [11], which is well known in classical statistics. The other is the empirical Bernstein inequality [12], which is a new mathematical tool developed for finite data analysis in machine learning. The empirical Bernstein inequality is the second key in this proof. It enables us to show a relation to the linearized uncertainty explained later. By combining these inequalities, contractivity, and Taylor expansion, we obtain Theorem 1.

It is important that $\delta(x^n, \epsilon)$ is dependent on only data $x^n$ and user-specified constant $\epsilon$, and that it is independent of the true parameter $\phi$. (The probability distribution of $\delta(x^n, \epsilon)$ depends on $\phi$.) Let us introduce a data-dependent interval,

$$L_I(x^n) := \Phi[\phi^{LS}_n(x^n) - \delta(x^n, \epsilon), \phi^{LS}_n(x^n) + \delta(x^n, \epsilon)]$$

Theorem 1 guarantees that, with probability at least $1 - \epsilon$, we obtain $L_I(x^n)$ including $\phi$. Such an interval is called a confidence interval with $(1 - \epsilon)$-confidence level. For example, when we choose $\epsilon = 0.01$, we obtain a confidence interval $I_{\epsilon = 0.01}(x^n)$ that includes $\phi$ with probability at least 99%. The upper bound $\delta$ becomes larger as we choose smaller $\epsilon$. This means that, if we require a higher confidence level for a size-fixed upper bound, we need a larger number of measurement trials.

Confidence interval and confidence level are well known concepts in classical statistics, and there are many statistical techniques to calculate them for finite data [13]. Most of these techniques are, however, based on the normal distribution approximation (NDA), and a confidence interval calculated with NDA is called an approximate confidence interval. NDA is valid when the number of measurement trials, $n$, is sufficiently large (the central limit theorem), but it is not clear which $n$ can be considered as sufficiently large. Therefore, it is not rigorous to apply approximate confidence intervals for finite data in experiments. In contrast to an approximate confidence interval, a confidence interval calculated without any assumption on probability distribution is called an exact confidence interval. To the best of our knowledge, our result is the first exact confidence interval for quantum metrology.

**Analysis.**— We explain a relation between $\delta$ and the linearized uncertainty (LU), $(\delta\phi)_{LU} := B_{LU}/\sqrt{n}$, where

$$B_{LU} := \sqrt{\frac{\mathbb{E}[\|\phi\|_2^2]}{\mathbb{E}[\|\phi\|_2^4]}} \left| \frac{d}{d\phi}[\phi^{LS}_n] \right|$$

By definition, $\delta_2$ decreases as $O(1/\sqrt{n})$, and the coefficient of the dominant term is given by $\sqrt{2V_n \ln \frac{1}{\epsilon} / \epsilon |d\phi^{LS}_n|}$. This coefficient converges to $B_{LU} \sqrt{2\ln \frac{1}{\epsilon}}$ in the limit of $n$ going to infinity because $V_n$ and $\phi^{LS}_n$ converge to $\mathbb{V}[\|\phi\|_2]$ and $\phi$, respectively. So, we would expect that $\delta$ have the scaling same as the LU with respect to $n$ and $\epsilon$. Actually we can prove the following inequality:

$$\lim_{n \to \infty} \left\{ \sqrt{n} \cdot \mathbb{E}[|\delta(x^n, \epsilon)|] \right\} \leq B_{LU} \sqrt{2\ln \frac{1}{\epsilon}},$$

where $\mathbb{E}$ denotes the expectation with respect to data $x^n$. The proof is shown in the Supplemental Material. The logic mentioned above and Eq. (10) guarantee that, on average, $\delta(x^n, \epsilon)$ scales the same as the LU. The upper bound $\delta$ shows the HL scaling, whenever the LU shows it. This is important especially in noisy cases. The quantum enhancement of precision can be suppressed when the dynamical process $\kappa_\phi$ is noisy [14], and recently there are many proposals for recovering the quantum enhancement with respect to LU [14] and CRB [16–18]. Eq. (11) indicates that the recovery method with respect to LU also works well for $\delta$.

**Example.**— We apply our result to a Ramsey interferometer with $N$ atoms. For a separable probe state of $N$ atoms, the LU scales as the SQL scaling, $O(1/\sqrt{N})$. On the other hand, for an entangled state like a GHZ state, the LU can scale as the HL scaling, $O(1/N)$. We consider two combinations of initial state and measurement. One is a combination of a separable state $\frac{1}{\sqrt{2}}(|e\rangle + |g\rangle)\otimes\rho_n$, and the measurement of the total energy, and the other is that of a GHZ state $\frac{1}{\sqrt{2}}(ee\cdots e + gg\cdots g)$ and the measurement of the parity, where $|e\rangle$ and $|g\rangle$ are excited and ground states of an atom, respectively. In the cases, we have $B_{LU} = 1/\sqrt{N}$ for the separable state and $B_{LU} = 1/N$ for the GHZ state.

We performed Monte Carlo simulations for the cases with $N = 1$, $n = 1$, $\phi_{min} = 0$, $\phi_{max} = \pi/400$, $\phi = \pi/400$, and $\epsilon = 0.1$ (90%-confidence level). The details are given in the Supplemental Material. In order to analyze typical behaviors of $\delta$, we calculated expectations of $\delta$ and compared them to expectations of $|\phi^{LS}_n - \phi|$. In both cases, (a) and (b), the vertical axes are for expected deviations. Solid and dashed (black) lines are $\mathbb{E}[|\delta(x^n, \epsilon)|]$ and $\mathbb{E}[|\phi^{LS}_n(x^n) - \phi|]$ for the separable state, respectively. Chained and dotted (red) lines are $\mathbb{E}[|\delta(x^n, \epsilon)|]$ and $\mathbb{E}[|\phi^{LS}_n(x^n) - \phi|]$ for the GHZ state, respectively. The expectations were calculated by a Monte Carlo sampling with 5000 repetitions. In panel (a), the horizontal axis is the number of atoms, $N$. Plots in the panel express the scaling of the expected deviations with respect to $N$ with a fixed number of measurement trials, $n = 3000$. The expectations of $\delta$ are larger than those of $|\phi^{LS}_n - \phi|$, which is consistent with Theorem 1. Panel (a) also indicates that, up to $N = 100$, the expectation of $\delta$ for the GHZ state scales as the HL scaling, although that for the separable state scales as the SQL scaling. In panel (b), the horizontal axis is the number of measurement trials, $n$. Plots in the panel express the scaling of the expected deviations with respect to $n$ with a fixed number of atoms, $N = 100$. The expectations of $\delta$ for both states scale as $O(1/\sqrt{N})$, and $\delta$ for the GHZ state is, on average, 10 ($= \sqrt{N}$ in the panel) times smaller than $\delta$ for the separable state.
FIG. 2. Numerical result on expected deviations ($\mathbb{E}\left[\delta\right]$ with $\epsilon = 0.1$, and $\mathbb{E}\left[\left|\phi_n^{LS} - \phi\right|\right]$) in a Ramsey interferometer using a separable or GHZ initial state of $N$ atoms. Expectations were calculated by a Monte Carlo sampling with 5000 repetitions. Panel (a) is for their $N$-dependency with $n = 3000$, and panel (b) is for their $n$-dependency with $N = 100$. Both panels indicate that $\delta$ for the GHZ state shows the HL scaling up to $N = 100$ for finite $n$.

In conclusion of the numerical simulations, the expectations of $\delta$ with 90%-confidence level are larger than the expectations of the actual deviations for both separable and entangled states, which is consistent with Theorem 1. Furthermore, Fig. 2 indicates that, compared to the separable state, the entangled state gives smaller deviation of estimates and smaller error bar $\delta$. Eq. (13) guarantees that $\delta$ shows the HL scaling for asymptotically large $n$ whenever the LU shows the scaling, and Fig. 2 indicates that $\delta$ can also show the quantum enhancement of precision for finite $n$. Note that the Ramsey interferometer is mathematically equivalent to a Mach-Zehnder interferometer [19], which means that $\delta$ can show the quantum enhancement of precision in an optical interferometer with a $N00N$ states.

Discussion.— In Theorem 1, it is assumed that we perfectly know $\rho$, $\Pi$, and the functional form of $\kappa_{\phi}$. This assumption may not be valid when there exists a systematic error in experiments. In the standard approach of quantum metrology theory, a model for the systematic error is introduced, and it is assumed that the model correctly characterizes the error and that we know the value of a noise parameter in the model [20, 21]. Theorem 1 is applicable for such a perfectly known systematic error. Even if the model is correct, however, the value that we think as the noise parameter may be different from the true value in an experiment. Theorem 1 and the standard approach are not directly applicable for such a partially unknown systematic error, but we can obtain an exact confidence interval for quantum metrology with a partially unknown systematic error, by modifying Theorem 1 based on the worst case of the noise parameter. The explicit form of the exact confidence interval is given in the Supplemental Material.

Summary.— We considered a general setting of quantum metrology, proposing a least squares estimator and deriving an explicit formula of an exact confidence interval for the estimator with arbitrary finite number of measurement trials larger than 2. We showed that the size of the interval, $\delta$, scales the same as the linearized uncertainty, which is a popular benchmark in the standard approach of quantum metrology, for asymptotically large number $n$ of measurement trials. This means that $\delta$ asymptotically shows the Heisenberg limit scaling whenever the linearized uncertainty shows the scaling. As an example, we applied our results to a Ramsey interferometer with $N$ atoms and performed Monte Carlo simulations for $N = 1 \sim 100$ and $n = 1 \sim 10000$. The numerical result indicates that, when a GHZ state is used as an initial state, $\delta$ shows the Heisenberg limit scaling for finite $n$. It means that $\delta$ can also exhibit the quantum entrapment of precision for finite $n$. To the best of our knowledge, this is the first result that makes it possible to rigorously guarantee an estimation precision in quantum metrology with finite data, and we hope it finds application in the analysis of experimental data.

The author acknowledges financial support by JSPS Postdoctoral Fellowships for Research Abroad (H25-32), the German Science Foundation (grant CH 843/2-1), the Swiss National Science Foundation (grants PP00P2-128455, 20CH21-138799 (CHIST-ERA project CQC)), the Swiss National Center of Competence in Research ‘Quantum Science and Technology (QSIT)’ and the Swiss State Secretariat for Education and Research supporting COST action MP1006.
Appendix A: Notations and Assumptions

In this section, for convenience we give a summary of notation and list of assumptions.

1. Summary of Notation

1. On setting of quantum metrology

- $\mathcal{H}$: a Hilbert space characterizing the probe system. In the example of the Ramsey interferometer in Sec. D, we consider a case of $\mathcal{H} = (\mathbb{C}^2)^\otimes N$, but Theorem 1 holds for arbitrary dimensional Hilbert space, which can be infinite dimensional.
- $\rho$: the density matrix/operator on $\mathcal{H}$ characterizing an initial state of the probe system.
- $\phi$: the unknown parameter to be estimated, which takes a value in $\Phi := [\phi_{\min}, \phi_{\max}]$.
- $\kappa_\phi$: a trace-preserving and completely positive map characterizing the dynamics in step 2, which is parameterized by $\phi$.
- $\rho_0 := \kappa_\phi(\rho)$: the density matrix/operator characterizing the state after the dynamical process.
- $\Pi = \{\Pi_x\}_{x \in X}$: a POVM characterizing the measurement, which is not necessarily a projective measurement. For simplicity, we consider a discrete and finite $X$, but Theorem 1 holds for a continuous $X$ as well.
- $a := \min_{x \in X} x$: the minimal value in the possible measurement outcomes.
- $b := \max_{x \in X} x$: the maximal value in the possible measurement outcomes.

2. On measurement statistics

- $p(x|\rho_0, \Pi) := \text{Tr}[\rho_0 \Pi_x]$: the probability that we obtain an outcome $x$.
- $E[\Pi|\rho_0] := \sum_{x \in X} x \cdot p(x|\rho_0, \Pi)$: the expectation of a measurement outcome of $\Pi$.
- $f$: the functional form of $E[\Pi|\rho_0]$, i.e., $f(\phi) := E[\Pi|\rho_0]$.
- $\mathcal{R}_f$: the range of $f$, i.e., $\mathcal{R}_f := \{f(\phi)|\phi \in \Phi\}$.
- $g := f^{-1}$: the inverse of $f$ on $\Phi$. $g(E[\Pi|\rho_0]) = \phi$ holds.
- $\nu[\Pi|\rho_0] := \sum_{x \in X} (x - E[\Pi|\rho_0])^2 p(x|\rho_0, \Pi)$: the variance of a measurement outcome of $\Pi$.
- $n$: the number of measurement trials.
- $x_i$: an outcome of the $i$-th measurement trial ($1 \leq i \leq n$).
- $x^n := \{x_1, \ldots, x_n\}$: data obtained, which is a set of $n$ outcomes.

- $S_n := \frac{1}{n} \sum_{i=1}^n x_i$: the sample mean of the data.
- $S_n^{LS} := \arg \min_{x' \in \mathcal{R}_f} |x' - S_n|$: the projection of $S_n$ on $\mathcal{R}_f$. The contractivity, $|S_n^{LS} - f(\phi)| \leq |S_n - f(\phi)|$, holds for any $\phi \in \Phi$ and data.
- $V_n := \frac{1}{n} \sum_{i=1}^n \left( x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2$: the sample variance of the data. $E[V_n] = V[\Pi|\rho_0]$ holds for any $n \geq 2$.
- $V_{\max} := \frac{1}{n} \left( b - a \right)^2$: the maximal variance. $V_{\max} \geq V[\Pi|\rho_0]$ holds.
- $\text{BLU} := \sqrt{\text{Var}[\Pi|\rho_0]}$: the linearized uncertainty (LU) normalized by $\sqrt{n}$. The statistical meaning is explained in Sec. A.3.

3. On estimator

- $\phi_{\text{DI}}$: a direct inversion (DI) estimator defined by $\phi_{\text{DI}}^n(x^n) := g(S_n)$.
- $\phi_{\text{DI}}^{LS}(x^n)$: a direct inversion estimate of $x^n$. It does not necessarily exist for any data. Even if it exists, it can be out of $\Phi$.
- $\phi_{\text{LS}}$: a least squares (LS) estimator defined by $\phi_{\text{LS}}^n(x^n) := \arg \min_{\phi \in \Phi} |S_n - f(\phi)|$.
- $\phi_{\text{LS}}^{LS}(x^n)$: a least square estimate of data $x^n$. It exists for any data, and it is always in $\Phi$. $\phi_{\text{LS}}^{LS}(x^n) = g(S_n^{LS})$ and $S_n^{LS} = f(\phi_{\text{LS}}^{LS}(x^n))$ hold.

4. On error bars

- $\epsilon$: a user-specified parameter taking a value in $(0,1)$, which determines the confidence level of $\delta$.
- $L := \max_{\phi \in \Phi} \left\{ \left. \left| \frac{df}{d\phi}(\phi') \right| \right|^{3/2} \right. \frac{\hat{\sigma}^2(\phi')}{\nu[\Pi|\rho_0]} \right\}^{-1/2}$: a constant characterizing a nonlinearity of $f(\phi) = E[\Pi|\rho_0]$.
- $\delta_1(x^n, \epsilon) := \frac{1}{\sqrt{n}} \left\{ \left( \nu[\Pi|\rho_0] \right)^{1/2} \right\}$: an error bar derived with Hoeffding’s inequality (Lemma 1).
- $\delta_2(x^n, \epsilon) := \frac{1}{\left( \nu[\Pi|\rho_0] \right)^{1/2}} \left\{ \left( \nu[\Pi|\rho_0] \right)^{1/2} \right\}$: an error bar derived with the empirical Bernstein inequality (Lemma 2).
- $\delta(x^n, \epsilon) := \min\{\delta_1(x^n, \epsilon), \delta_2(x^n, \epsilon)\}$: an error bar with $(1-\epsilon)$-confidence level.

5. $I_\epsilon$: an exact confidence interval with $(1-\epsilon)$-confidence level, defined by $I_\epsilon(x^n) := \Phi \cap \{\phi_{\text{LS}}^{LS}(x^n) - \delta(x^n, \epsilon), \phi_{\text{LS}}^{LS}(x^n) + \delta(x^n, \epsilon)\}$.

We give a sketch of a procedure for calculating a confidence interval $I_\epsilon$ in Figure 51.
2. List of Assumptions

Theorem 1 holds under the following four assumptions.

A1. We know $\rho$, $\Pi$, and the functional form of $\kappa_\phi$, i.e.,
we know $f$.

A2. The measurement outcomes are bounded, i.e.,
$-\infty < a \leq x \leq b < +\infty$, $\forall x \in \mathcal{X}$.

A3. $f$ is injective for $\phi \in \Phi$.

A4. The derivative of $f$ is always non-zero on $\Phi$, i.e.,
$\frac{df}{d\phi} \neq 0$, $\forall \phi \in \Phi$.

Assumption A1 is the standard assumption not only in
quantum metrology, but also in statistical parameter
estimation. Assumption A2 is necessary for the use of Hoeffding’s inequality (Lemma 1) and empirical Bernstein
inequality (Lemma 2) in the proof of Theorem 1. Unbounded
outcomes can exist theoretically, but outcomes are always bounded in experiments. So, assumption A2
is natural in experiments. Assumption A3 is necessary for
the uniqueness of the least squares estimates for any
data, and assumption A4 is necessary for avoiding the
divergence of $\delta_1$ and $\delta_2$. In Sec. [A3], we explain that
assumptions A3 and A4 are required in the use of the
linearized uncertainty, which means that A3 and A4 are implicitly assumed in the standard approach using LU.

3. Linearized uncertainty and Assumptions

We explain a relation between the root mean squared
error (RMSE) and LU (Eq. [A8]), which clarifies the
role of assumptions A3 and A4 for the LU. The root
mean squared error of an estimator $\phi^{\text{est}}$ is defined by

$$ (\delta\phi)_{\text{RMSE}}(\phi^{\text{est}}, n|\phi) := \sqrt{\mathbb{E}[(\phi^{\text{est}}(x^n) - \phi)^2]} \quad (A1) $$

Let us choose the DI estimator, $\phi^{\text{DI}}$ as the estimator. The
DI estimates do not necessarily exist for any data. Assumption A3 guarantees the existence of the DI estimates
only for $S_n \in \mathcal{R}_f$. When $S_n$ is out of $\mathcal{R}_f$, the DI
estimate may not exist. A3 is a necessary condition for the
existence, but it is not a sufficient condition. However, let us ignore this fact, i.e., we assume that DI estimates
exist for any data. By definition,

$$ f(\phi^{\text{DI}})_n = S_n \quad (A2) $$
holds. We have

$$ \mathbb{V}(\Pi|\rho_\phi) = n \cdot \mathbb{E}[(S_n - \mathbb{E}[\Pi|\rho_\phi])^2] \quad (A3) $$

$$ = n \cdot \mathbb{E}[(f(\phi^{\text{DI}})_n - f(\phi))^2]. \quad (A4) $$

We apply the Taylor expansion to $f$,

$$ f(\phi^{\text{DI}})_n = f(\phi) + \frac{df}{d\phi}(\phi) \cdot (\phi^{\text{DI}}_n - \phi) + O((\phi^{\text{DI}}_n - \phi)^2) \quad (A5) $$

and suppose that $n$ is sufficiently large that the nonlinear
terms in the Taylor expansion, $O((\phi^{\text{DI}}_n - \phi)^2)$, is negligible. Then

$$ \mathbb{V}(\Pi|\rho_\phi) = n \cdot \left( \frac{df}{d\phi}(\phi) \right)^2 \mathbb{E}[(\phi^{\text{DI}}_n - \phi)^2] \quad (A6) $$

$$ \approx n \cdot \left( \frac{df}{d\phi}(\phi) \right)^2 (\delta\phi)_{\text{RMSE}}(\phi^{\text{DI}}_n, n|\phi)^2 \quad (A7) $$

holds. Since $\frac{df}{d\phi} \neq 0$ holds from assumption A4, and we obtain

$$ (\delta\phi)_{\text{RMSE}}(\phi^{\text{DI}}_n, n|\phi) \approx \frac{1}{\sqrt{n}} \frac{\mathbb{V}(\Pi|\rho_\phi)}{|\frac{df}{d\phi}(\phi)|} \quad (A8) $$

$$ = (\delta\phi)_{\text{LU}} \left( := \frac{B_{\text{LU}}}{\sqrt{n}} \right). \quad (A9) $$

Eq. [A8] means that the linearized uncertainty is an
approximated RMSE of the DI estimator, which is derived
by ignoring the existence problem of the estimator and
the nonlinearity of $f$. This is the reason why we call
$(\delta\phi)_{\text{LU}}$ a linearized uncertainty.

In the derivation of Eq. [A8], the following two
conditions are required in addition to assumptions A3 and
A4.

C1. DI estimates exist for $S_n \notin \mathcal{R}_f$.

C2. The number of measurement trials $n$ is sufficiently
large that the nonlinearity of $f$ around $\phi$ is negligible.

In the standard approach using LU, A1, A3, A4, C1, and
C2 are implicitly assumed. On the other hand, Theorem
1 does not require C1 and C2. Especially the disuse of C2
is important to analyze finite data, because it is unclear
which $n$ can be considered as “sufficiently” large in C2.
Appendix B: Proof of Theorem 1

In this section we show the proof of Theorem 1. We derive an upper bound of $\Pr[|\phi_{n}^{LS} - \phi| > \delta] = 1 - \Pr[|\phi_{n}^{LS} - \phi| \leq \delta]$. It is difficult to directly analyze this quantity, because $\phi_{n}^{LS}$ is a nonlinear function of $S_n$ and is a biased estimator. On the other hand, the following two lemmas hold for $S_n$.

Lemma 1 (Hoeffding’s inequality [11]) Let $X$ be a random variable with $X \in [a, b]$ and $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random variables satisfying $X_i = X(i = 1, \ldots, n)$, respectively. Then for any $0 < \epsilon < 1$ and $n \geq 1$,

$$\Pr \left[ \left| S_n - \mathbb{E}[X] \right| > \sqrt{\frac{2}{n} V_{\max} \ln \frac{2}{\epsilon}} \right] \leq \epsilon$$

(B1)

holds.

Lemma 2 (Empirical Bernstein inequality [12]) Let $X$ be a random variable with $a \leq X \leq b$ and $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random variables satisfying $X_i = X(i = 1, \ldots, n)$, respectively. Then for any $0 < \epsilon < 1$ and $n \geq 2$,

$$\Pr \left[ \left| S_n - \mathbb{E}[X] \right| > \sqrt{\frac{2}{n} V_n(X^n) \ln \frac{4}{\epsilon} + \frac{8(b-a)}{3(n-1)} \ln \frac{4}{\epsilon}} \right] \leq \epsilon$$

(B2)

holds[23], where

$$V_n(X^n) := \frac{1}{n-1} \sum_{i=1}^{n} \left( X_i - \frac{1}{n} \sum_{j=1}^{n} X_j \right)^2.$$  

(B3)

Note that $V_{\max}$ in Hoeffding’s inequality is independent of data, and that $V_n(X^n)$ in the empirical Bernstein inequality is dependent of data.

First, we reduce the analysis of $|\phi_{n}^{LS}(x^n) - \phi|$ to that of $|S_n - \mathbb{E}[\Pi|\rho_{\phi}|]|$. Let $r$ denote the argument of $g$. Using the Taylor expansion of $g(r)$ around $S_n^{LS}$ up to the 2nd order with the remainder in the Lagrange form, we obtain the following inequality.

$$\left| \phi_{n}^{LS}(x^n) - \phi \right| = \left| g(S_n^{LS}) - g(r) \right|$$

(B4)

$$= \left| \frac{dg}{dr} \left( S_n^{LS} \right) (r - S_n^{LS}) + \frac{1}{2} \frac{d^2g}{dr^2}(r') (r - S_n^{LS})^2 \right|$$

(B5)

$$\leq \left| \frac{dg}{dr} \left( S_n^{LS} \right) \right| \cdot |r - S_n^{LS}| + \frac{1}{2} \left| \frac{d^2g}{dr^2}(r') \right| \cdot |r - S_n^{LS}|^2$$

(B6)

where $r'$ is some real number between $r$ and $S_n^{LS}$. By combining Eq. (B6) with $\left| \frac{d^2g}{dr^2}(r') \right| = \left| \left( \frac{df}{dr} (\phi') \right)^{-1} \cdot \frac{d^2f}{dr^2} (\phi') \right| \leq L$ and the contractivity (Eq. (10)), we obtain

$$\left| \phi_{n}^{LS}(x^n) - \phi \right| \leq \left| \frac{dg}{dr} \left( S_n^{LS} \right) \right| \cdot |S_n - r| + \frac{1}{2} L |S_n - r|^2.$$  

(B7)

Then $\delta < |\phi_{n}^{LS} - \phi|$ implies

$$\delta < \left| \frac{dg}{dr} \left( S_n^{LS} \right) \right| \cdot |S_n - r| + \frac{1}{2} L |S_n - r|^2.$$  

(B8)

By solving this quadratic inequality with $\delta > 0$, we can show that Eq. (B8) is equivalent to

$$|S_n - r| > \frac{1}{L} \left( \sqrt{\left| \frac{dg}{dr} \left( S_n^{LS} \right) \right|^2 + 2\delta L - \left| \frac{dg}{dr} \left( S_n^{LS} \right) \right|^2} \right)$$

(B9)

By substituting $\delta = \delta_1$ and $\delta = \delta_2$ into Eq. (B9), we obtain

$$\frac{1}{L} \left( \sqrt{\left| \frac{dg}{dr} \left( S_n^{LS} \right) \right|^2 + 2\delta_1 L - \left| \frac{dg}{dr} \left( S_n^{LS} \right) \right|^2} \right) = \sqrt{\frac{2}{n} V_{\max} \ln \frac{2}{\epsilon}},$$

(B10)

$$\frac{1}{L} \left( \sqrt{\left| \frac{dg}{dr} \left( S_n^{LS} \right) \right|^2 + 2\delta_2 L - \left| \frac{dg}{dr} \left( S_n^{LS} \right) \right|^2} \right) = \sqrt{\frac{2}{n} V_n(X^n) \ln \frac{4}{\epsilon} + \frac{8(b-a)}{3(n-1)} \ln \frac{4}{\epsilon}},$$

(B11)

where we used the equalities $\frac{dg}{dr}(\phi_{n}^{LS}) = 0$ from assumption A3 and $\frac{df}{dr}(\phi_{n}^{LS}) \neq 0$ from assumption A4 (note that $\phi_{n}^{LS}(x^n) \in \Phi$ holds for any $x^n$). From Lemmas 1 and 2, we obtain

$$\Pr \left[ |\phi_{n}^{LS} - \phi| > \delta \right] \leq \Pr \left[ |S_n - \mathbb{E}[\Pi|\rho_{\phi}|]| > \min \left\{ \sqrt{\frac{2}{n} V_{\max} \ln \frac{2}{\epsilon}}, \sqrt{\frac{2}{n} V_n(X^n) \ln \frac{4}{\epsilon} + \frac{8(b-a)}{3(n-1)} \ln \frac{4}{\epsilon}} \right\} \right] \leq \epsilon.$$  

(B12)

(B13)

□

Appendix C: Proof of Eq. (13)

Here we show the proof of Eq. (13).

$$\lim_{n \to \infty} \left\{ \sqrt{n} \cdot \mathbb{E}[\delta(x^n, \epsilon)] \right\} \leq \lim_{n \to \infty} \left\{ \sqrt{n} \cdot \mathbb{E}[\delta_2(x^n, \epsilon)] \right\}$$

(C1)

$$= \lim_{n \to \infty} \left\{ \mathbb{E} \left[ \left| \frac{dg}{dr} \left( S_n^{LS} \right) \right| \sqrt{V_n} \right] \right\} \sqrt{2 \ln \frac{4}{\epsilon}}$$

(C2)

$$\leq \lim_{n \to \infty} \left\{ \mathbb{E} \left[ \left| \frac{dg}{dr} \left( S_n^{LS} \right) \right|^2 \right] \sqrt{\mathbb{E}[V_n]} \right\} \sqrt{2 \ln \frac{4}{\epsilon}}$$

(C3)

$$= \lim_{n \to \infty} \mathbb{E} \left[ \left| \frac{dg}{dr} \left( S_n^{LS} \right) \right|^2 \right] \sqrt{\mathbb{E}[V_n]} \sqrt{2 \ln \frac{4}{\epsilon}}.$$  

(C4)
where we used the Cauchy-Schwarz inequality and the equality $\mathbb{E}[V_n] = \mathbb{V}[\Pi | \rho_\phi]$. From the Taylor expansion, we have
\[
\frac{dg}{dr}(S_{n}^{\text{LS}}) = \frac{dg}{dr}(r) + O(|S_{n}^{\text{LS}} - r|).
\] (C5)
At the limit of $n$ to infinity, $|S_{n}^{\text{LS}} - r|$ converges to 0 because of the contractivity, $|S_{n}^{\text{LS}} - r| \leq |S_n - r|$, and the law of large numbers. Then
\[
\lim_{n \to \infty} \mathbb{E} \left[ \frac{dg}{dr}(S_{n}^{\text{LS}}) \right]^2 = \left| \frac{dg}{dr}(r) \right|^2
\] (C6)
holds, and we obtain
\[
\lim_{n \to \infty} \{ \sqrt{\mathbb{E} [\delta(x^n, \epsilon)]} \} \leq \frac{\sqrt{\mathbb{V} [\Pi | \rho_\phi]}}{\sqrt{2N}} \cdot 2 \ln \frac{4}{\epsilon} \quad (C7)
\]
and
\[
\frac{\sqrt{\mathbb{V} [\Pi | \rho_\phi]}}{d \phi} \cdot \sqrt{2 \ln \frac{4}{\epsilon}}. \quad (C8)
\]
\[\square\]

### Appendix D: Example: Ramsey interferometer

In this section, we explain the details of the Ramsey interferometer analyzed in the main text. When we use a separable state of $N$ atoms for the initial state, the LU scales as the SQL scaling, $O(1/\sqrt{N})$. On the other hand, when we use an entangled state, the LU can scale as the HL scaling, $O(1/N)$. The procedure of the Ramsey interferometer is as follows.

1. Prepare an initial state $|\phi\rangle$ of $N$ atoms. Each atom is a two-level system.
2. Each atom independently undergoes a free evolution, $\exp \left( i \frac{\phi}{2} \sigma_3 \right)$.
3. After the evolution, we perform a $\frac{\pi}{2}$-pulse along an axis, $\cos \phi_0 \cdot \sigma_1 + \sin \phi_0 \cdot \sigma_2$, where $\phi_0$ is a reference phase to be user-tuned.
4. Perform a projective measurement of an observable, $A$.
5. Repeat 1 to 4 a number $n$ of times.

We consider the following two combinations of the initial state $|\psi\rangle$ and measured observable $A$.

1. **Product state and energy measurement**
   Let us choose a product state,
   \[
   |\psi\rangle = \left[ \frac{1}{\sqrt{2}} (|e\rangle + |g\rangle) \right]^\otimes N \quad \text{(D1)}
   \]
as the initial state, where $|e\rangle$ and $|g\rangle$ are the excited and ground states, respectively. We observe the total energy,
   \[
   J_3 := \sum_{j=1}^{N} \sigma_3^{(j)}, \quad \text{(D2)}
   \]
   where $\sigma_3^{(j)} := I^{(j-1)} \otimes \sigma_3 \otimes I^{(N-j)}$. The set of possible measurement outcomes is $\mathcal{X} = \{-N, -N+1, \ldots, N-1, N\}$. In this combination, the probability distribution is given by
   \[
p(x | \rho_\phi, \Pi) = \frac{N!}{(N+x)!((N-x)/2)!} \cdot \left\{ \frac{1 + \sin(\phi - \phi_0)}{2} \right\}^{N-x} \left\{ \frac{1 - \sin(\phi - \phi_0)}{2} \right\}^{x-N/2} \quad \text{(D3)}
   \]
   and we obtain the following equalities:
   \[
   f(\phi) = N \sin(\phi - \phi_0), \quad \text{(D4)}
   \]
   \[
   \frac{df}{d\phi} = N \cos(\phi - \phi_0), \quad \text{(D5)}
   \]
   \[
   \mathbb{V}[\Pi | \rho_\phi] = N \cos^2(\phi - \phi_0), \quad \text{(D6)}
   \]
   \[
   V_{\text{max}} = N^2, \quad \text{(D7)}
   \]
   \[
   B_{\text{LU}} = \frac{1}{\sqrt{N}}, \quad \text{(D8)}
   \]
   \[
   \left\{ \frac{df}{d\phi} \right\}^{-3} \cdot \frac{d^2 f}{d\phi^2} = -\frac{1}{N^2} \sin(\phi - \phi_0). \quad \text{(D9)}
   \]

From Eq. (D4), $f$ is a periodic function with period $2\pi$. In order to satisfy assumptions A3 and A4, the size of $\Phi$ must be at most smaller than $\pi$.

2. **A GHZ state and parity measurement**
   Let us choose a GHZ state,
   \[
   |\psi\rangle = \frac{1}{\sqrt{2}} (|e\rangle |e\rangle \cdots |e\rangle + |g\rangle |g\rangle \cdots |g\rangle) \quad \text{(D10)}
   \]
as the initial state. We observe the parity,
   \[
   P = (+1)^{N_g} (-1)^{N_e}, \quad \text{(D11)}
   \]
   where $N_g$ and $N_e$ are the particle number operators for $|g\rangle$ and $|e\rangle$, respectively. The set of possible measurement outcomes is $\mathcal{X} = \{-1, +1\}$. In this combination, the probability distribution is given by
   \[
p(x | \rho_\phi, \Pi) = \frac{1}{2} \left\{ 1 + x \cdot \cos \left( \phi - \phi_0 + \frac{\pi}{2} \right) \right\} \quad \text{(D12)}
   \]
In this case, we have the following equalities:

\[
\begin{align*}
    f(\phi) &= \cos N \left( \phi - \phi_0 + \frac{\pi}{2} \right), \\
    \frac{df}{d\phi} &= -N \sin N \left( \phi - \phi_0 + \frac{\pi}{2} \right),
\end{align*}
\]

\[
\begin{align*}
    \mathbb{V}[\Pi|\phi_0] &= \sin^2 N \left( \phi - \phi_0 + \frac{\pi}{2} \right), \\
    V_{\text{max}} &= 1, \\
    B_{\text{LU}} &= \frac{1}{N}, \\
    \left\{ \frac{df}{d\phi} \right\}^{-3} \cdot \frac{d^2f}{d\phi^2} &= \frac{1}{N} \cos N \left( \phi - \phi_0 + \frac{\pi}{2} \right),
\end{align*}
\]

From Eq. (D13), \( f \) is a periodic function with period \( 2\pi/N \). In order to satisfy assumptions A3 and A4, the size of \( \Phi \) must be at most smaller than \( \pi/N \).

Note that \( J_3 \) and \( P \) are commuting, and that these are reduced from the measurement of particle numbers for each energy level in an experiment [24]. Mathematically these are different observables, but the same measurement apparatus is used for both of two in the experiment. So, in cases (1) and (2), their initial states are different, their observables are different, and their POVMs are same. In general, the LU is larger than or equivalent to the CRB, but in cases (1) and (2) their LUs coincide with their CRBs, respectively. This is the reason why we choose different observables.

We performed a Monte Carlo simulation for cases (1) and (2) with the following parameters: \( \phi_{\text{min}} = 0 \), \( \phi_{\text{max}} = \pi/400 \), \( \phi = \pi/4000 \), and \( \epsilon = 0.1 \) (90\%-confidence level). The reference phases are chosen as \( \phi_0 = -\pi/8 \) for case (1) and \( \phi_0 = \pi/2 - \pi/10N \). We show the result in Fig. S2 where \( B_{\text{LU}} \) and \( \phi_{\text{max}} - \phi_{\text{min}} \) are added to Fig. 2. In both panels (a) and (b) of Fig. S2 vertical axes are for expected deviations, \( \mathbb{E}[\delta(x^n, \epsilon)] \) and \( \mathbb{E}[\phi_n^{LS}(x^n) - \phi] \). Solid and dashed (black) lines are \( \mathbb{E}[\delta(x^n, \epsilon)] \) and \( \mathbb{E}[\phi_n^{LS}(x^n) - \phi] \) for the separable state, respectively. Chained and dotted (red) lines are \( \mathbb{E}[\delta(x^n, \epsilon)] \) and \( \mathbb{E}[\phi_n^{LS}(x^n) - \phi] \) for the GHZ state, respectively. The expectations were calculated by a Monte Carlo sampling with 5000 repetitions. Each horizontal axis in panels (a) and (b) is for the number of atoms \( N \) and the number of measurement trials \( n \), respectively. As explained in the main text, these panels indicate that \( \delta \) for the GHZ state shows the Heisenberg limit scaling, \( O(1/N) \) for finite \( n \). The scalings of \( \mathbb{E}[\delta(x^n, \epsilon)] \) with respect to \( N \) and \( n \) are independent of \( \epsilon \), and the quantum enhancement on \( \delta \) appears not only for \( \epsilon = 0.1 \) but also for other values of \( \epsilon \).

Solid blue and orange lines in Fig. S2 are \( B_{\text{LU}} \sqrt{2 \ln 2} \) for cases (1) and (2), respectively. Eq. (13) guarantees that \( \mathbb{E}[\delta(x^n, \epsilon)] \) becomes smaller than the new lines in the limit of \( n \) going to infinity. Panel (b) indicates that two lines for \( \mathbb{E}[\delta(x^n, \epsilon)] \) (solid black and chained red lines) become closer to the blue and orange lines as \( n \) becomes larger, respectively. However, lines for \( \mathbb{E}[\delta(x^n, \epsilon)] \) are still larger than \( B_{\text{LU}} \sqrt{2 \ln 2} \) at \( n = 10000 \). This means that \( n = 10000 \) cannot be considered as a “sufficiently” large number.

Solid green lines in Fig. S2 are for \( \phi_{\text{max}} - \phi_{\text{min}} = \pi/400 \approx 0.008 \). Because \( \phi_n^{LS}(x^n) \) is always included in \( \Phi \).
for any data,

$$|\phi_n^{LS} - \phi| \leq \phi_{max} - \phi_{min}$$

(D19)

holds with probability 1. So, $\phi_{max} - \phi_{min}$ gives a trivial 100%-confidence interval. In panel (a) of Fig. S3 $\mathbb{E}[\delta]$ for the separable state (solid black line) is larger than $\phi_{max} - \phi_{min}$ (the solid green line) for all $N$ between 1 to 100. This means that, on average, $n = 3000$ is not enough for obtaining a nontrivial 90%-confidence interval in case (1), while the number is enough in case (2) with $N \geq 10$. In panel (b), $\mathbb{E}[\delta]$ for the separable state (solid black line) becomes smaller than $\phi_{max} - \phi_{min}$ (solid green line) for $n \geq 6000$. If we perform measurement trials more than 6000 times for the separable state, it is expected to obtain a non-trivial 90%-confidence interval.

Finally we analyze the behaviors of $\delta_1$ and $\delta_2$. In both panels (a) and (b) of Fig. S3 vertical axes are for expected deviations, $\mathbb{E}[\delta_1(x^n, \epsilon)]$ and $\mathbb{E}[\delta_2(x^n, \epsilon)]$. Solid and dashed (black) lines are $\mathbb{E}[\delta_1(x^n, \epsilon)]$ and $\mathbb{E}[\delta_2(x^n, \epsilon)]$ for the separable state, respectively. Solid and dashed (red) lines are $\mathbb{E}[\delta_1(x^n, \epsilon)]$ and $\mathbb{E}[\delta_2(x^n, \epsilon)]$ for the GHZ state, respectively. In panel (a), the horizontal axis is the number of atoms, $N$. Plots in the panel express the scaling of the expected $\delta_1$ and $\delta_2$ with respect to $N$ with a fixed number of measurements, $n = 3000$. $\mathbb{E}[\delta_1]$ for the separable state (solid black) is almost constant, although the other three plots decrease as $N$ becomes large. This is caused by the difference between scalings of $V_{max}$ and $V_n$. Roughly speaking, $V_n$ and $\nabla[\Pi]_{\rho_o}$ has the same scaling with respect to $N$. From Eqs. (D5), (D6), (D7), (D9), (D13), (D16), and (D18), we have the following scalings of $\delta_1$ and $\delta_2$.

$$\delta_1 = \begin{cases} O(1) & \text{(the separable state)} \\ O(1/N) & \text{(the GHZ state)} \end{cases} ,$$

(D20)

$$\delta_2 = \begin{cases} O(1) & \text{(the separable state, small $n$)} \\ O(1/\sqrt{N}) & \text{(the separable state, sufficiently large $n$)} \\ O(1/N) & \text{(the GHZ state)} \end{cases}$$

The scaling of dashed black line is between $O(1/\sqrt{N})$ and $O(1)$, which means that $n = 3000$ is not small and is not sufficiently large. In panel (b), the horizontal axis is the number of measurement trials, $n$. Plots in the panel express the scaling of the expected $\delta_1$ and $\delta_2$ with respect to $n$ with a fixed number of atoms, $N = 100$. Plots for the separable and GHZ states have a same behavior, i.e., $\mathbb{E}[\delta_1] < \mathbb{E}[\delta_2]$ for small $n$ and $\mathbb{E}[\delta_1] > \mathbb{E}[\delta_2]$ for large $n$. For small $n$, a correction term, $\frac{\delta b-a}{N(n-1)} \ln \frac{1}{\epsilon}$, in $\delta_2$ is not negligible, and $\delta_1 < \delta_2$ holds. For large $n$, the correction term becomes negligible, and $\delta_1 > \delta_2$ becomes true because $V_n \leq V_{max}$ holds.

**Appendix E: Partially unknown noises**

In the standard analysis of noisy quantum metrology, it is assumed that a model of the noise is correct, and that a value of a parameter characterizing the noise is known \([14, 18, 20, 21, 23, 26]\). This is the case that the noise is perfectly known. In experiments, however, it is unusual that the noise is perfectly known.

Here let us consider the case that the noise is partially unknown, i.e., the noise model is correct, but we do not know the value of a noise parameter in the model. Suppose that there is an imperfection of the preparation of initial state and that it is characterized by a noise model with a parameter $\eta_1$. The time evolution is characterized.
by the true parameter of interest $\phi$ and noise parameter $\eta_2$. There is an imperfection in the measurement apparatus, and it is characterized by a noise model with a parameter $\eta_3 = (\eta_1, \eta_2, \eta_3)$, which is a LS estimate calculated from data and incorrect values of $\eta_3$, but we do not know a region $E$ including the true noise parameters, i.e., $\eta \in E$. In this case, the probability distribution of the measurement outcome is given by

$$ p(x|\phi, \eta) = \text{Tr} [\kappa_{\phi,\eta_2}(\rho_0) \Pi_{x,\eta_3}] . \tag{E1} $$

We know the function form of the probability distribution, but we do not know the true values of $\phi$ and $\eta$. Then the functional forms of $f$ and $g$ depend on the values of $\eta$, and the value of $\delta$ depends on $\eta$ as well. To clarify this noise-dependency of $\phi_n^{LS}$ and $\delta$, let us use new notations, $\phi_n^{LS}(x^n, \eta)$ and $\delta(x^n, \epsilon, \eta)$.

Let $\eta' := (\eta_1', \eta_2', \eta_3')$ denote the values that we think as the true values of $\eta$. In general, $\eta'$ and $\eta$ are different. We want to evaluate the difference between $\phi_n^{LS}(x^n, \eta')$, which is a LS estimate calculated from data and incorrect noise parameter, and $\phi$. We have

$$ |\phi_n^{LS}(x^n, \eta') - \phi| \leq |\phi_n^{LS}(x^n, \eta) - \phi| + |\phi_n^{LS}(x^n, \eta) - \phi_n^{LS}(x^n, \eta')| \tag{E2} $$

$$ \leq |\phi_n^{LS}(x^n, \eta) - \phi_n^{LS}(x^n, \eta)| + \delta(x^n, \epsilon, \eta) \tag{E3} $$

$$ \leq \max_{\eta \in E} \{ |\phi_n^{LS}(x^n, \eta) - \phi_n^{LS}(x^n, \eta)| + \delta(x^n, \epsilon, \eta) \} \tag{E4} $$

where Eq. \ref{eq:E5} holds with probability at least $1 - \epsilon$. Let us define

$$ \tilde{\delta}(x^n, \epsilon, \eta') := \max_{\eta \in E} \{ |\phi_n^{LS}(x^n, \eta') - \phi_n^{LS}(x^n, \eta)| + \delta(x^n, \epsilon, \eta) \} \tag{E5} $$

We obtain the following theorem.

**Lemma 3** For any number of measurement trials $n \geq 2$ and user-specified constant $0 < \epsilon < 1$,

$$ |\phi_n^{LS}(x^n, \eta') - \phi| \leq \tilde{\delta}(x^n, \epsilon, \eta') \tag{E6} $$

holds with probability at least $1 - \epsilon$.

Lemma 3 provides an exact confidence interval for quantum metrology with partially unknown noise.

The first term in the R.H.S. of Eq. \ref{eq:E6} is the effect of the partially unknown noise. This is a systematic error. The second term in the R.H.S. of Eq. \ref{eq:E6} corresponds to the statistical error. When a noise is partially unknown and we choose an incorrect value for the noise parameter, any estimator $\phi^{est}$ cannot converge to the true parameter $\phi$. So, when $n$ goes to infinity, $\delta$ converges to $0$ but $\tilde{\delta}$ does not. To avoid this problem in the case that the noise is partially unknown, we need to estimate the parameter of interest $\phi$ and noise parameters $\eta$ both. This simultaneous estimation of $\phi$ and $\eta$ is a theoretically interesting and practically important problem, but it is out of the main topic of this paper.

---

[1] V. Giovannetti, S. Lloyd, and L. Maccone, Science \textbf{306}, 1330 (2004).
[2] V. Giovannetti, S. Lloyd, and L. Maccone, Nature Photonics \textbf{5}, 222 (2011).
[3] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński (2014), quant-ph/1202.5270.
[4] J. G. Rarity, P. R. Tapster, E. Jakeman, T. Larchuk, R. A. Campos, M. C. Teich, and B. E. A. Saleh, Phys. Rev. Lett. \textbf{65}, 1348 (1990).
[5] V. Meyer, M. A. Rowe, D. Kielpinski, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland, Phys. Rev. Lett. \textbf{86}, 5870 (2001).
[6] A. Widera, O. Mandel, M. Greiner, S. Kreim, T. W. Hänsch, and I. Bloch, Phys. Rev. Lett. \textbf{92}, 160406 (2004).
[7] Note1, there are two types of the CRBs. One is about for finite data, and the other is for infinite data. The sufficient and necessary condition for the attainability of the CRBs for finite data that the probability distribution of measurement outcomes is an affine function of the parameter $\phi$. In quantum metrology the probability distribution is not affine of $\phi$, and the CRBs are not attainable for finite data. It becomes attainable at infinite data.
[8] Note2, note that $n$ and $N$ are different. $N$ is the number of particles in a probe system used for each measurement trial, and $n$ is the number of measurement trials.
[9] C. R. Rao, \textit{Linear Statistical Inference and Its Applications}, Wiley series in probability and statistics (Wiley, New York, 2002), 2nd ed., (originally published in 1973).
[10] R. R. Bahadur, Sankhyā \textbf{22}, 229 (1960).
[11] W. Hoeffding, J. Am. Stat. Assoc. \textbf{58}, 13 (1963).
[12] A. Maurer and M. Pontil, in \textit{COLT 09: Proc. of The 22nd Annual Conference on Learning Theory} (2009).
[13] E. L. Lehmann and J. P. Romano, \textit{Testing Statistical Hypotheses} (Springer, New York, 2005). 3rd ed.
[14] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Phys. Rev. Lett. \textbf{79}, 3865 (1997).
[15] Y. Matsuzaki, S. C. Benjamin, and J. Fitzsimons, Phys. Rev. A \textbf{84}, 012103 (2011).
[16] A. W. Chin, S. F. Huelga, and M. B. Plenio, Phys. Rev. Lett. \textbf{109}, 233601 (2012).
[17] R. Chaves, J. B. Brask, M. Markiewicz, J. Kołodyński, and A. Acín, Phys. Rev. Lett. \textbf{111}, 120401 (2013).
[18] W. Dür, M. Skotiniotis, F. Fröwis, and B. Kraus, Phys. Rev. Lett. \textbf{112}, 080801 (2014).
[19] H. Lee, P. Kok, and J. P. Dowling, Journal of Modern Optics \textbf{49}, 2325 (2002).
[20] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, Nature Physics \textbf{7}, 406 (2011).
[21] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Gutá, Nature Communications \textbf{3}, 1063 (2012).
[22] S. Amari and H. Nagaoka, \textit{Methods of Information Geometry}, Translations of Mathematical Monographs (Americal Mathematical Society, 2000).
[23] Note3, some coefficients in Eq. \ref{eq:E2} are different from...
the corresponding inequality in [12], because we have a proof of Eq. (B2) but we could not prove the original inequality.

[24] D. Leibfried, M. D. Barrett, T. Schaetz, J. Britton, J. Chiaverini, W. M. Itano, J. D. Jost, C. Langer, and D. J. Wineland, Science 304, 1476 (2004).

[25] G. Arrad, Y. Vinkler, D. Aharonov, and A. Retzker, Phys. Rev. Lett. 112, 150801 (2014).

[26] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin, Phys. Rev. Lett. 112, 150802 (2014).