

General Optimality of the Heisenberg Limit for Quantum Metrology

Marcin Zwierz,¹ Carlos A. Pérez-Delgado,^{1,2} and Pieter Kok^{1,*}

¹*Department of Physics and Astronomy, University of Sheffield, Hounsfield Road, Sheffield, S3 7RH, United Kingdom*

²*Department of Physics and Astronomy, University of Sussex, Falmer, Brighton, East Sussex, BN1 9QH, United Kingdom*

(Received 3 June 2010; published 27 October 2010)

Quantum metrology promises improved sensitivity in parameter estimation over classical procedures. However, there is a debate over the question of how the sensitivity scales with the resources and the number of queries that are used in estimation procedures. Here, we reconcile the physical definition of the relevant resources used in parameter estimation with the information-theoretical scaling in terms of the query complexity of a quantum network. This leads to a completely general optimality proof of the Heisenberg limit for quantum metrology. We give an example of how our proof resolves paradoxes that suggest sensitivities beyond the Heisenberg limit, and we show that the Heisenberg limit is an information-theoretic interpretation of the Margolus-Levitin bound, rather than Heisenberg's uncertainty relation.

DOI: 10.1103/PhysRevLett.105.180402

PACS numbers: 03.65.Ta, 03.67.-a, 42.50.Lc

Parameter estimation is a fundamental pillar of science and technology. Caves [1] showed that quantum mechanical systems can, in principle, produce greater sensitivity over classical methods, and many quantum parameter estimation protocols have been proposed since [2]. The field of quantum metrology started with the work of Helstrom [3,4], who derived the minimum value for the mean square error in a parameter in terms of the density matrix of the quantum system and a measurement procedure. This was a generalization of a known result in the classical parameter estimation, called the Cramér-Rao bound. Braunstein and Caves [5] showed how this bound can be formulated for the most general state preparation and measurement procedures.

The quantum Cramér-Rao bound is typically formulated in terms of the Fisher information. One of the central questions in quantum metrology is how the Fisher information scales with the physical resources used in the measurement procedure. We usually consider two scaling regimes: First, in the *standard quantum limit* (SQL) [6] or *shot-noise limit* the Fisher information is constant, and the error scales with the inverse square root of the number of times T we make a measurement. Second, in the *Heisenberg limit* [7] the error is bounded by the inverse of the physical resources. Typically, these are expressed in terms of the size N of the probe system, e.g., (average) photon number. However, it has been clearly demonstrated that this form of the limit is not universally valid. For example, Beltrán and Luis [8] showed that the use of classical optical nonlinearities can lead to an error with average photon number scaling $N^{-3/2}$. Boixo *et al.* [9] devised a parameter estimation procedure that sees the error scale with N^{-k} with $k \in \mathbb{N}$, and Roy and Braunstein [10] constructed a procedure that achieves an error that scales with 2^{-N} . The central question is then, what is the real fundamental Heisenberg limit for quantum

metrology? We could redefine this limit accordingly to scale as 2^{-N} , but in practice this bound will never be tight.

In this Letter, we give a natural definition of the relevant physical resources for quantum metrology based on the general description of a parameter estimation procedure, and we prove the fundamental bound on the mean squared error based on this resource count. We will show that the resource count is proportional to the size of the probe system only if the interaction between the object and the probe is nonentangling over the systems constituting the probe. First, we study the query complexity of quantum metrology networks, which will lead to a resource count given by the expectation value of the generator of translations in the parameter φ . Second, we prove that the mean error in φ is bounded by the inverse of this resource count. We argue that this is the fundamental Heisenberg limit for quantum metrology. Furthermore, we show that it is a form of the Margolus-Levitin bound, as opposed to Heisenberg's uncertainty relation. Finally, we illustrate how this general principle can resolve paradoxical situations in which the Heisenberg limit seems to be surpassed.

The most general parameter estimation procedure is shown in Fig. 1(a). Consider a probe system prepared in an initial quantum state $\rho(0)$ that is evolved to a state $\rho(\varphi)$ by $U(\varphi) = \exp(-i\varphi\mathcal{H})$. This is a unitary evolution when we include the relevant environment in our description, and it includes feed-forward procedures. The Hermitian operator \mathcal{H} is the generator of translations in φ , the parameter we wish to estimate. The system is subjected to a generalized measurement M , described by a positive operator valued measure (POVM) that consists of elements \hat{E}_x , where x denotes the measurement outcome. These can be discrete or continuous (or a mixture of both). The probability distribution that describes the measurement data is given by the Born rule $p(x|\varphi) = \text{Tr}[\hat{E}_x\rho(\varphi)]$, and the maximum amount of information

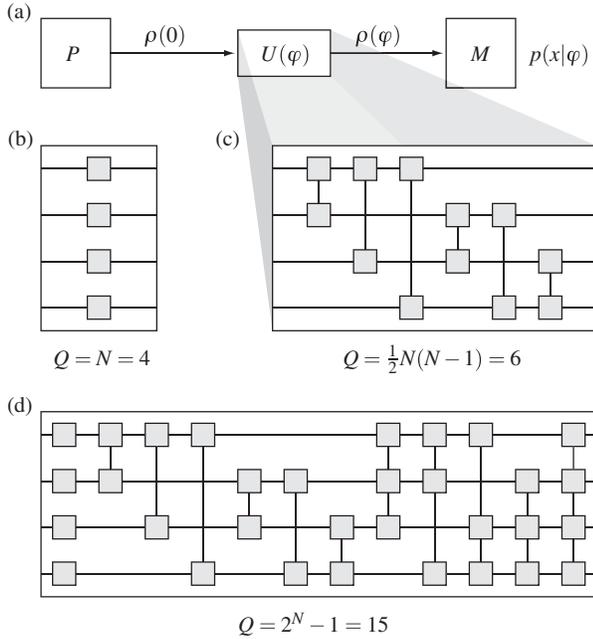


FIG. 1. (a) General parameter estimation procedure involving state preparation P , evolution $U(\varphi)$, and generalized measurement M with outcomes x , which produces a probability distribution $p(x|\varphi)$. In terms of quantum networks, the evolution can be written as a number of queries of the parameter φ . (b) Example for $N = 4$ of the usual situation described by \mathcal{H}_{GLM} [the grey box represents $O_j(\varphi)$]. (c) For $\mathcal{H}_{\text{BFCG}}$ the number of queries Q does not always equal the number of systems. (d) For \mathcal{H}_{RB} all possible subsets of systems perform a single query. The number of queries scales exponentially with the number of systems.

about φ that can be extracted from this measurement is given by the Fisher information

$$F(\varphi) = \int dx \frac{1}{p(x|\varphi)} \left(\frac{\partial p(x|\varphi)}{\partial \varphi} \right)^2. \quad (1)$$

This leads to the quantum Cramér-Rao bound [3,5]

$$\delta\varphi \geq \frac{1}{\sqrt{TF(\varphi)}}, \quad (2)$$

where $(\delta\varphi)^2$ is the mean square error in the parameter φ , and T is the number of times the procedure is repeated. The SQL is obtained when the Fisher information is a constant with respect to T , and the Heisenberg limit is obtained in a single-shot experiment ($T = 1$) when the Fisher information scales quadratically with the resource count. The SQL and the Heisenberg limit therefore relate to two fundamentally different quantities, T and F , respectively. We need to reconcile the meaning of these two limits if we want to compare them in a meaningful way.

To solve this problem, we can define an unambiguous resource count for parameter estimation by recognizing that a quantum parameter estimation protocol can be written as a quantum network acting on a set of quantum systems, with repeated “black-box” couplings of the network to the system we wish to probe for the parameter φ [11]. A black box is a function that can be univariate or

multivariate. When the function is multivariate, a *query* to the black box must take the form of multiple input parameters. Likewise, when the operator that describes the fundamental interaction between the queried system and the probe is a two-body interaction, then a query can consist only of precisely two input bodies. The scaling of the error in φ is then determined by the *query complexity* of the network. The number of queries Q is not always identical to the number of physical systems N in the network.

In Figs. 1(b)–1(d) we consider three examples. The quantum network in (b) was analyzed by Giovannetti, Lloyd, and Maccone [11]. Suppose that each grey box in Fig. 1 is a unitary gate $O_j(\varphi) = \exp(-i\varphi H_j)$, where $j = 1, \dots, N$ denotes the system, and H_j is a positive Hermitian operator. It is convenient to define the generator of the joint queries as $\mathcal{H}_{\text{GLM}} = \sum_j H_j$, because all H_j commute with each other. The number of queries Q is then equal to the number of terms in \mathcal{H}_{GLM} , or $Q = N$. In Fig. 1(c) the black box is bivariate. This is a type of Hamiltonian considered by Boixo, Flammia, Caves, and Geremia [9], and takes the form $\mathcal{H}_{\text{BFCG}} = \sum_{k=1}^N \sum_{j=1}^k H_j \otimes H_k$. A physical query to a black box characterized by $O_{jk} = \exp(-i\varphi H_j \otimes H_k)$ must consist of two systems, labeled j and k . Since each pair interaction is a single query, the total number of queries is $\binom{N}{2}$. Finally, in Fig. 1(d) we depict the network corresponding to the protocol by Roy and Braunstein [10]. The number of terms in the corresponding generator \mathcal{H}_{RB} is given by $2^N - 1$, and the number of queries is therefore $Q = 2^N - 1$. The entangling power of the black-box operation over multiple input systems accounts for the superlinear scaling of Q with N . Only when \mathcal{H} does not have any entangling power across the input are we guaranteed to have $Q = O(N)$. This is in agreement with Refs. [9,10], where $\sqrt{F(\varphi)}$ scales superlinearly in N but is always linear in Q , as defined here. Since we have a systematic method for increasing N (and Q) given the atomic interaction H_j , this uniquely defines the asymptotic query complexity of the network. Since both T and Q count the number of queries, this allows us to meaningfully compare the SQL with the Heisenberg limit.

Given that in Eq. (2) $\sqrt{F(\varphi)} \leq Q$, we have to find a general procedure that bounds Q , based on the physical description of the estimation protocol in Fig. 1(a). Previously, we showed that Q is the number of black-box terms in \mathcal{H} , and a straightforward choice for the resource count is therefore $|\langle \mathcal{H} \rangle| \leq O(Q)$. An important subtlety occurs when \mathcal{H} corresponds to a proper Hamiltonian. The origin of the energy scale has no physical meaning, and the actual value of $|\langle \mathcal{H} \rangle|$ can be changed arbitrarily. Hence, we must fix the scale such that the ground state has zero energy (equivalently, we may choose $|\langle \mathcal{H} - h_{\min} I \rangle|$, where h_{\min} is the smallest eigenvalue and I the identity operator). In most cases, this is an intuitive choice. For example, it is natural to associate zero energy to the vacuum state and add the corresponding amount of energy

for each added photon. Technically, this corresponds to the normal ordering of the Hamiltonian of the radiation field in order to remove the infinite vacuum energy. Slightly less intuitive is that the average energy of N spins in a Greenberger-Horne-Zeilinger state $(|\uparrow\rangle^{\otimes N} + |\downarrow\rangle^{\otimes N})/\sqrt{2}$ is no longer taken to be zero, but rather $N/2$ times the energy splitting between $|\uparrow\rangle$ and $|\downarrow\rangle$.

Other seemingly natural choices are the variance and the seminorm. For example, if we write $\mathcal{H} \equiv \sum_j^Q A_j$, the variance is $(\Delta\mathcal{H})^2 = \langle(\sum_j^Q A_j)^2\rangle - \langle\sum_j^Q A_j\rangle^2 = \sum_j^Q \langle L_j \rangle - \sum_{j,k}^Q \langle A_j \rangle \langle A_k \rangle \leq cQ^2$ for some positive number c and positive operator L_j . This gives $\Delta\mathcal{H} \leq O(Q)$, where, e.g., in Ref. [9] $Q = O(N^2)$. Similarly, $|\langle\mathcal{H}\rangle| = \sum_j^Q |\langle A_j \rangle| \leq O(Q)$ since all expectation values are positive and finite. In other words, in terms of the scaling behavior with Q , we can use either the variance or the expectation value. However, there are important classes of quantum systems for which the variance of the energy diverges, such as systems with a Breit-Wigner (or Lorentzian) spectrum [12,13]. Similarly, the seminorm does not exist for a large class of states, such as optical Gaussian states. In these cases the resource count, and by implication the scaling of the error, would be ill defined. By contrast, $|\langle\mathcal{H}\rangle|$ always exists and is always positive. Also, from a physical perspective the higher-order moments do not describe ‘‘amounts’’ in the same way as the first moment does, and they refer instead to the shape of the distribution. This is a further argument that $|\langle\mathcal{H}\rangle|$ is the natural choice for the resource count. Sometimes, it is unclear how the query complexity is defined, for example, when the estimation procedure does not involve repeated applications of the gates $O_j(\varphi)$, or when an indeterminate number of identical particles, such as photons, are involved. Nevertheless, the generator \mathcal{H} is always well defined in any estimation procedure, and we can always use its expectation value to define the relevant resource count.

The resource count in terms of $|\langle\mathcal{H}\rangle|$ is completely general for all possible quantum networks. For interactions $U(\varphi)$ where we include feed-forward and arbitrary unitary gates between queries, we can use an argument by Giovannetti *et al.* [11] to show that $|\langle\mathcal{H}\rangle| = |\langle(\partial U(\varphi)/\partial\varphi)U^\dagger(\varphi)\rangle|$ is unaffected by the intermediate unitary gates, and the scaling is still determined by Q .

After establishing the appropriate resource count, we can prove the optimality of the Heisenberg limit in its most general form. The Fisher information can be related to a statistical distance s on the probability simplex spanned by $p(x|\varphi)$. Consider two probability distributions $p(x|\varphi)$ and $p(x|\varphi) + dp(x)$. The infinitesimal statistical distance between these distributions is given by $ds^2 = \int dx [dp(x|\varphi)]^2 / p(x|\varphi)$ [14,15]. Dividing both sides by $(d\varphi)^2$ we obtain

$$\left(\frac{ds}{d\varphi}\right)^2 = \int dx \frac{1}{p(x|\varphi)} \left(\frac{\partial p(x|\varphi)}{\partial\varphi}\right)^2 = F(\varphi), \quad (3)$$

which relates the Fisher information to the rate of change of the statistical distance.

When we count the resources that are used in a parameter estimation procedure, we must make sure that we do not leave anything out, and this can be guaranteed by including in our description the environment that the estimation procedure couples to. This reduces the quantum states to pure states, which means that we can use Wootters’ distance [15] between quantum states as the statistical distance:

$$s(\psi, \phi) = \arccos(|\langle\psi|\phi\rangle|), \quad (4)$$

where $|\psi\rangle$ and $|\phi\rangle$ are two pure states in the larger Hilbert space, and $s(\psi, \phi)$ is the angle between them. The distance between the probe state $\rho(0)$ and the evolved state $\rho(\varphi)$ can then be represented by the pure states $|\psi(0)\rangle$ and $|\psi(\varphi)\rangle$, respectively, and the unitary evolution is given by

$$|\psi(\varphi)\rangle = \exp(-i\varphi\mathcal{H})|\psi(0)\rangle. \quad (5)$$

Here, we place no restriction on \mathcal{H} , other than fixing the energy scale if necessary. We can place an upper bound on the derivative of Wootters’ distance by evaluating the differential of s in Eq. (4) and using Eq. (5) [16]:

$$\frac{ds}{d\varphi} \leq |\langle\mathcal{H}\rangle|. \quad (6)$$

Combining this with Eqs. (2) and (3) leads to the Cramér-Rao bound

$$(\delta\varphi)^2 \geq \frac{1}{T} \left(\frac{ds}{d\varphi}\right)^{-2} \geq \frac{1}{T|\langle\mathcal{H}\rangle|^2}. \quad (7)$$

When all resources are used in a single-shot ($T = 1$) experiment, the error in φ is bounded by

$$\delta\varphi \geq \frac{1}{|\langle\mathcal{H}\rangle|}. \quad (8)$$

Since $|\langle\mathcal{H}\rangle|$ is the resource count in the parameter estimation procedure, this is the Heisenberg limit. It is always positive and finite; in the limit where $|\langle\mathcal{H}\rangle| \rightarrow 0$, there are no resources available to estimate φ , and $\delta\varphi$ cannot be bounded. In general, the bound is not tight. Indeed, only carefully chosen entangled systems can achieve this bound [11]. This completes the proof of the optimality of the Heisenberg limit in the most general case.

In addition to Eqs. (3) and (6), the Fisher information is also bounded by the variance of \mathcal{H} according to $F(\varphi) \leq 4(\Delta\mathcal{H})^2$ [17]. This leads to a (single-shot) quantum Cramér-Rao bound

$$\delta\varphi \geq \frac{1}{2\Delta\mathcal{H}}. \quad (9)$$

However, since $\Delta\mathcal{H}$ is not a resource count, such as the average photon number, but rather a variance (or uncertainty), this is not the Heisenberg limit. In fact, it is Heisenberg’s uncertainty relation for the parameter φ and its conjugate operator \mathcal{H} . Any parameter estimation procedure must respect both bounds, and the Heisenberg limit

in Eq. (8) may not be attained for a particular input state because the bound in Eq. (9) prevents it from doing so.

The term ‘‘Heisenberg limit’’ was introduced by Holland and Burnett [7], who referred to the number-phase uncertainty relation in Heitler’s work [18]. However, as our optimality proof and the subsequent discussion indicate, the Heisenberg limit is *not* an uncertainty relation, since it relates the uncertainty of the parameter to the *first* moment of the conjugate observable \mathcal{H} , rather than the second. It turns out, instead, that the Heisenberg limit is intimately connected to the Margolus-Levitin bound on the time it takes for a quantum system to evolve to an orthogonal state [16,19,20]. To see this, we can formally solve Eq. (6) by separation of variables, yielding

$$\int_0^\varphi d\varphi' \geq \frac{1}{|\langle \mathcal{H} \rangle|} \int_0^{\pi/2} ds \Rightarrow \varphi \geq \frac{\pi}{2} \frac{1}{|\langle \mathcal{H} \rangle|}. \quad (10)$$

We can therefore identify the Heisenberg limit with the Margolus-Levitin bound on the speed of dynamical evolution of quantum systems when \mathcal{H} is the Hamiltonian. The (generalized) uncertainty relation, on the other hand, can be identified with the Mandelstam-Tamm bound [16]. Both limits are completely general (assuming the existence of $\Delta\mathcal{H}$) and complement each other.

Finally, we demonstrate that our proof applies to continuous variable systems as well as discrete systems, by considering the procedure of Beltrán and Luis [8]. The construction is as follows: The evolution $O(\varphi)$ is generated by an optical nonlinearity proportional to the square of the photon number operator \hat{n}^2 acting on a single-mode coherent state $|\psi(0)\rangle = |\alpha\rangle$. The evolved state before detection is given by $|\psi(\varphi)\rangle = \exp(-i\varphi\hat{n}^2)|\alpha\rangle$, and the mean square error in φ is calculated as $\delta\varphi \simeq \frac{1}{4}\langle\hat{n}\rangle^{-3/2} = \frac{1}{4}|\alpha|^{-3}$, to leading order in the average photon number $\langle\hat{n}\rangle$. Since here the average energy is directly proportional to the average photon number, this procedure seems to surpass the Heisenberg limit. To resolve this paradox, we note that the generator of translations in φ is *not* the photon number operator \hat{n} , but rather the higher-order nonlinearity $\mathcal{H} = \hat{n}^2$. The appropriate resource count is therefore $|\langle\mathcal{H}\rangle| = \langle\hat{n}^2\rangle$, instead of the average photon number $\langle\hat{n}\rangle$. It is easily verified that to leading order $\delta\varphi$ is theoretically bounded by $1/\langle\hat{n}^2\rangle = 1/|\alpha|^4$. Hence the parameter estimation procedure does not even attain the Heisenberg limit.

Formally, we can attain the Heisenberg limit in this setup with the following modification of the input state and the measurement. Consider the single-mode input state $|\psi_0\rangle = (|0\rangle + |N\rangle)/\sqrt{2}$, where $|0\rangle$ denotes no photons, and $|N\rangle$ denotes N photons. The state of the probe before detection is then given by $|\psi(\varphi)\rangle = \exp(-i\varphi\hat{n}^2)|\psi_0\rangle = (|0\rangle + e^{-i\varphi N^2}|N\rangle)/\sqrt{2}$. We define the measurement observable $X = |0\rangle\langle N| + |N\rangle\langle 0|$. Hence, for the final state $|\psi(\varphi)\rangle$ we calculate $\langle X \rangle = \langle \psi_\varphi | X | \psi_\varphi \rangle = \cos(N^2\varphi)$ and $\Delta X = \sin(N^2\varphi)$. Using the standard expression for the mean

squared error, we find that $\delta\varphi = \Delta X |d\langle X \rangle/d\varphi|^{-1} = N^{-2}$. Since $|\langle\mathcal{H}\rangle| = \langle\hat{n}^2\rangle = \frac{1}{2}N^2$, the Heisenberg limit is attained. This is a formal demonstration that the Heisenberg limit can be attained according to quantum mechanics, even though we currently do not know how to implement it.

In conclusion, we demonstrated that the Heisenberg limit is optimal for all parameter estimation procedures in quantum metrology. The correct resource to take into account is (the expectation value of) the generator of the translations in the parameter. In the case of most optical phase estimation protocols, this reduces to the average photon number. Contrary to the origin of the term ‘‘Heisenberg limit,’’ it is not a generalized uncertainty relation, but rather an expression of the Margolus-Levitin bound on the speed of dynamical evolution for quantum states.

We thank Jonathan Dowling for establishing the etymology of the term ‘‘Heisenberg limit,’’ and Sam Braunstein for valuable comments on the manuscript. This research was funded by the White Rose Foundation.

*p.kok@sheffield.ac.uk

- [1] C. M. Caves, *Phys. Rev. D* **23**, 1693 (1981).
- [2] P. Kok and B. W. Lovett, *Introduction to Optical Quantum Information Processing* (Cambridge University Press, Cambridge, England, 2010).
- [3] C. W. Helstrom, *Phys. Lett.* **25A**, 101 (1967).
- [4] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [5] S. L. Braunstein and C. M. Caves, *Phys. Rev. Lett.* **72**, 3439 (1994).
- [6] C. W. Gardiner and P. Zoller, *Quantum Noise* (Springer-Verlag, Berlin, 2004), 3rd ed., p. 322.
- [7] M. J. Holland and K. Burnett, *Phys. Rev. Lett.* **71**, 1355 (1993).
- [8] J. Beltrán and A. Luis, *Phys. Rev. A* **72**, 045801 (2005).
- [9] S. Boixo, S. T. Flammia, C. M. Caves, and J. M. Geremia, *Phys. Rev. Lett.* **98**, 090401 (2007).
- [10] S. Roy and S. L. Braunstein, *Phys. Rev. Lett.* **100**, 220501 (2008).
- [11] V. Giovannetti, S. Lloyd, and L. Maccone, *Phys. Rev. Lett.* **96**, 010401 (2006).
- [12] G. Breit and E. Wigner, *Phys. Rev.* **49**, 519 (1936).
- [13] J. Uffink, *Am. J. Phys.* **61**, 935 (1993).
- [14] A. Bhattacharyya, *Bull. Calcutta Math. Soc.* **35**, 99 (1943).
- [15] W. K. Wootters, *Phys. Rev. D* **23**, 357 (1981).
- [16] P. J. Jones and P. Kok, *Phys. Rev. A* **82**, 022107 (2010).
- [17] S. L. Braunstein, C. M. Caves, and G. J. Milburn, *Ann. Phys. (N.Y.)* **247**, 135 (1996).
- [18] W. Heitler, *The Quantum Theory of Radiation* (Clarendon Press, Oxford, 1954), 3rd ed., p. 65.
- [19] N. Margolus and L. B. Levitin, *Physica D (Amsterdam)* **120**, 188 (1998).
- [20] V. Giovannetti, S. Lloyd, and L. Maccone, *Phys. Rev. A* **67**, 052109 (2003).