Using Entanglement Against Noise in Quantum Metrology

Rafal Demkowicz-Dobrzański[1] and Lorenzo Maccone[2]
[1]Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, Poland
[2]Dipartimento di Fisica and INFN Sezione Pavia, University of Pavia, via Bassi 6, I-27100 Pavia, Italy
(Received 21 July 2014; published 19 December 2014)

We analyze the role of entanglement among probes and with external ancillas in quantum metrology. In the absence of noise, it is known that unentangled sequential strategies can achieve the same Heisenberg scaling of entangled strategies and that external ancillas are useless. This changes in the presence of noise; here we prove that entangled strategies can have higher precision than unentangled ones and that the addition of passive external ancillas can also increase the precision. We analyze some specific noise models and use the results to conjecture a general hierarchy for quantum metrology strategies in the presence of noise.

DOI: 10.1103/PhysRevLett.113.250801 PACS numbers: 06.20.-f, 03.65.Ta, 03.67.Ac, 42.50.Lc

Quantum metrology [1,2] describes parameter estimation techniques that, by sampling a system N times, achieve precision better than the $1/\sqrt{N}$ scaling of the central limit theorem of classical strategies. Different schemes can beat such a limit (Fig. 1): (i) entanglement-free classical schemes where N/n independent probes sense the system sequentially, thus rescaling the parameter—and hence the error—by n for each probe [3,4]; (ii) entangled parallel schemes that employ a collective entangled state of the N probes that sample the system in parallel [5–8]; (iii) passive ancilla schemes, where the N probes may also be entangled with noiseless ancillas; (iv) active ancilla-assisted schemes (comprising all of the previous cases) that also encompass all schemes employing feedback, where adaptive procedures are described as unitary operations acting on the probes and ancillas between the sensing and the final measurement [9,10].

In the noiseless case, classical single-probe sequential schemes (i) can attain the same $1/N$ precision as parallel entangled ones (ii) at the expense of an N-times longer sampling time, whereas passive and active ancilla schemes (iii) and (iv) offer no additional advantage [9,11]. In this Letter we analyze the performance of these strategies in the presence of specific noise models and use the results to conjecture a general hierarchy of protocols. Noise in quantum metrology has been extensively studied—e.g., see Refs. [12–25]—but the main focus was on comparing parallel-entangled with parallel-unentangled strategies [12–14] which do not match in the noiseless case. Single-probe states are typically less sensitive to decoherence and simpler to prepare than entangled states, so it would seem [3] that the sequential strategy should be preferable in the presence of noise. Our first result is that this is not true: in the presence of noise (here we analyze dephasing, erasure, and damping), entanglement among probes increases the precision over the sequential strategy, even though it fails to do so in the noiseless case, and we provide a quantitative characterization of this advantage. Our second result is to show that (ii) and (iii) are, in general, asymptotically inequivalent by demonstrating that (iii) is strictly better than (ii) for amplitude-damping noise. Our third result is to show that the bounds to parallel-entangled strategies (ii) and (iii) derived for a large class of noise models [13,14] also apply asymptotically in N to the most general strategies (iv), suggesting that active ancilla-based schemes are not helpful in increasing the precision in the presence of noise [26]. Finally, we use our results to conjecture a general hierarchy of quantum metrology schemes valid in the presence of any uncorrelated noise:

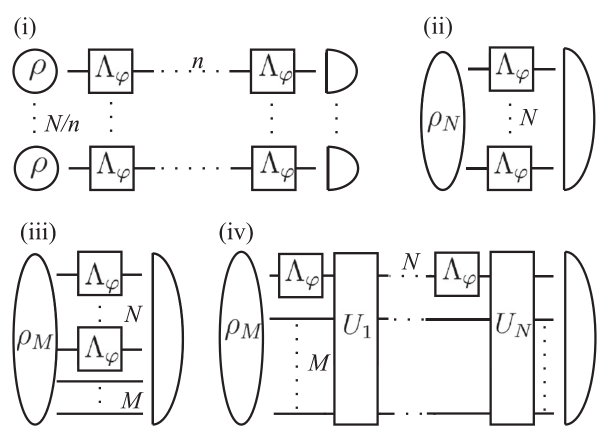

FIG. 1. Quantum metrology strategies. The maps Λ_φ encode the parameter φ to be estimated. (i) sequential scheme: Λ_φ acts n times sequentially on N/n input probes ρ (this is an entanglement-free classical scheme); (ii) entangled parallel scheme: an entangled state of N probes ρ_N goes through N maps Λ_φ in parallel; (iii) passive ancilla scheme: the N probes are also entangled with M noiseless ancillas; (iv) active ancilla-assisted scheme: the action of N channels Λ_φ is interspersed with arbitrary unitaries U_i representing interactions of the probe with ancillas. [All of the other schemes can be derived from (iv) choosing swap or identity unitaries U_i.]

$$\begin{aligned}
&(\mathrm{i}) = (\mathrm{ii}) = (\mathrm{iii}) = (\mathrm{iv}) \quad \text{decoherence free},\\
&(\mathrm{i}) < (\mathrm{ii}) = (\mathrm{iii}) = (\mathrm{iv}) \quad \text{dephasing, erasure},\\
&(\mathrm{i}) < (\mathrm{ii}) < (\mathrm{iii}) \stackrel{?}{=} (\mathrm{iv}) \quad \text{amplitude-damping},\\
&(\mathrm{i}) \stackrel{?}{\leq} (\mathrm{ii}) \leq (\mathrm{iii}) \stackrel{?}{=} (\mathrm{iv}) \quad \text{general conjecture}.
\end{aligned} \tag{1}$$

Namely, in general, sequential strategies (i) are worse [27] than parallel-entangled ones (ii), which might in some cases be improved by entangling the probes with noiseless ancillas (iii), but there is no additional asymptotic gain from using active ancilla-aided schemes (iv). Question marks represent our conjectures and the equality symbol "=" should be interpreted as asymptotically equivalent, though in the decoherence-free case as well as in the case of equality between (ii) and (iii) for erasure and dephasing noise, this is a strict equality for any finite N. We emphasize that even in the cases where our bounds are equivalent, the related strategies may not be if the bounds are not achievable, which must be verified on a case-by-case basis.

Schemes that employ quantum-error correction [28–30] are in general of type (iv), so our claim might be misinterpreted as saying that error correction schemes are useless. Instead, what we will say is simply that their asymptotic precision can also be achieved through (possibly unknown) strategies of types (ii) and (iii); e.g., the noise models considered in [28–30] allow for decoupling the decoherence from the parameter sensing transformation at short evolution times; so, the bounds derived for (ii) and (iii) also allow for the possibility of better than $1/\sqrt{N}$ scaling [31].

The outline of the Letter is as follows. We first introduce the quantum Cramer-Rao bound for the strategies (i)–(iv), and derive some general bounds for their quantum Fisher information. We then prove a gap in precision between (i) and (ii), the equivalence of (ii), (iii), and (iv) in case of dephasing and erasure noise, and finally the inequivalence of (ii) and (iii) for amplitude damping.

The map Λ_φ that writes the parameter φ on the state ρ of the probe acts as

$$\rho_\varphi = \Lambda_\varphi(\rho) = \sum_k K_k^\varphi \rho K_k^{\varphi\dagger}, \tag{2}$$

with K_k^φ being the Kraus operators. The precision of an estimation strategy can be gauged through the root-mean-square error $\Delta\varphi$ of the measurement of φ. It is lower bounded by the quantum Cramer-Rao bound [2,5–8], $\Delta\varphi \geq 1/\sqrt{\nu F(\rho_\varphi)}$, where ν is the number of times the estimation is repeated and $F(\rho)$ is the quantum Fisher information (QFI) of a state ρ [2,5,6]. This bound is guaranteed to be achievable, in general, only asymptotically for $\nu \to \infty$, but in case of noise models with QFI scaling linearly with the number of probe particles N it is also tight for a single shot setting, $\nu = 1$, provided one considers the asymptotics, $N \to \infty$ [32].

The QFIs for the schemes (i)–(iv) are defined as

$$F^{(\mathrm{i})} = \max_{\rho,n} F\{[\Lambda_\varphi^n(\rho)]^{\otimes N/n}\}, \tag{3}$$

$$F^{(\mathrm{ii})} = \max_{\rho_N} F[\Lambda_\varphi^{\otimes N}(\rho_N)], \tag{4}$$

$$F^{(\mathrm{iii})} = \max_{\rho_M} F[\Lambda_\varphi^{\otimes N} \otimes \mathbb{1}^{\otimes M}(\rho_M)], \tag{5}$$

$$F^{(\mathrm{iv})} = \max_{\rho_M,\{U_i\}} F[U_N \Lambda_\varphi \ldots U_1 \Lambda_\varphi(\rho_M)], \tag{6}$$

where ρ denotes an input state of a single probe and we look for the optimal sequential-parallel splitting of the N probes in n channels for strategies (i), ρ_N is the global state of N probes in (ii), and ρ_M denotes the global probes-ancilla input state in (iii) and (iv). In the formula for $F^{(\mathrm{iv})}$, the U_i's act on all the probes while Λ_φ without loss of generality may be assumed to act on the first probe only. Because of the convexity of the QFI, the optimal input probes are pure.

The hierarchy conjecture (1) should be understood in terms of corresponding inequalities on QFIs: $F^{(\mathrm{ii})} \leq F^{(\mathrm{iii})}$ is obvious as (ii) is a special case of (iii)—the inequality may be strict, as is the case of the amplitude damping discussed below; $F^{(\mathrm{iii})} \leq F^{(\mathrm{iv})}$ is also easy to show since taking swap operators U_i in (iv), one can obtain the action of parallel channels on an entangled input state (iii). It is less trivial to determine the cases when inequalities turn to equalities and the corresponding schemes become asymptotically equivalent. Finally, the $F^{(\mathrm{i})} \leq F^{(\mathrm{ii})}$ inequality is more challenging to prove, in general, but we show that it holds strictly for dephasing, erasure, or amplitude damping, proving the advantage of parallel schemes [33]. We also present general tools to derive bounds for (iv) and show that they are asymptotically equivalent to known bounds for (ii) and (iii). Moreover, since these bounds are saturable for dephasing and erasure using (ii) schemes, there is no asymptotic advantage of (iv) over the simpler (ii) and (iii) in these cases.

Calculating QFI explicitly for large N is, in general, not possible, but bounds to it are known. The most versatile ones employ the nonuniqueness of the Kraus representation [13,14,34]; Λ_φ is unchanged if one replaces K_k^φ with $\tilde{K}_k^\varphi = \sum_l u_{kl}^\varphi K_l^\varphi$, where u^φ is an arbitrary φ-dependent unitary matrix. This produces bounds on the maximal QFI of a transformation Λ_φ in terms of minimization over the possible Kraus representations [14,34]:

$$\max_\rho F[\Lambda_\varphi(\rho)] \leq 4 \min_{\{K_k^\varphi\}} \Big\| \sum_k \dot{K}_k^{\varphi\dagger} \dot{K}_k^\varphi \Big\|, \tag{7}$$

where $\dot{K}_k^\varphi = (\partial K_k^\varphi / \partial\varphi)$ and $\|\cdot\|$ is the operator norm. The above inequality becomes an equality provided one replaces Λ_φ with a trivially extended channel $\Lambda_\varphi \otimes \mathbb{1}$, which represents the possibility of entangling the probes with an ancilla [34]. This immediately implies that the bounds derived for (ii) will also be valid for (iii).

We now recall the known bounds for $F^{(\mathrm{ii/iii})}$ and derive a new bound for $F^{(\mathrm{iv})}$ using the minimization of Eq. (7). Bounds for (ii) and (iii) are equivalent (as argued above), so we use a combined notation (ii/iii). For any Kraus representation K_k^φ of a single channel Λ_φ, one can write a *product* Kraus representation for channels $\Lambda_\varphi^{\otimes N}$, $U_1\Lambda_\varphi \ldots U_N\Lambda_\varphi$ corresponding to schemes (ii/iii) and (iv), respectively: $K_{\boldsymbol{k}}^{\varphi(\mathrm{ii/iii})} = K_{k_N}^\varphi \otimes \cdots K_{k_1}^\varphi$; $K_{\boldsymbol{k}}^{\varphi(\mathrm{iv})} = U_N K_{k_N}^\varphi \ldots U_1 K_{k_1}^\varphi$, where $\boldsymbol{k} = \{k_1, \ldots, k_N\}$.

For (ii/iii) the minimization (7) gives a simple bound expressed in terms of single channel Kraus operators [34]:

$$F^{(\mathrm{ii/iii})} \leq 4\min_{K_k^\varphi} N\|\alpha\| + N(N-1)\|\beta\|^2 \leq 4 \min_{K_k^\varphi, \beta=0} N\|\alpha\|, \tag{8}$$

with $\alpha \equiv \sum_k \dot{K}_k^{\varphi\dagger}\dot{K}_k^\varphi$ and $\beta \equiv \sum_k \dot{K}_k^{\varphi\dagger} K_k^\varphi$. The last inequality in (8) may be used without a loss of efficiency for large N provided there is a Kraus representation for which $\beta = 0$ (it exists for many noisy maps), which immediately implies linear QFI scaling with N [14,34]. The minimization in Eq. (8) can be easily performed using the semidefinite programming [14,35].

The derivation of the general bound for (iv) again uses (7) and a product Kraus representation. It gives (see the Supplemental Material [36] for details)

$$\begin{aligned} F^{(\mathrm{iv})} &\leq 4\min_{K_k^\varphi} N\|\alpha\| + N(N-1)\|\beta\|(\|\alpha\| + \|\beta\| + 1) \\ &\leq 4 \min_{K_k^\varphi, \beta=0} N\|\alpha\|. \end{aligned} \tag{9}$$

Importantly, the asymptotic form of the bound is equivalent to Eq. (8), the one derived for (ii/iii) if $\beta = 0$ is feasible.

It is worth noting that less powerful but more intuitive methods based on the concept of minimization over classical or quantum simulations of the channel [14,35,39], originally proposed to derive bounds for (ii/iii), can also be applied to (iv). In the classical-and-quantum-simulation method [39] one formally replaces the action of Λ_φ with a parameter-independent map Λ and a parameter-dependent ancillary system σ_φ, so that for any ρ, $\Lambda_\varphi(\rho) = \Lambda(\rho \otimes \sigma_\varphi)$. Since QFI is nonincreasing under parameter-independent maps, $F(\Lambda(\rho \otimes \sigma_\varphi)) \leq F(\sigma_\varphi)$, which for the schemes (ii/iii) implies that $F[\Lambda_\varphi^{\otimes N}(\rho_N)] \leq F[\Lambda^{\otimes N}(\rho_N \otimes \sigma_\varphi^{\otimes N})] \leq NF(\sigma_\varphi)$ [14,35,39]. It was not noticed before that the same method can be applied to (iv), as the scheme can be rewritten as one black-box quantum operation $\tilde{\Lambda}$ fed with $\sigma_\varphi^{\otimes N}$ (see Fig. 2), resulting in

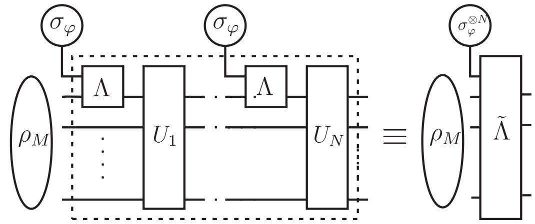

FIG. 2. Depiction of quantum channel simulation applied to the most general adaptive scheme (iv). It shows that, for a given simulation Λ, σ_φ, the QFI of the scheme is bounded by $F^{(\mathrm{iv})} = F[\Lambda(\rho_M \otimes \sigma_\varphi^{\otimes N})] \leq NF(\sigma_\varphi)$.

$$F^{(\text{ii-iv})} \leq N\min_{\Lambda,\sigma_\varphi} F(\sigma_\varphi). \tag{10}$$

This bound often coincides with the asymptotic bound in Eq. (8), e.g., in the case of erasure or dephasing [14,35] (but not amplitude damping; see Ref. [35]).

We now analyze dephasing, erasure, and amplitude-damping noise. Let $|0\rangle$, $|1\rangle$ be the eigenbasis of the phase encoding unitary $U_\varphi = |0\rangle\langle 0| + e^{i\varphi}|1\rangle\langle 1|$. We assume that the dephasing is defined with respect to the same basis, so the corresponding Kraus operators read

$$K_0 = \mathbb{1}\left(\frac{1+\sqrt{\eta}}{2}\right)^{1/2}, \qquad K_1 = \sigma_z\left(\frac{1-\sqrt{\eta}}{2}\right)^{1/2}, \tag{11}$$

where $\mathbb{1} = |0\rangle\langle 0| + |1\rangle\langle 1|$, $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$, and $\sqrt{\eta}$ is the decoherence rate of the off-diagonal terms in the density matrix. Since both Kraus operators commute with the unitary U_φ, we can separate the noise map from the sampling and consider a total evolution of the form

$$\rho_\varphi = \Lambda_\varphi(\rho) = \sum_k K_k U_\varphi \rho U_\varphi{}^\dagger K_k{}^\dagger. \tag{12}$$

Instead, for erasure noise the probe is untouched with probability η while with probability $1-\eta$ its state is replaced with one in a subspace orthogonal to the subspace where the estimation takes place (again the noise map and U_φ commute and the map can be written in a Kraus form; see the Supplemental Material [36]). The erasure map is isomorphic to optical loss applied to a state with a fixed number of distinguishable photons with transmission coefficient η in both arms of an interferometer [14,40]. Finally, Kraus operators for amplitude damping read

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{\eta} \end{pmatrix}, \qquad K_1 = \begin{pmatrix} 0 & \sqrt{1-\eta} \\ 0 & 0 \end{pmatrix}, \tag{13}$$

where η represents the probability of a particle switching from the excited to the ground state.

We start with calculating $F^{(\mathrm{i})}$ to assess the performance of entanglement-free strategies. In the case of erasure, since in the noiseless case the optimal probe state is

$|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, while the probability of an erasure event does not depend on the state itself, the optimal input state remains the same and yields $F[\Lambda_\varphi(|+\rangle\langle+|)] = \eta$. For dephasing and amplitude damping, the situation is less obvious but the optimal probe state is again $|+\rangle$ and the QFI is again η [35]; see the Supplemental Material [36] for a simple proof in the case of dephasing.

To calculate $F^{(\mathrm{i})}$ it remains to optimize the number n of sequential maps for each probe; see Fig. 1. Using n maps in a sequence increases the overall phase rotation n times at the cost of increasing the decoherence parameter η to η^n, whereas considering parallel channels simply adds their QFIs. Therefore, $F\{[\Lambda_\varphi^n(\rho_+)]^{\otimes N/n}\} = N/n \cdot n^2\eta^n$. This is the same formula which would be obtained for (ii) with the input $N00N$ state [17,41–43]. Treating $1 \le n \le N$ as a continuous parameter [17], the optimal value $n = [\ln(1/\eta)]^{-1}$, provided $e^{-1} \le \eta \le e^{-1/N}$, which corresponds to [44]

$$F^{(\mathrm{i})} = \frac{N}{e\ln(1/\eta)}. \tag{14}$$

For erasure and dephasing, we use the inequality (8) to calculate (see Supplemental Material [36])

$$F_Q^{(\mathrm{ii/iii})} \lesssim \frac{N\eta}{1-\eta}. \tag{15}$$

Importantly, this bound is asymptotically saturable for both models with a scheme (ii) where the optimal input probes are prepared in spin-squeezed states for atomic systems [13,15], or in squeezed states of light for optical implementations [14,40,45].

In order to inspect the benefits of entangled-based strategies over sequential ones, we plot in Fig. 3 the ratio

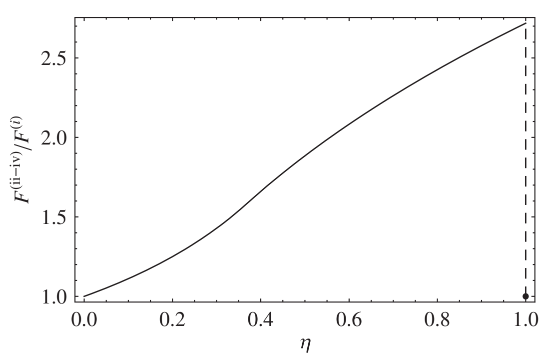

FIG. 3. Advantage of entangled-based over entanglement-free schemes for erasure, dephasing, and amplitude damping, quantified as an asymptotic ratio of achievable quantum QFIs as a function of the decoherence parameter η. For $\eta \to 1$ the ratio approaches exp(1), but for the perfectly noiseless case $\eta = 1$, the advantage vanishes, which is depicted by a dot. In the case of amplitude damping, a further improvement is possible (bounded by a factor of 4) when using strategies (iii) and (iv) instead of (ii).

of formulas in Eqs. (15) and (14) as a function of η. Note that the entanglement-enhancement factor is bounded by bounded by exp(1), a result known in frequency estimation schemes in the limit vanishing interrogation times [12,13], which in our case corresponds to $\eta \to 1$. We stress, however, that in the noiseless case $\eta = 1$, all four metrology schemes perform equally well, achieving the Heisenberg scaling. Finally, regarding scheme (iv), we note that since the asymptotic bound on $F_Q^{(\mathrm{iv})}$ coincides with the bound on $F_Q^{(\mathrm{ii/iii})}$ (as $\beta = 0$) and the latter is asymptotically saturable using (ii) for erasure and dephasing, this immediately implies that there is no asymptotic benefit in using (iv) in these cases.

One can also derive the corresponding bound for the amplitude damping (see the Supplemental Material [36]) which reads $F_Q^{(\mathrm{ii/iii})} \lesssim 4N\eta/(1-\eta)$. This bound, however, is not tight for the (ii) strategies, which has been proven recently in Ref. [46] using an alternative method based on the calculus of variations—the actual tight bound for (ii) in fact coincides with Eq. (15). This makes the case of amplitude damping distinct from the other two and opens up the possibility of proving the asymptotic benefits of using the ancillas; see below.

Analyzing the role of ancillas, we have already shown that they are useless in the case of dephasing and erasure. Surprisingly, this not so in the case of the amplitude-damping noise. In this case, as mentioned above and proven in Ref. [46], bound (15) is tight for (ii). A numerical search for optimal ancilla assisted strategies (iii) for the small

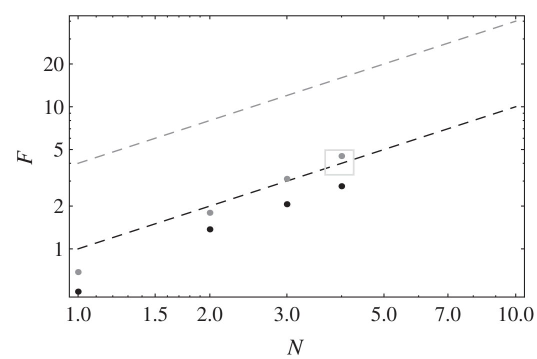

FIG. 4. Comparison between the yield of the amplitude-damping channel with and without passive ancillas for exemplary decoherence parameter $\eta = 0.5$ as a function of the number N of maps employed in the estimation: attainable QFI without ancillas, strategy (ii) (black circles); attainable QFI with passive ancillas, strategy (iii) (gray circles); asymptotically tight upper bound for the QFI for (ii) strategies from Ref. [46] (dashed black line); our universal bound for QFI for both passive (iii) and active (iv) ancillas (dashed gray curve)—no strategy can achieve better precision. The gray box emphasizes that for $N = 4$, and hence also asymptotically, the strategy (iii) can beat the bound for all strategies of type (ii). (More details are in the Supplemental Material [36].)

number of probes where $N \leq 4$ gives a QFI that exceeds the bound (15) for $\eta \lesssim 0.5$; see Fig. 4. Most importantly, this advantage of strategy (iii) over (ii) will also be preserved in the asymptotic limit since the bound (15) is linear in N and the same linear gain can be achieved by simply repeating the experiment, e.g., using the optimal 4 particle strategy $N/4$ times. This gives a (numerical) proof that (ii) is strictly less powerful than (iii) for amplitude damping.

In conclusion, we have presented a hierarchy for the performance of quantum metrology in the presence of dephasing, erasure, and amplitude-damping noise, and we have illustrated a conjecture on how this hierarchy can be extended to arbitrary noise models, based on new general bounds. In this hierarchy, entanglement-free schemes perform worse than entangled ones, and in some cases schemes with passive ancillas perform better than unaided ones, even though they are all equivalent in the noiseless case.

R. D. D. thanks Marcin Jarzyna for the useful discussions. This research work was supported by the FP7 IP project SIQS, cofinanced by the Polish Ministry of Science and Higher Education.

[1] V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004).

[2] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Photonics **5**, 222 (2011).

[3] A. Luis, Phys. Rev. A **65**, 025802 (2002).

[4] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Nature (London) **450**, 393 (2007).

[5] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Ann. Phys. (N.Y.) **247**, 135 (1996).

[6] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[7] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).

[8] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).

[9] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).

[10] W. van Dam, G. M. D'Ariano, A. Ekert, C. Macchiavello, and M. Mosca, Phys. Rev. Lett. **98**, 090501 (2007).

[11] A. M. Childs, J. Preskil, and J. Renes, J. Mod. Opt. **47**, 155 (2000).

[12] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. Ekert, M. Plenio, and J. Cirac, Phys. Rev. Lett. **79**, 3865 (1997).

[13] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, Nat. Phys. **7**, 406 (2011).

[14] R. Demkowicz-Dobrzański, J. Kolodynski, and M. Guta, Nat. Commun. **3**, 1063 (2012).

[15] D. Ulam-Orgikh and M. Kitagawa, Phys. Rev. A **64**, 052106 (2001).

[16] A. Shaji and C. M. Caves, Phys. Rev. A **76**, 032111 (2007).

[17] U. Dorner, R. Demkowicz-Dobrzanski, B. Smith, J. Lundeen, W. Wasilewski, K. Banaszek, and I. Walmsley, Phys. Rev. Lett. **102**, 040403 (2009),

[18] J. Kolodynski and R. Demkowicz-Dobrzanski, Phys. Rev. A **82**, 053804 (2010).

[19] S. Knysh, V. N. Smelyanskiy, and G. A. Durkin, Phys. Rev. A **83**, 021804 (2011).

[20] C. M. Caves, Phys. Rev. D **23**, 1693 (1981).

[21] M. Xiao, L.-A. Wu, and H. J. Kimble, Phys. Rev. Lett. **59**, 278 (1987).

[22] M. A. Rubin and S. Kaushik, Phys. Rev. A **75**, 053805 (2007).

[23] R. Demkowicz-Dobrzański, Laser Phys. **20**, 1197 (2010).

[24] S. Alipour, M. Mehboudi, and A. T. Rezakhani, Phys. Rev. Lett. **112**, 120405 (2014).

[25] S. Alipour and A. T. Rezakhani, arXiv:1403.8033.

[26] Our claim is stronger than the one in Ref. [13] that if (ii) and (iii) are limited by a $1/\sqrt{N}$ precision scaling, so is (iv). We derive an explicit bound for (iv) and show that it asymptotically coincides with the bound for (ii) and (iii).

[27] They are equivalent in the noiseless case, however.

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

[29] W. Dur, M. Skotiniotis, F. Fröwis, and B. Kraus, Phys. Rev. Lett. **112**, 080801 (2014).

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

[31] R. Chaves, J. B. Brask, M. Markiewicz, J. Kolodynski, and A. Acin, Phys. Rev. Lett. **111**, 120401 (2013).

[32] M. Jarzyna and R. Demkowicz-Dobrzański, arXiv:1407.4805 [New J. Phys. (to be published)].

[33] This claim may seem to contradict the one in Refs. [41,42,47], where equivalence of the sequential and parallel-entangled strategy is proven for dephasing. The contradiction is only an apparent one, as no optimization over the input state is performed there: only the response of the channel is analyzed.

[34] A. Fujiwara and H. Imai, J. Phys. A **41**, 255304 (2008).

[35] J. Kolodynski and R. Demkowicz-Dobrzański, New J. Phys. **15**, 073043 (2013).

[36] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.113.250801, which includes Refs. [37,38], for further details on the results presented here.

[37] K. Macieszczak, M. Fraas, and R. Demkowicz-Dobrzański, New J. Phys. **16**, 113002 (2014).

[38] K. Macieszczak, arXiv:1312.1356.

[39] K. Matsumoto, arXiv:1006.0300.

[40] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kolodynski, arXiv:1405.7703 [Prog. Opt. (to be published)].

[41] A. Shaji and C. M. Caves, Phys. Rev. A **76**, 032111 (2007).

[42] S. Boixo, C. M. Caves, A. Datta, and A. Shaji, Laser Phys. **16**, 1525 (2006).

[43] L. Maccone, Phys. Rev. A **88**, 042109 (2013).

[44] For completeness we should add that if $\eta < e^{-1}$ we take $n = 1$, which gives $F = N\eta$, while for $\eta > e^{-1/N}$ we take $n = N$, which gives $F = N^2\eta^N$. Note that in the asymptotic limit $N \to \infty$ and $\eta < 1$, we can ignore the case $\eta > e^{-1/N}$.

[45] R. Demkowicz-Dobrzański, K. Banaszek, and R. Schnabel, Phys. Rev. A **88**, 041802(R) (2013).

[46] S. I. Knysh, E. H. Chen, and G. A. Durkin, arXiv:1402.0495.

[47] S. Boixo and C. Heunen, Phys. Rev. Lett. **108**, 120402 (2012).