



General Optimality of the Heisenberg Limit for Quantum Metrology

Author

Zwierz, Marcin, A. Pe´rez-Delgado, Carlos, Kok, Pieter

Published

2010

Journal Title

Physical Review Letters

DOI

[10.1103/PhysRevLett.105.180402](https://doi.org/10.1103/PhysRevLett.105.180402)

Downloaded from

<http://hdl.handle.net/10072/42768>

Griffith Research Online

<https://research-repository.griffith.edu.au>

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, UK*

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

(Dated: October 27, 2010)

It has long been known that quantum metrology promises improved sensitivity in parameter estimation over classical procedures. However, there is an extensive debate over the question how the sensitivity scales with the resources (such as the average photon number) and number of phase gates 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 how our proof resolves paradoxes that suggest sensitivities beyond the Heisenberg limit, and we clarify the precise relationship between the Heisenberg limit and Heisenberg's uncertainty principle.

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

Parameter estimation is a fundamental pillar of science and technology, and improved measurement techniques for parameter estimation have often led to scientific breakthroughs and technological advancement. The field of quantum metrology started with the work of Helstrom [1, 2], 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 generalisation of a known result in classical parameter estimation, called the Cramér-Rao bound. Braunstein and Caves [3] showed how this bound can be formulated for the most general state preparation and measurement procedures. While it is generally a hard problem to show that the Cramér-Rao bound can be attained in a given setup, at least it gives an upper limit to the precision of quantum parameter estimation. Caves [4] showed that quantum mechanical systems can in principle produce greater sensitivity over classical methods. A second, related problem is to find an expression for the mean square error in terms of the resources used in the parameter estimation procedure. In other words, is there a general optimal scaling of the quantum Cramér-Rao bound? We usually consider two scaling regimes: (i) the *standard quantum limit* (SQL) [5] or *shot-noise limit*, which is typically given by the inverse square root of the number of times T we make a measurement, and (ii) the *Heisenberg limit* [6], in which the mean error scales linearly inversely with the resource count. Often the standard quantum limit and the Heisenberg limit can be compared directly in terms of e.g., the average photon number. However, as we will see shortly, they refer to two fundamentally different quantities.

The question is now twofold: First, what is the appropriate resource count for the Cramér-Rao bound? And second, is the Heisenberg limit (i.e., linear scaling) also the ultimate limit of the quantum Cramér-Rao bound? For many common cases the first question is easily answered: when in an optical parameter estimation procedure each photon probes the system of interest once, the appropriate resource count is the photon number. In most cases the Heisenberg limit is then given by the inverse photon number N^{-1} , as expected. However,

there are cases where the error in a parameter can scale much more favourably with respect to the average photon number. For example, Boixo *et al.* [7] devised a parameter estimation procedure that sees the error scale with N^{-2} , and Roy and Braunstein [8] construct a procedure that achieves an error that scales with 2^{-N} . In the case of continuous variables, Beltrán and Luis [9] showed that the use of classical optical nonlinearities can lead to an error with average photon number scaling $N^{-3/2}$. It is tempting to conclude that these protocols beat the Heisenberg limit, but this is not the case. To see this, we first need to establish the exact nature of the resource count that is to be used. We will see that in some cases this is *not* the photon number or average energy used in the procedure.

This paper is organized as follows: First, we will argue that the proper resource count is 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. This will shed light on the precise relation between the Heisenberg limit and Heisenberg's uncertainty principle. Finally, we show how the proposal by Beltrán and Luis is reconciled with our proof.

The most general parameter estimation procedure is shown in Fig. 1a). 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 into 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 about φ that can be extracted from this measurement is

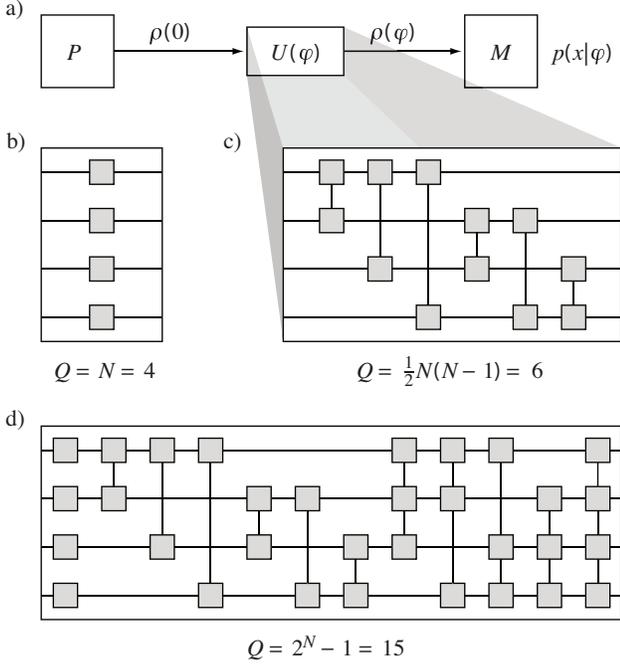


FIG. 1: a) General parameter estimation procedure involving state preparation P , evolution $U(\varphi)$ and generalized measurement M with outcomes x , which produces a probability distribution $p(x|\varphi)$. In terms of quantum networks, the evolution can be written as a number of queries of the parameter φ . b) Example for $N = 4$ of the usual situation described by \mathcal{H}_{GLM} , where each system performs a single query, and the number of queries equals the number of systems (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: any two systems can jointly perform a single query, and the number of queries then scales quadratically with 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.

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 [1, 3]

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

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 recognising that a

quantum parameter estimation protocol can be written as a quantum network acting on a set of quantum systems, with repeated couplings of the network to the system we wish to probe for the parameter φ . The couplings that introduce φ into the network can be considered *queries*, and 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.

As an example, consider the quantum networks presented in Fig. 1. The quantum network in b) was analysed by Giovannetti, Lloyd, and Maccone [10]. 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 $\mathcal{H}_{\text{GLM}} = \sum_j H_j$. The number of queries Q is then equal to the number of terms in \mathcal{H}_{GLM} , or $Q = N$. Next, the quantum network corresponding to the procedure of Boixo, Flammia, Caves, and Geremia is shown in Fig. 1c), where each query consists of a joint interaction on two modes. In other words, the queries are given by the unitary gates $O_{ij} = \exp(-i\varphi H_i \otimes H_j)$ where $i \neq j$. Since all H_i commute with each other, the generator of the joint queries can be written as

$$\mathcal{H}_{\text{BFCG}} = \sum_{i>j} H_i \otimes H_j. \quad (3)$$

The number of terms in $\mathcal{H}_{\text{BFCG}}$, and therefore the query complexity with respect to the number of systems, is given by $Q = \frac{1}{2}N(N-1) = O(N^2)$. Finally, the network corresponding to the protocol of Roy and Braunstein is given in Fig. 1d). It is easy to see that the number of terms in the corresponding generator \mathcal{H}_{RB} is given by $2^N - 1$, and the number of queries is therefore $Q = 2^N - 1$. Since we have a systematic method for increasing N (and Q) given the gate O_j , this defines an 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.

From our arguments about the query complexity of quantum networks, it is clear that the resource count must be related to the generator of the joint queries \mathcal{H} . The most natural way to map this operator to a number (i.e., the *count*) is by taking the expectation value $\langle \mathcal{H} \rangle$. However, we have to address an important subtlety. When \mathcal{H} corresponds to a proper Hamiltonian, the fact that the origin of the energy scale has no physical meaning means that the actual value of $\langle \mathcal{H} \rangle$ can be changed arbitrarily. Hence, we must fix the scale such that the ground state (which may be degenerate) has zero energy. In most cases, this is an intuitive choice. For example, most people would agree that 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 GHZ 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 $|\downarrow\rangle$.

For more general interactions $U(\varphi)$ where we include feed-forward and arbitrary unitary gates between queries in Fig. 1, we can use an argument by Giovannetti *et al.* [10] to show that $\langle \mathcal{H} \rangle = \langle i(\partial U(\varphi)/\partial \varphi)U^\dagger(\varphi) \rangle$ is unaffected by the intermediate unitary gates, and the scaling is therefore still determined by Q .

Finally, one may argue that the resource count should be defined in terms of the variance or semi-norm of \mathcal{H} . Indeed, this is how the resource count is performed to date. 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 [11, 12]. In these cases the resource count, and by implication the scaling of the error, would be ill-defined. By contrast, the expectation value of \mathcal{H} always exists and is always positive. Moreover, when the variance exists, both expressions can be bounded by a linear function in the number of queries Q . Let $\mathcal{H} = \sum_j^Q A_j$. The variance can then be written as

$$\begin{aligned} (\Delta \mathcal{H})^2 &= \left\langle \left(\sum_j^Q A_j \right)^2 \right\rangle - \left\langle \sum_j^Q A_j \right\rangle^2 \\ &= \sum_j^Q \langle L_j \rangle - \sum_{j,k}^Q \langle A_j \rangle \langle A_k \rangle \leq cQ^2 \end{aligned} \quad (4)$$

for some positive number c and positive operator L_j , which gives $\Delta \mathcal{H} \leq O(Q)$ [7]. 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 behaviour with Q , we can use either the variance or the expectation value.

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.

After establishing the appropriate resource count, we are finally in a position to prove the optimality of the Heisenberg limit for quantum parameter estimation 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 $p(x)$ and $p(x) + dp(x)$. The infinitesimal statistical distance between these distributions is given by [13, 14]

$$ds^2 = \int dx \frac{1}{p(x)} [dp(x)]^2. \quad (5)$$

Dividing both sides by $(d\varphi)^2$ and including the dependence on φ in $p(x)$, we obtain

$$\left(\frac{ds}{d\varphi} \right)^2 = \int dx \frac{1}{p(x|\varphi)} \left(\frac{\partial p(x|\varphi)}{\partial \varphi} \right)^2 = F(\varphi), \quad (6)$$

which relates the Fisher information to the rate of change of the statistical distance (i.e., the speed of dynamical evolution).

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 [14] between quantum states as the statistical distance:

$$s(\psi, \phi) = \arccos(|\langle \psi | \phi \rangle|), \quad (7)$$

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

$$|\psi(\varphi)\rangle = \exp(-i\varphi \mathcal{H}) |\psi(0)\rangle. \quad (8)$$

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. (7) and using the Schrödinger equation implicit in Eq. (8) [15]:

$$\frac{ds}{d\varphi} \leq \langle \mathcal{H} \rangle. \quad (9)$$

Combining this with Eq. (6) and Eq. (2) leads to the Cramér-Rao bound

$$(\delta\varphi)^2 \geq \frac{1}{T} \left(\frac{ds}{d\varphi} \right)^{-2} = \frac{1}{T \langle \mathcal{H} \rangle^2}. \quad (10)$$

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

$$\delta\varphi \geq \frac{1}{\langle \mathcal{H} \rangle}. \quad (11)$$

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, and 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 [10]. This completes the proof of the optimality of the Heisenberg limit in the most general case.

In addition to Eqs. (6) and (9), the Fisher information is also bounded by the variance of \mathcal{H} [16]:

$$F(\varphi) \leq 4(\Delta \mathcal{H})^2. \quad (12)$$

This leads to a (single-shot) quantum Cramér-Rao bound

$$\delta\varphi \geq \frac{1}{2\Delta \mathcal{H}}. \quad (13)$$

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 Principle* for the parameter φ and its conjugate operator \mathcal{H} . Any parameter estimation procedure must respect both bounds, and the Heisenberg limit in Eq. (11) may not be attained, even in principle, because the bound in Eq. (13) prevents it from doing so.

The term ‘‘Heisenberg limit’’ was introduced by Holland and Burnett [6], who referred to the number-phase uncertainty relation in Heitler [17]. 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 [15, 18]. To see this, we can formally solve Eq. (9) by separation of variables, yielding

$$\int_0^\varphi d\varphi' \geq \frac{1}{\langle \mathcal{H} \rangle} \int_0^{\pi/2} ds \Rightarrow \varphi \geq \frac{\pi}{2} \frac{1}{\langle \mathcal{H} \rangle}. \quad (14)$$

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 principle, on the other hand, can be identified with the Mandelstam-Tamm bound [15]. 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 [9]. 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 not only does *not* beat the Heisenberg limit, it does not attain it!

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

$$\delta\varphi = \frac{\Delta X}{|d\langle X \rangle/d\varphi|} = \frac{1}{N^2}. \quad (15)$$

Since $\langle \mathcal{H} \rangle = \langle \hat{n}^2 \rangle = O(N^2)$, this procedure exactly attains the Heisenberg limit. Note that 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, but it requires careful consideration as to which resource is appropriate for expressing the scaling behaviour of the mean square error. 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 generalised uncertainty relation, but rather an expression of the Margolus-Levitin bound on the speed of dynamical evolution for quantum states.

Acknowledgements. 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.

* Electronic address: p.kok@sheffield.ac.uk

- [1] C. W. Helstrom, Phys. Letters **25A**, 101 (1967).
- [2] C. W. Helstrom, *Quantum detection and estimation theory* (Academic Press, 1976).
- [3] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
- [4] C. M. Caves, Phys. Rev. D **23**, 1693 (1981).
- [5] C. W. Gardiner and P. Zoller, *Quantum Noise*, p. 322 (Springer-Verlag, 2004), 3rd ed.
- [6] M. J. Holland and K. Burnett, Phys. Rev. Lett. **71**, 1355 (1993).
- [7] S. Boixo, S. T. Flammia, C. M. Caves, and J. M. Geremia, Phys. Rev. Lett. **98**, 090401 (2007).
- [8] S. Roy and S. L. Braunstein, Phys. Rev. Lett. **100**, 220501 (2008).
- [9] J. Beltrán and A. Luis, Phys. Rev. A **72**, 045801 (2005).
- [10] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 10401 (2006).
- [11] G. Breit and E. Wigner, Phys. Rev. **49**, 519 (1936).
- [12] J. Uffink, Am. J. Phys. **61**, 935 (1993).
- [13] A. Bhattacharyya, Bulletin of the Calcutta Mathematical Society **35**, 99 (1943).
- [14] W. K. Wootters, Phys. Rev. D **23**, 357 (1981).
- [15] P. J. Jones and P. Kok, arXiv:1003:4870 (2010).
- [16] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Ann. Phys. **247**, 135 (1996).
- [17] W. Heitler, *The Quantum Theory of Radiation*, p. 65 (Clarendon Press, Oxford, 1954), 3rd ed.
- [18] N. Margolus and L. B. Levitin, Physica D **120**, 188 (1998).