π-Corrected Heisenberg Limit

Wojciech Górecki,[1] Rafał Demkowicz-Dobrzański,[1] Howard M. Wiseman,[2] and Dominic W. Berry[3]
[1]*Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland*
[2]*Centre for Quantum Computation and Communication Technology (Australian Research Council), Centre for Quantum Dynamics, Griffith University, Brisbane, Queensland 4111, Australia*
[3]*Department of Physics and Astronomy, Macquarie University, Sydney, New South Wales 2109, Australia*

(Received 22 July 2019; revised manuscript received 28 November 2019; published 23 January 2020)

We consider the precision $\Delta\varphi$ with which the parameter φ, appearing in the unitary map $U_\varphi = e^{i\varphi\Lambda}$, acting on some type of probe system, can be estimated when there is a finite amount of prior information about φ. We show that, if U_φ acts n times in total, then, asymptotically in n, there is a tight lower bound $\Delta\varphi \geq \pi/[n(\lambda_+ - \lambda_-)]$, where λ_+, λ_- are the extreme eigenvalues of the generator Λ. This is greater by a factor of π than the conventional Heisenberg limit, derived from the properties of the quantum Fisher information. That is, the conventional bound is never saturable. Our result makes no assumptions on the measurement protocol and is relevant not only in the noiseless case but also if noise can be eliminated using quantum error correction techniques.

DOI: 10.1103/PhysRevLett.124.030501

Introduction and statement of result.—The Heisenberg limit (HL) is the central concept for the whole field of quantum metrology research, as it epitomizes the potential of optimal quantum metrology protocols to surpass standard schemes that are restricted by the so-called standard quantum limit (SQL) [1–9]. For the canonical example of interferometry to measure a stationary optical phase, these two limits are expressed in terms of the number of photon passes through the unknown phase. The HL for the estimation precision is conventionally given as $1/n$, while the SQL corresponds simply to the $1/\sqrt{n}$ shot-noise precision limit. This scaling improvement can be achieved using entangled photon states [10] or multiple passes [11], or a combination of both [12]. In all such cases, the essential feature is that phase is being accumulated coherently over the n uses of the probe system, unlike in standard schemes where each probe (photon) interferes only with itself and the whole procedure is repeated n times, gathering statistics that leads to $1/\sqrt{n}$ improvement of precision.

In a generalized phase estimation scenario, evolution of a probe system is given by a unitary $U_\varphi = \exp(i\varphi\Lambda)$, where Λ is an arbitrary Hermitian generator of the transformation. In what follows, we allow the spectrum of Λ to be arbitrary, apart from being bounded from above and from below by λ_+ and λ_-, respectively. Hence, the parameter φ is not necessarily restricted to the $[0, 2\pi)$ interval. Analysis of the problem using the concept of the quantum Fisher information (QFI) and the quantum Cramér-Rao (CR) bound leads to the following SQL and HL, respectively [1]:

$$\Delta\varphi_{\mathrm{SQL}} \geq \frac{1}{\sqrt{kn}(\lambda_+ - \lambda_-)}, \qquad \Delta\varphi_{\mathrm{HL}} \geq \frac{1}{\sqrt{k}n(\lambda_+ - \lambda_-)}, \tag{1}$$

where k is the number of repetitions of the experiment, n is the number of applications of the unitary U_φ in a single repetition of the experiment. The SQL corresponds to the situation when n independent interrogations of the probe system are performed in a single experiment, while the HL takes into account the most general interrogation scheme involving n uses of U_φ. That may include coherent sequential probes and entangled probes, as well as the most general adaptive schemes. Interestingly, in such noiseless unitary parameter estimation scenarios, there is no advantage to using adaptive strategies, as the simplest sequential scenario where the phase is being coherently imprinted on a single probe n times already leads to the above stated HL. The fundamental advantage that entanglement and adaptiveness offer emerges in the CR bound only when noise is present [13].

Importantly, by the nature of the CR bound, the above bounds are guaranteed to be saturable only in the limit of many repetitions, $k \to \infty$. On the other hand, when considering k repetitions, the total number of unitary operations involved is kn. The most general way of using those resources is to allow entanglement between *all*, in which case the HL in terms of total resources is $1/(kn)$ rather than $1/(\sqrt{k}n)$ scaling [14]. That corresponds to using $k = 1$ in Eq. (1), in which case saturability cannot be guaranteed.

In many papers, the saturability discussion focuses on the existence of a measurement for which the Fisher information of the corresponding probabilistic model coincides with the QFI [15]. Such a measurement indeed always exists, defined by the eigenbasis of the symmetric logarithmic derivative operator. However, even with such a measurement chosen, the existence of an estimator that saturates the CR bound is guaranteed in single-shot scenarios only if

 © 2020 American Physical Society

the value of the parameter is known exactly beforehand or in the case of a narrow class of probabilistic models called the exponential family [16].

In this Letter, we prove that the asymptotically tight HL includes an additional π factor,

$$\Delta\varphi_{\rm HL} \geq \frac{\pi}{n(\lambda_+ - \lambda_-)}. \tag{2}$$

More formally, $\lim_{n\to\infty} n\Delta\varphi_{\rm HL} \geq \pi/(\lambda_+ - \lambda_-)$. The condition under which we prove this bound is that any prior information is limited. For example, one could require that there be a prior probability distribution for φ that is bandlimited or piecewise constant, or that the measurement work over some finite region in φ space. Most importantly, the prior information must be independent of n. That is, if you are given more resources n, you cannot change the task by demanding a sharper prior distribution or further restricting the region of validity of your measurement. This requirement is necessary to make the HL a meaningful concept; otherwise, the prior information is comparable to the information from the measurement itself. The π-corrected limit we derive comes precisely from eliminating the influence of any (implicit or explicit) prior.

Equation (2) was conjectured in [17], but the argument was indirect and restricted to the standard parallel qubit phase estimation scheme with Gaussian prior, and the potential impact of adaptiveness was not analyzed. Since the HL is the key benchmark against which any theoretically conceived or experimentally implemented quantum-enhanced strategy is compared, it is essential to phrase it as an actual attainable limit, unlike its most commonly encountered form [Eq. (1)], which is not achievable even in principle. Our result is also timely, given the recent revival of interest in quantum error correction inspired metrological protocols that allow estimation with HL scaling even in the presence of some particular noise types [18–21]. The proper phrasing of the HL is vital not only for idealized noiseless metrological scenarios but also in the case of more realistic noisy ones.

In our analysis, we use the Bayesian approach to estimation, also called random parameter estimation, in which a probability distribution $p(\varphi)$ is given that describes the prior knowledge of φ. In the case where no prior is known, also called nonrandom parameter estimation [22], the usual approach is to require that the measurement be locally unbiased. Local unbiasedness allows one to derive the traditional (not π-corrected) form of the HL given in Eq. (1) and ensures that, typically, the estimator approximately achieves the claimed mean-square error (MSE) over some finite range of φ. However, it does not ensure that this finite region is independent of n. Since the true value of φ cannot be known before making the measurement, a useful measurement must work over a fixed region, even as n increases. Thus, it is appropriate to consider the average of the MSE over some fixed region. Mathematically, that is equivalent to Bayesian estimation with a flat prior over the region, and so there is no loss of generality in using a Bayesian approach.

It is necessary to exclude pathological priors, which could lead to arbitrarily high precision. We do this by requiring the prior to be well approximated by a finite bandwidth function (i.e., a function whose Fourier transform has bounded support). As we will show, this is not a very restrictive condition; in particular, the results will be valid for any prior that may be approximated by a weighted sum of flat priors with nonzero fixed width $\delta > 0$: $p(\varphi) \approx \sum_{l=-\infty}^{\infty} p(l\delta)\Theta(\frac{1}{2}\delta - |\varphi - l\delta|)$, where Θ is the Heaviside step function. That includes the case where the MSE is averaged over a finite region, used to address cases with an unknown prior. Provided the above regularity condition holds, the actual form of the prior becomes irrelevant for sufficiently large n. The intuition behind this result is that, in the limit of $n \to \infty$, the amount of data that can potentially be gathered on the parameter φ is unlimited and hence overwhelms any impact of the prior on the final precision.

Proof of result.—Consider the most general adaptive estimation scheme, Fig. 1(a). Here $|\psi\rangle$ is the input state of a probe system potentially entangled with an arbitrary number of ancillary systems, and V_i, $i \in \{1, \ldots, n\}$ are control unitary operations applied between interrogation steps, where the unknown parameter is imprinted on the probe system. The final state at the output $|\psi_\varphi^n\rangle$ is measured using a generalized measurement described by a positive operator valued measure $\{M_{\tilde{\varphi}} d\tilde{\varphi}\}$, where the index $\tilde{\varphi}$ represents the estimated value of the parameter upon attaining that outcome from the measurement. The minimal expected (in the Bayesian sense) mean-square error in the estimation thus reads

$$\Delta^2\varphi = \min_{|\psi\rangle, \{M_{\tilde{\varphi}}\}, \{V_i\}} \iint d\tilde{\varphi} d\varphi\, p(\varphi) \langle\psi_\varphi^n| M_{\tilde{\varphi}} |\psi_\varphi^n\rangle (\tilde{\varphi} - \varphi)^2. \tag{3}$$

Let us analyze the structure of the state $|\psi\rangle$ as it evolves through the subsequent gates and control operations. Each gate multiplies components of the state, as decomposed in the Λ eigenbasis, by one of the $e^{i\varphi\lambda}$ factors (λ represents some eigenvalue of Λ), while control operations V_i perform a basis change. In the end, after n coherent interrogations of the unknown parameter φ, the final state will have the following structure:

$$|\psi_\varphi^n\rangle = \int_{n\lambda_-}^{n\lambda_+} c(\mu) e^{i\varphi\mu} |g_\mu\rangle d\mu, \tag{4}$$

where $c(\mu)$ are complex amplitudes and $|g_\mu\rangle$ are some normalized vectors, which, in general, will not be orthogonal. The key feature of this state is that it has bandwidth bounded by $n(\lambda_+ - \lambda_-)$, as will any inner products with this state.

Now let $p_L(\varphi)$ be the prior with a finite bandwidth L, so that its Fourier transform is supported on an interval of

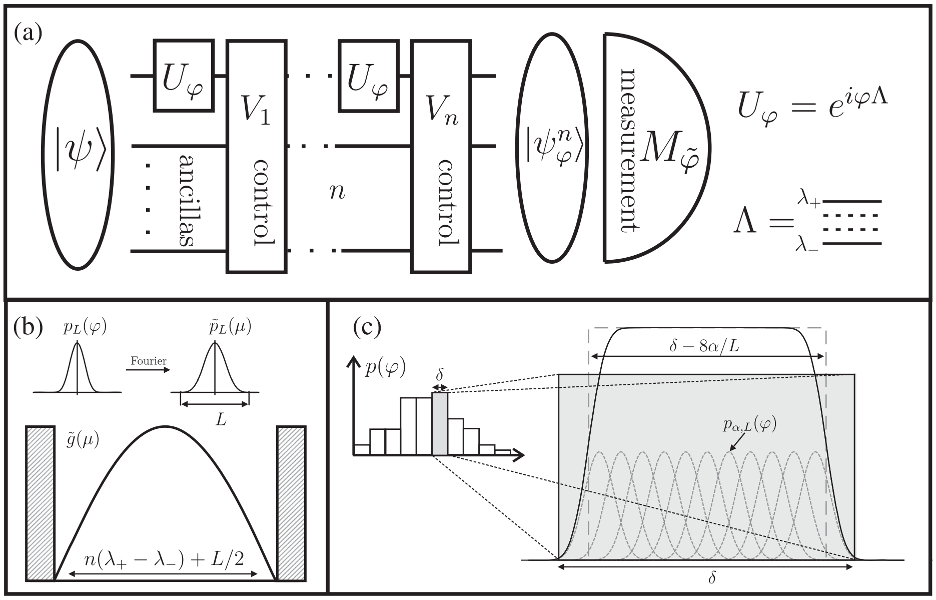

FIG. 1. Graphical representation of the proof. (a) A general adaptive phase estimation protocol with the total number of phase gates n. (b) For a finite bandwidth prior $p_L(\varphi)$ we derive the bound, which is equivalent to finding the minimum energy eigenstate in an infinite potential well with width $n(\lambda_+ - \lambda_-) + L/2$. (c) A rectangular prior may be approximated to any desired accuracy by a convolution of a slightly narrower rectangular prior with a finite bandwidth Kaiser window function, to derive the bound.

length L. It is known that, in all single-generator unitary estimation problems with quadratic cost, one may restrict the class of measurements to rank-one projective ones [23,24]. Let us assume for a moment a fixed input state $|\psi\rangle$ and a measurement basis $\{|\chi\rangle\}$ with a corresponding estimator $\tilde\varphi_\chi$. The corresponding cost reads

$$\Delta^2\varphi = \int d\chi \int d\varphi p_L(\varphi)|\langle\psi_\varphi^n|\chi\rangle|^2(\tilde\varphi_\chi - \varphi)^2. \qquad (5)$$

Now, from [25], it is possible to find a function $w_L(\varphi)$ with bandwidth $L/2$ such that $|w_L(\varphi)|^2 = p_L(\varphi)$. Let us define functions $f_\chi(\varphi) = \langle\psi_\varphi^n|\chi\rangle$ [which have bandwidths bounded by $n(\lambda_+ - \lambda_-)$] and $g_\chi(\varphi) = w_L(\varphi)f_\chi(\varphi)$. We have

$$\Delta^2\varphi = \int d\chi \int d\varphi |g_\chi(\varphi)|^2(\tilde\varphi_\chi - \varphi)^2. \qquad (6)$$

The product of two bandlimited functions gives a function that is bandlimited by the sum of the bandwidths, so $g_\chi(\varphi)$ has a bandwidth at most $n(\lambda_+ - \lambda_-) + L/2$. We can then write

$$\Delta^2\varphi = \int p_\chi C_{\chi,\tilde\varphi_\chi} d\chi \geq \min_\chi C_{\chi,\tilde\varphi_\chi}, \qquad (7)$$

with

$$p_\chi = \int |g_\chi(\varphi)|^2 d\varphi, \qquad (8)$$

$$C_{\chi,\tilde\varphi_\chi} = \frac{1}{p_\chi}\int |g_\chi(\varphi)|^2(\tilde\varphi_\chi - \varphi)^2 d\varphi. \qquad (9)$$

The task of minimizing $C_{\chi,\tilde\varphi_\chi}$ for a given $\tilde\varphi_\chi$ is equivalent to that of minimization for $\tilde\varphi_\chi = 0$, because the optimal $g_\chi(\varphi)$ can be shifted by any amount without altering the bandwidth. We may also set $p_\chi = 1$ (this just sets the normalization of the function g). After applying the Fourier transform, the minimization problem reads

$$\min_{\tilde g(\mu)} \int_{n\lambda_- - L/4}^{n\lambda_+ + L/4} \left|\frac{d\tilde g(\mu)}{d\mu}\right|^2 d\mu, \qquad (10)$$

subject to the constraints

$$\int_{n\lambda_- - L/4}^{n\lambda_+ + L/4} |\tilde g(\mu)|^2 d\mu = 1, \qquad \tilde g(n\lambda_- - L/4) = 0,$$
$$\tilde g(n\lambda_+ + L/4) = 0. \qquad (11)$$

Here the boundary conditions are due to the restricted bandwidth of $g(\varphi)$. This is equivalent to the problem of finding the minimum energy eigenstate in an infinite potential well with width $n(\lambda_+ - \lambda_-) + L/2$. The solution for $\tilde g(\mu)$ that achieves the minimum is a sine curve [see Fig. 1(b)], which gives $\min C_{\chi,\tilde\varphi_\chi} = \pi^2/[n(\lambda_+ - \lambda_-) + L/2]^2$, and hence

$$\Delta^2\varphi \geq \frac{\pi^2}{[n(\lambda_+ - \lambda_-) + L/2]^2}. \qquad (12)$$

That is, provided the prior has a limit L on its bandwidth, the Heisenberg limit (2) holds in the limit of large n—i.e., when the prior correction $L/2$ becomes negligible.

Now we need only prove that any reasonable prior $p(\varphi)$ may be well approximated by a finite bandwidth function. To first show this informally, let us introduce a family of non-negative normalized finite bandwidth functions

$$p_{\alpha,L}(\varphi) = \mathcal{N}_\alpha L \mathrm{sinc}^4\left(\pi\alpha\sqrt{(L\varphi/4\alpha)^2 - 1}\right), \qquad (13)$$

where L is the bandwidth, α is a parameter that controls the size of the tails, and $\mathcal{N}_\alpha$ is a normalization factor that $\approx 4\sqrt{2}\pi^4\alpha^{7/2}e^{-4\pi\alpha}$ for α large. This $p_{\alpha,L}(\varphi)$ is proportional to the fourth power of the Fourier transform of the Kaiser window [26] of width $L/4$, so has bandwidth L. The function $p_{\alpha,L}(\varphi)$ has width $8\alpha/L$, beyond which the tails are exponentially suppressed in α, like $e^{-4\pi\alpha}$. Thus, for large L it approximates the Dirac delta function and therefore any reasonable prior may be approximated by its convolution with $p_{\alpha,L}$, i.e., $p(\varphi) \approx (p_{\alpha,L} * p)(\varphi)$, for which the bandwidth is also L.

In particular, we may show that, for any prior of the form $p(\varphi) \approx \sum_{l=-\infty}^{\infty} p(l\delta)\Theta(\frac{1}{2}\delta - |\varphi - l\delta|)$, the following bound holds:

$$\Delta^2\varphi \geq \frac{\pi^2}{[n(\lambda_+ - \lambda_-)]^2}\left(1 - \sqrt{\frac{8\log\Delta}{\Delta}}\right), \qquad (14)$$

where $\Delta = n(\lambda_+ - \lambda_-)\delta$. This gives

$$\lim_{n\to\infty} n^2\Delta^2\varphi_{\mathrm{HL}} \geq \frac{\pi^2}{(\lambda_+ - \lambda_-)^2}, \qquad (15)$$

which after taking the square root of both sides proves Eq. (2).

Here we just sketch the reasoning leading to the above claim, whereas the complete proof with all the technical details is presented in the Supplemental Material [27]. First, we lower bound the minimal cost in the case that the prior is actually a single rectangular prior of width δ at $l = 0$. Note that this will also be a legitimate lower bound for the original problem where the prior is a weighted sum of such rectangular priors, as the optimal strategy for this original problem cannot perform better than the optimal strategy when we additionally know to which δ interval the value of our parameter is restricted. Then, we approximate the rectangular prior via a distribution obtained by convolving a slightly narrower rectangular distribution (narrower by $8\alpha/L$) with the $p_{\alpha,L}$ function, as defined in Eq. (13) [see Fig. 1(c)]. By setting α and L appropriately, we can guarantee that all the deviations of the resulting cost due to modifications of the prior from the strictly rectangular one introduce no more than $\tilde{\mathcal{O}}(\Delta^{-1/2})$ ($\tilde{\mathcal{O}}$ indicates that logarithmic multipliers are ignored) relative correction compared with the cost corresponding to the $p_{\alpha,L}$ distribution. Finally, the cost corresponding to the $p_{\alpha,L}$ distribution can be bounded using Eq. (12) thanks to the finite bandwidth property of $p_{\alpha,L}$.

In the case of nonrandom parameter estimation [22], we average the MSE over a φ interval of size δ. This is equivalent to using a flat prior of width δ, so our bound holds in this case also, and this is true regardless of whether or not the estimators are unbiased. The conventional scenario of unbiased estimators with no average does typically mean that the measurement works well for φ within some finite region. However, the size of that region may depend on n. That is, as n increases, the region of validity for an unbiased measurement may shrink with n. This is exactly the case for the NOON states [28] that maximize the CR bound, saturating the conventional HL in Eq. (1). For these states, the probability distributions obtained are periodic in φ with period π/n, and the interval in φ where the CR-saturating measurement is useful shrinks exactly as fast in n as the CR bound itself.

To reiterate, to make the HL a meaningful concept, the bound should not rely on using more and more prior information as n increases, because that would mean that the prior information can be comparable to the information from the measurement itself. To derive a HL that describes the information obtained from the measurement, one must do as we have done and take a prior that is independent of n or require a region of validity that is independent of n. Then the information about φ in the limit $n \to \infty$ comes only from the measurement on the state. Eliminating the influence of prior information is what gives our additional factor of π.

Discussion.—Having derived the bound, let us now discuss its saturability. It is known that, in the case of a standard phase estimation problem with $u_\varphi = e^{i\varphi\sigma_z/2}$ applied in parallel to n qubits, the optimal Bayesian strategy for flat prior distribution $p(\varphi) = 1/(2\pi)$ ($\varphi \in [-\pi, \pi]$) in the limit of large n yields $\Delta\varphi \to \pi/n$ [29]. Note also, that the optimal strategy involves application of the so-called covariant measurements [30], and a covariant measurement strategy will yield the same average cost irrespective of the form of prior. Hence the π/n limit is saturable in the case of an arbitrary prior as well. In the case of an estimation problem with a general generator Λ, we can say that, provided the prior is supported on an interval smaller than $2\pi/(\lambda_+ - \lambda_-)$, we can directly adapt the reasoning from the standard qubit phase estimation scheme by considering our elementary system as a qubit with only two accessible states being the eigenstates of Λ corresponding to λ_+ and λ_-. This way we obtain $\Delta\varphi_{\mathrm{HL}} = \pi/[n(\lambda_+ - \lambda_-)]$. However, if our prior is broader, then clearly using this strategy we will not be able to discriminate between phases that differ by a multiple of $2\pi/(\lambda_+ - \lambda_-)$, as they effectively would lead to the same output state. In order to discriminate between these phases, we would need to use additional eigenstates of Λ

corresponding to intermediate eigenvalues λ (provided they are available). If we use levels corresponding to eigenvalues that differ by ϵ, we may discriminate between all phases that differ by less than $2\pi/\epsilon$. Note that, for our purposes, the minimal level splitting ϵ may be effectively obtained as a difference between sums of a certain finite number of energy levels $\epsilon = \sum_{i\in\{i_1,\ldots,i_s\}} \lambda_i - \sum_{j\in\{j_1,\ldots,j_s\}} \lambda_j$, and a result may be smaller than the minimal level splitting in the Λ itself. Since the discrimination error drops exponentially with the number of resources used, we may sacrifice sublinear (in n) uses of the channel for the purpose of this additional discrimination task and this will not affect the final scaling.

The phase estimation problem may be viewed as a special case of a more general frequency estimation problem, where the probe system is allowed to be interrogated for the total interrogation time T and the goal is to estimate a frequencylike parameter ω entering into the Hamiltonian of the system as $H = \omega G$, with G being some Hermitian operator. The total interrogation time T may be split into a number of shorter evolution steps each lasting time $t = T/n$. Assuming the prior distribution $p(\omega)$ satisfies the regularity assumption and can be written as a sum of rectangular priors of some finite width δ_ω, we may repeat the whole reasoning as presented above by formally identifying $\varphi = \omega t$, $\Lambda = -G$, and $n = T/t$ and arrive at

$$\Delta\omega \geq \frac{\pi}{T(\lambda_+ - \lambda_-)}. \tag{16}$$

This is the valid asymptotically saturable bound for the most general frequency estimation adaptive strategies in the limit of long total interrogation time T. In particular, in all the cases where, despite the presence of noise, the Heisenberg scaling is being recovered via, e.g., application of quantum error correction inspired techniques [18–21,31–33], it is the above bound that should be used as a operationally meaningful figure of merit of such protocols and not the standard QFI-based one.

Finally, since Eq. (14) is not tight for finite n, the form of the exact achievable bound in the nonasymptotic case is an interesting open question for future research.

We thank Michael Hall, Marcin Jarzyna, Paweł Kasprzak, Jan Kolodynski, and Yuval Sanders for fruitful discussions. W. G. and R. D. D. acknowledge support from the National Science Center (Poland) Grant No. 2016/22/E/ST2/00559. D. W. B. is funded by Australian Research Council Discovery Projects No. DP160102426 and No. DP190 102633. H. M. W. is funded by the Australian Research Council Centre of Excellence Program CE170100012.

[1] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).
[2] M. G. A. Paris, Int. J. Quantum. Inform. **07**, 125 (2009).
[3] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Photonics **5**, 222 (2011).
[4] G. Toth and I. Apellaniz, J. Phys. A **47**, 424006 (2014).
[5] R. Demkowicz-Dobrzanski, M. Jarzyna, and J. Kołodyński, in *Progress in Optics*, edited by E. Wolf (Elsevier, New York, 2015), Vol. 60, pp. 345–435.
[6] R. Schnabel, Phys. Rep. **684**, 1 (2017).
[7] C. L. Degen, F. Reinhard, and P. Cappellaro, Rev. Mod. Phys. **89**, 035002 (2017).
[8] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Rev. Mod. Phys. **90**, 035005 (2018).
[9] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Nat. Photonics **12**, 724 (2018).
[10] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, Nature (London) **429**, 161 (2004).
[11] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Nature (London) **450**, 393 (2007).
[12] S. Daryanoosh, S. Slussarenko, D. W. Berry, H. M. Wiseman, and G. J. Pryde, Nat. Commun. **9**, 4606 (2018).
[13] R. Demkowicz-Dobrzański and L. Maccone, Phys. Rev. Lett. **113**, 250801 (2014).
[14] M. Hayashi, S. Vinjanampathy, and L. C. Kwek, J. Phys. B **52**, 015503 (2018).
[15] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
[16] A. V. Fend, Ann. Math. Stat. **30**, 381 (1959).
[17] M. Jarzyna and R. Demkowicz-Dobrzaski, New J. Phys. **17**, 013010 (2015).
[18] P. Sekatski, M. Skotiniotis, J. Kołodyński, and W. Dür, Quantum **1**, 27 (2017).
[19] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Phys. Rev. X **7**, 041009 (2017).
[20] D. Layden and P. Cappellaro, npj Quantum Inf. **4**, 30 (2018).
[21] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Nat. Commun. **9**, 78 (2018).
[22] S. M. Kay, in *Fundamentals of Statistical Signal Processing: Estimation Theory*, edited by S. E. Alan V. Oppenheim (Prentice Hall, Englewood Cliffs, NJ, 1993).
[23] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[24] K. Macieszczak, M. Fraas, and R. Demkowicz-Dobrzaski, New J. Phys. **16**, 113002 (2014).
[25] R. Boas, Jr, M. Kac *et al.*, Duke Math. J. **12**, 189 (1945).
[26] J. Kaiser and R. Schafer, IEEE Trans. Acoust. Speech Signal Process. **28**, 105 (1980).
[27] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.124.030501 for technical details of the proof.
[28] P. Kok, H. Lee, and J. P. Dowling, Phys. Rev. A **65**, 052104 (2002).
[29] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. **85**, 5098 (2000).
[30] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam, 1982).
[31] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin, Phys. Rev. Lett. **112**, 150802 (2014).
[32] W. Dür, M. Skotiniotis, F. Frowis, and B. Kraus, Phys. Rev. Lett. **112**, 080801 (2014).
[33] G. Arrad, Y. Vinkler, D. Aharonov, and A. Retzker, Phys. Rev. Lett. **112**, 150801 (2014).