PAPER • OPEN ACCESS

Probabilistic metrology or how some measurement outcomes render ultra-precise estimates

To cite this article: J Calsamiglia *et al* 2016 *New J. Phys.* **18** 103049

View the article online for updates and enhancements.

Related content

- Flexible resources for quantum metrology
  Nicolai Friis, Davide Orsucci, Michalis Skotiniotis et al.

- Optimal quantum states for frequency estimation
  F Fröwis, M Skotiniotis, B Kraus et al.

- Units of rotational information
  Yuxiang Yang, Giulio Chiribella and Qinheping Hu

Recent citations

- Weak-value amplification and optimal parameter estimation in the presence of correlated noise
  Josiah Sinclair *et al*

New Journal of Physics

The open access journal at the forefront of physics

CrossMark

OPEN ACCESS

PAPER

Probabilistic metrology or how some measurement outcomes render ultra-precise estimates

J Calsamiglia[1], B Gendra, R Muñoz-Tapia and E Bagan

Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, E-08193 Bellaterra (Barcelona), Spain
[1]   Author to whom any correspondence should be addressed.

E-mail: John.Calsamiglia@uab.cat

Abstract

We show on theoretical grounds that, even in the presence of noise, probabilistic measurement strategies (which have a certain probability of failure or abstention) can provide, upon a heralded successful outcome, estimates with a precision that exceeds the deterministic bounds for the average precision. This establishes a new ultimate bound on the phase estimation precision of particular measurement outcomes (or sequence of outcomes). For probe systems subject to local dephasing, we quantify such precision limit as a function of the probability of failure that can be tolerated. Our results show that the possibility of abstaining can set back the detrimental effects of noise.

1. Introduction

Quantum-enhanced precision measurements and sensors are some of the most disruptive quantum technologies [1], with applications across various disciplines, e.g., optical communications [2, 3], cryptography [4], brain and heart medical diagnosis via atomic magnetometry [5, 6], biological measurements [7, 8], and are critical in gravitational-wave detectors [9, 10] and GPS and other current technologies that rely on atomic clocks [11–14].

The above are examples of the so-called quantum metrology problems. In broad terms a metrology problem can be cast as a four step process: the preparation of a probe, its controlled evolution that imprints in the probe the (continuous) unknown parameter to be estimated, a measurement on the modified probe, and a final data-processing to obtain the value of the unknown parameter. The accuracy of the estimation is limited by the experimental imperfections and, ultimately, by the noise inherent in any quantum measurement. Classically, one can reduce the effects of noise in a given setup by repeating the experiment on n independent preparations of the probe [15]. The uncertainty of the estimation is thereby reduced by a factor $n^{-1/2}$ (the so-called standard quantum limit, SQL). However, in a fully quantum mechanical setting, the possibility of using entangling operations in the preparation and in the measurement steps gives rise to an uncertainty that scales as n^{-1} (the so-called Heisenberg limit or scaling).

Recent experimental advances that enable an unprecedented control of diverse optical and condensed matter systems at a quantum level make quantum metrology an extremely timely field of research [16]. In the last years, the agenda of quantum-enhanced metrology has been put under scrutiny by a number of results [12, 17–21] that show that under quite generic (local, uncorrelated and markovian) experimental noise the quantum enhancement amounts to a constant factor rather than a quadratic improvement. The field has revamped in search for alternative schemes that push forward the limits and circumvent or diminish the detrimental effect of noise. This has entailed the study of particular systems with non-trivial noise-models [22–26], and nonlinear interactions [6, 27] that enable quantum error-correction codes [28–31].

Most quantum metrology schemes found in the literature and their corresponding bounds are deterministic. That is, these schemes are optimized in order to provide a valid estimate for each possible measurement outcome, in such a way that the average precision is maximized. Recently it has been shown that for a fixed probe state, and in the absence of noise, the precision of some particular (favorable) outcomes can be

greatly enhanced, well beyond the limits set for deterministic strategies [32–35]. The possibility to post-select, i.e., to abstain from providing an estimate some times, can even change the uncertainty from SQL to Heisenberg scaling. It has also been shown that the limit on the precision of these probabilistic metrology strategies agrees with that found for deterministic strategies when the optimization over probe states is performed. So, for pure states, probabilistic metrology can compensate a bad choice of probe state, or in other words, it can attain the optimal precision bounds in situations where the probe state is a given.

Here we study the performance of probabilistic metrology in the presence of noise. We will show that probabilistic metrology can substantially lessen the effects of local dephasing noise, although not enough to overcome the infamous loss of asymptotic Heisenberg scaling [17, 18]. In addition, and in contrast to the noiseless ideal case, the ultimate precision bounds for probabilistic metrology will be shown to exceed those attained by deterministic strategies optimized over probe states.

To put these results into context, we recall that in most quantum metrology schemes the probe is a composite made up of a large number of elementary quantum systems [1–35]. We then envisage the following situation: an ensemble of fifty-thousand two-level atoms has been prepared in a known probe state ρ and is awaiting for a signal coming from a supernova, such as a gravitational wave or some byproduct of a gamma burst. The experiment is designed in a way that the signal will leave an imprint on the state of the atoms, $\rho \to \rho_\theta$, that will depend on the value of some relevant physical parameter θ of the signal. The experimenter will perform a measurement on the atoms and will try to infer from the outcome the unknown value of θ. The experimental set-up is perfectly calibrated and characterized. At a certain time the long-awaited event occurs. Conditional on the obtained measurement outcome, the experimentalist reports a value of $\theta = 2.23$ with a mean square error of $\sigma^2 = 10^{-7}$. How should the community react if the reported error is smaller than the (deterministic) bounds found in current literature, e.g., $\langle \sigma^2 \rangle > 1/n = 2 \times 10^{-4}$?

The direct answer is that the community should celebrate the result without reservation. The error obtained in a single outcome can be smaller than the corresponding limits found in the literature, which are based on bounds on the average error over all outcomes. It is no surprise that the precision depends on what particular observation one happens to obtain: some observations are better, more informative, than others. The apparent contradiction disappears when one realizes that ultra-precise outcomes can only occur if infra-precise outcomes also exist (so as to respect the deterministic bounds).

Here we show that by a suitable choice of measurement it is possible to obtain *ultra-precise* outcomes, whose precision is still limited, but goes beyond the (average) bounds established for deterministic quantum metrology protocols. We also show that such ultra-precise outcomes can only occur with a small probability.

In a scenario where each outcome can have a different precision, the criterion for optimality is by no means unique. In a classical setting there is no compromising choice to be made. One can find the optimal estimator and precision for each measurement outcome independently. However, in the quantum case one needs to fix the POVM through some criteria. The ultimate quantum limit is obtained by choosing a POVM that produces an outcome with the highest possible precision; deterministic protocols are optimized in order to produce the highest possible precision on average (over all possible outcomes). The protocols that we study here under the name of probabilistic quantum metrology interpolate between these two cases, by optimizing the precision of successful outcomes with the constrain that they occur with some prescribed probability.

Understanding the power of probabilistic operations in general quantum tasks is a highly non-trivial and relevant problem. Probabilistic operations introduce through normalization a very particular nonlinearity that is in stark contrast with the linearity of quantum deterministic operations. Many no-go theorems stem from linearity and probabilistic operations might revoke them, turning the once-thought impossible into possible. Notable examples include unambiguous state discrimination, whereby non-orthogonal states can be distinguished with no error [36] (see [37, 38] for generalizations); probabilistic cloning [39] the KLM scheme [40], whereby Bell measurements can be realized by linear optical elements [41]; also, related to the current work (see discussion) we find probabilistic amplification [42–47] and weak-value amplification [48–52].

Although the probability of attaining the ultimate bounds is often small, at a fundamental level it is important to distinguish between ultimate versus *de facto* quantum limits. No matter how unlikely an event is, once it occurs it is a certainty; and certainties cannot violate ultimate bounds [53]. This fundamental distinction has also motivated the definition of a complexity class in quantum computing [54]. All in all, post-selection can be considered a resource per se in quantum information tasks, and this paper is devoted to the study of its power for metrological tasks in realistic noisy scenarios.

Our work becomes particularly relevant in applications where: (a) *there are high demands on precision*. We are already at a stage where quantum metrology is required to push the limits of precision. Hence it might well be that a known optimal deterministic protocol, e.g., the phase covariant measurement, fails to provide the precision required for a specific task, whereas a probabilistic scheme does not. There are tasks for which having an estimate below a certain precision is *at least* as bad as having no estimate at all. For instance, when locating a tumor in radiation therapy, or some deeply buried magnetized material for its extraction, missing the true

position of the target by more than certain threshold value can have disastrous consequences. (b) *Resources are fixed*. As in the first example above, it might be impossible to repeat an experiment (for a given instance of the unknown parameter).

In real life applications, we care about the precision attainable in absolute terms. We wish to add consistent error bars to our estimates, not just demonstrate a particular scaling of the uncertainty as the number of resources increases. Our results hold both for finite and asymptotically large number of resources. Fixing the number of resources is not only necessary to state optimality in a clear-cut way. In many situations the limitations on the available resources are patent. It might be a given, as in the measurement of the magnetization of a particular magnet; or in the measurement of a parameter that changes rapidly with time (a requirement in some feedback schemes). Last but not least, there are situations in which the experiment is impossible to reproduce because the event under study is uncontrollable, e.g., a supernova or the arrival of a gravitational wave. It is precisely in observational astronomy where the (classical) bayesian approach to statistical inference, on which our analysis relies, is widely used in current studies (see for instance [56]).

This paper is organized as follows. Section 2 describes the theoretical framework of this work. We state the general working assumptions and discuss the statistical approach(es) used throughout the paper. Section 3 contains the core findings of our work. To ease the presentation, we have divided it into different sections that contain different results. We give general expressions for the uncertainty and probability of success for covariant measurements, closed expressions for asymptotically large number of resources as well as the optimal probe states and the ultimate precision bounds. We also address the issue of the information left in the system after a discarded event. In the last section we state our main conclusions and discuss possible implementations of our scheme. More technical results are presented in the appendices. In particular, appendix A introduces specific notation that is used in the derivation of some of these results.

2. Framework

In this section we introduce our framework in detail. We set out our physical assumptions and goals, and discuss why, in view of these assumptions, the Bayesian formulation suits our purpose better, while also allowing for a straightforward extension to the probabilistic case. Alternatively, the minimax approach, or worst case scenario, is also well suited. For the problem at hand, the later is shown to be equivalent to the Bayesian formulation. The relationship with the frequentist approach is discussed at the end of the section.

Our framework consists of the following:

(a) *A model*. We assume there exists a model that formalizes the 'state of the world' and the measurement. In our case the former is given by the quantum state ρ_θ (see section 3) of a physical system of interest, where θ is a real parameter. It can, of course, take into account noise sources and other experimental imperfections. The measurement outcomes and the state of the world are related through a measurement model (Born's rule in our case) that gives the probability distribution of the outcomes for a given θ. The true value of θ is assumed to be unknown, while the rest of the model parameters are known with high precision. Our goal is to infer the value of θ from the observed measurement outcome.

(b) *A fixed amount of resources*. We view the size of the probe state ρ_θ as the amount of resources of the problem. More precisely, we quantify the amount of resources by the number n of qubits that the estate ρ_θ describes. Accordingly, if an experiment consists of repeating a measurement on independent copies of a system of N qubits a number ν of times, the total number of resources used is $n = \nu N$.

(c) *An optimization over measurements*. To establish ultimate bounds on the precision that can be achieved with given resources, we optimize over all measurements, the most general being a *single collective* measurement on the whole available resources, i.e., on ρ_θ. Hence, in our framework the inference protocols are single-shot. Namely, in each instance of the problem the state of the world is labeled by a particular (unknown) value of θ, and the measurement returns a single outcome χ out of the various possible outcomes of the measurement. Based on χ, an estimate θ_χ of the true value of θ is produced.

Note that this characterization is fully general and includes strategies where each qubit or subsystem is measured individually, in which case every collective outcome is labelled by a sequence of individual measurement output labels. Without loss of generality, we can label the measurement outcomes by the value $\hat\theta$ of the corresponding estimate (see section 3).

(d) *A report of precision*. After a successful completion, the scheme should return an estimate $\hat\theta$ of the true parameter θ, together with suitable error bars. Error bars are essential in any scientific or technological

discipline, as they quantify the confidence one should place in conclusions drawn from existing data. Such an assesment of the precision should be quantified bearing in mind that the whole set-up is single-shot, i.e., the experiment will not be necessarily repeated, and that the true value of the parameter is unknown. Hence, the precision assessment can only be based on the measurement outcome and on the precise knowledge of the model in (a).

To carry out (c) and (d) we need to introduce a so-called cost or loss function $\ell(\theta, \hat{\theta})$ that quantifies how well our estimate $\hat{\theta}$ agrees with the true value of θ. There are a priory infinitely many such functions, but two common choices in metrology are the quadratic loss function $\ell_q(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2$, and $\ell_p(\theta, \hat{\theta}) = 4\sin^2[(\theta - \hat{\theta})/2]$ if θ belongs to a periodic domain. Note that they are equivalent to leading order in $\hat{\theta} - \theta$, when the estimate approaches the true value, $\hat{\theta} \approx \theta$.

The Bayesian formulation offers a very natural and rigorous way to assign a quantitative precision measure to a particular outcome $\hat{\theta}$. In this formulation the unknown parameter θ is treated as a random variable and it is assigned a (prior) probability $\pi(\theta)$. This probability reflects the knowledge we have on the state of the world prior to the measurement. After performing the measurement, the observed outcome and our knowledge of the model is used to update the prior $\pi(\theta)$ to the posterior probability distribution $p(\theta|\hat{\theta})$. Using Bayes' rule, we can write $p(\theta|\hat{\theta}) = p(\hat{\theta}|\theta)\pi(\theta)/p(\hat{\theta})$, with $p(\hat{\theta}) = \int d\theta\, p(\hat{\theta}|\theta)\pi(\theta)$. Then the uncertainty of an outcome $\hat{\theta}$ is defined as

$$L_{\hat{\theta}} = \int d\theta\, p(\theta|\hat{\theta})\ell(\theta, \hat{\theta}). \qquad (1)$$

Thus, $L_{\hat{\theta}}$ quantifies how the unknown value θ is scattered around its estimate $\hat{\theta}$, in the light of the information gathered by the measurement.

In general, the various outcomes of a given measurement have different precision. Hence, to quantify the overall performance of a metrology scheme by a single figure of merit we take the average uncertainty over all possible measurement outcomes

$$L = \int d\hat{\theta}\, p(\hat{\theta})\, L_{\hat{\theta}} = \int d\hat{\theta} \int d\theta\, p(\theta, \hat{\theta})\, \ell(\theta, \hat{\theta}), \qquad (2)$$

where the joint probability is $p(\theta, \hat{\theta}) = p(\hat{\theta}|\theta)\pi(\theta)$. Equation (2) is the expected loss L, given by the average over all possible values of the unknown parameter θ and over all measurement outcomes.
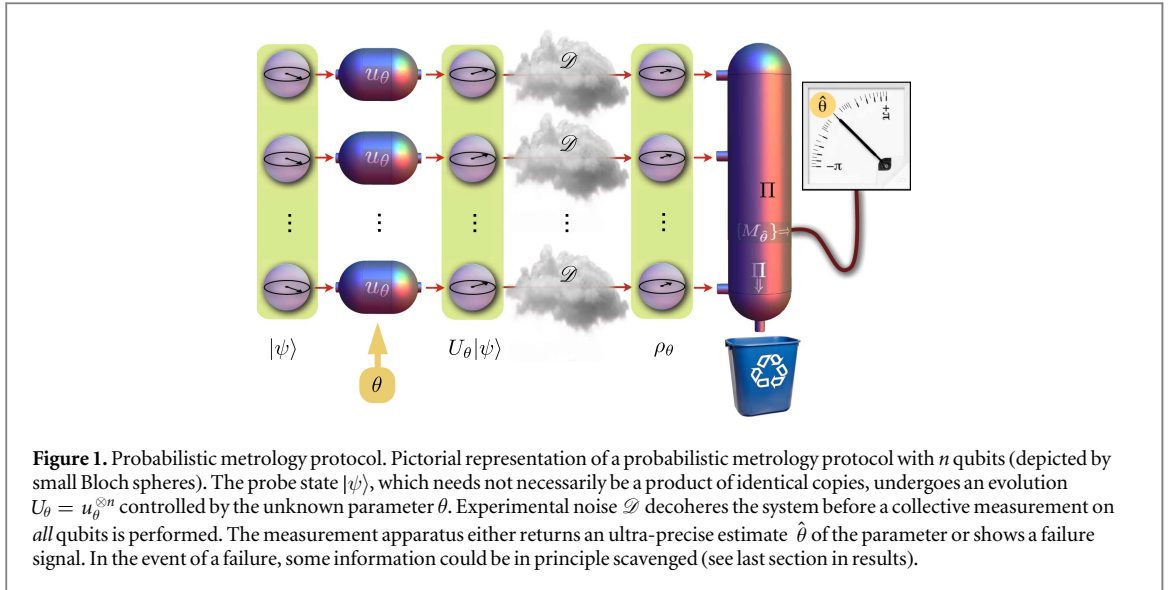
With this in mind, we now focus on probabilistic metrology. As discussed in the introduction, we can improve performance if we give up on the idea of deterministic protocols, by allowing for failures to perform the tasks they have been designed for. Accordingly, probabilistic metrology protocols will either succeed and provide a precise estimate $\hat{\theta}$, or warn of failure (abstain). Following these premises, the figure of merit for such protocols are given by the average uncertainty of the successful outcomes, i.e., by

$$L_s = \int d\hat{\theta}\, p(\hat{\theta}|\text{succ})L_{\hat{\theta}}. \qquad (3)$$

Conditional expectations such as this are the cornerstone of bayesian estimation. Their use is wide-spread and established in a number of disciplines, such as control theory or signal processing, where an accurate and rigorous assessment of the precision is required—see for instance [57]. In order to give a complete characterization of the probabilistic protocol, one should supplement the attained uncertainty L_s, with the corresponding probability of success, S. We will derive the tradeoff curve $L_s(S)$ that gives the minimum uncertainty for every fixed value of the success probability S. In particular, by computing $\lim_{S\to 0} L_s(S)$ we will show that there is an *ultimate quantum limit* in the precision of an estimate inferred from *any outcome* of a quantum measurement.

At this point, it should be clear that a probabilisitic protocol, as defined above, is not meant to be repeated until it succeeds [58]. Obviously, such a strategy would be ultimately deterministic (it will always end up providing an estimate) and, thus, it could not outperform the optimal deterministic protocol for the same total amount of resources. Only with some pre-establish success probability can probabilistic metrology provide a guaranteed precision for a given amount of resources.

We next outline an alternative approach that is often used in quantum metrology and point out the differences with the global single-shot framework defined above. The so-called pointwise approach aims at minimizing the dispersion of the estimates $\hat{\theta}$ that results from the noise inherent to quantum measurements. It assumes that the true value of the parameter θ is fixed (i.e., it is not a random variable), and that the metrology protocol can be repeated an arbitrary number of times; it is a frequentist framework. It is customary to quantify the precision of the protocol by the mean square error

Figure 1. Probabilistic metrology protocol. Pictorial representation of a probabilistic metrology protocol with n qubits (depicted by small Bloch spheres). The probe state $|\psi\rangle$, which needs not necessarily be a product of identical copies, undergoes an evolution $U_\theta = u_\theta^{\otimes n}$ controlled by the unknown parameter θ. Experimental noise $\mathscr{D}$ decoheres the system before a collective measurement on *all* qubits is performed. The measurement apparatus either returns an ultra-precise estimate $\hat{\theta}$ of the parameter or shows a failure signal. In the event of a failure, some information could be in principle scavenged (see last section in results).

$$\text{MSE}_\theta = \int d\hat{\theta}\, p(\hat{\theta}|\theta)\ell_q(\theta, \hat{\theta}), \tag{4}$$

which indeed gives a measure of how the estimates would scatter around the true value if the protocol is repeated many times. Note that if a prior $\pi(\theta)$ were supplied, one could compute the average over θ, thus recovering the expression of the Bayesian expected loss in equation (2) for the quadratic loss function.

The celebrated quantum Cramer–Rao bound [59] provides a lower bound to MSE_θ that can be readily computed. In addition, one can often argue that the Cramer–Rao bound can be attained in the asymptotic limit of large number of resources by a suitable two-step adaptive protocol. However, the assumptions under which the quantum Cramer–Rao bound holds, and the conditions under which the bound is attained entail some subtleties that are often ignored and that can lead to erroneous conclusions [60, 61] and misleading accounting of resources (see for instance [62, 63]). In the particular case of probabilistic metrology, the direct application of the pointwise approach can lead to unphysical results, as pointed out in [64].

The Bayesian approach has been widely used to assess the performance of quantum information protocols such as teleportation, state estimation, universal cloning and quantum memories. Despite its many advantages, which include a straightforward accounting of resources and its validity even for a small (non-asymptotic) number of resources, it also has some drawbacks: optimal bounds are usually hard to compute and there is no general prescription to choose the prior $\pi(\theta)$. In the case at hand (estimation of a phase θ) these drawbacks can be easily evaded, as the symmetry of the problem simplifies the calculations significantly while providing a valid justification (Laplaces principle of insufficient reason) to choose a uniform prior on $(-\pi, \pi]$.

There is still another approach that suits our framework and does not require a prior distribution: the mini-max approach, whereby the average over the unknown parameter θ is replaced by its worst-case value:

$$L_{wc} = \sup_{\theta \in (-\pi, \pi]} \int d\hat{\theta}\, p(\hat{\theta}|\theta, \text{succ})\ell(\theta, \hat{\theta}), \tag{5}$$

where the optimization is over all possible quantum measurements. As shown in appendix F, for phase estimation the optimal worst-case loss, equation (5), and the expected loss, equation (3) are equivalent.

3 Results

3.1. Optimal probabilistic measurement for *n*-qubits
In the scope of this paper, metrology aims at estimating the parameter θ that determines the unitary evolution, $U_\theta := u_\theta^{\otimes n}$, of a probe system of n qubits in the presence of local decoherence, where $u_\theta = \exp(i\theta|1\rangle\langle 1|)$.

As depicted in figure 1, the initial n-partite pure state $|\psi\rangle\langle\psi| = \psi$ (this shorthand notation will be used throughout the paper) is prepared and is let evolve. The state is affected by uncorrelated dephasing noise, which can be modeled by independent phase-flip errors occurring with probability $p_f = (1 - r)/2$ for each qubit. Its action on the n-qubits is described by a map $\mathscr{D}$ that commutes with the Hamiltonian, so that it could as well be understood as acting before or during the phase imprinting process.

Next, the experimentalist performs a suitable measurement on $\rho_\theta = \mathscr{D}(U_\theta \psi U_\theta^\dagger)$ and, based on its outcome, decides whether to abstain or to produce an estimate $\hat{\theta}$ for the unknown parameter θ. Note that this decision is

based solely on the outcome of the measurement as, naturally, the actual value of θ is unknown to the experimentalist. Our aim is to find the optimal protocol, e.g., the measurement that gives the most accurate estimates for a given probe state and for a given maximum probability of abstention.

Motivated by the periodicity of the phase, we quantify the uncertainty of the estimated phase $\hat{\theta}$ by the periodic loss function $\ell_{\mathrm{p}}(\theta, \hat{\theta})$ defined in section 2, and to assess the performance of the protocol we use the expected loss defined in equations (1) and (3)

$$
\begin{aligned}
\sigma^2 \equiv L_{\mathrm{s}} &= \iint \mathrm{d}\hat{\theta}\,\mathrm{d}\theta\, p(\theta|\hat{\theta})\, p(\hat{\theta}|\mathrm{succ})\, \ell_{\mathrm{p}}(\theta, \hat{\theta}) \\
&= \iint \mathrm{d}\hat{\theta}\,\mathrm{d}\theta\, p(\theta, \hat{\theta}|\mathrm{succ})\, \ell_{\mathrm{p}}(\theta, \hat{\theta}) \\
&= \frac{1}{S} \iint \mathrm{d}\hat{\theta}\,\mathrm{d}\theta\, p(\theta, \hat{\theta}, \mathrm{succ})\, \ell_{\mathrm{p}}(\theta, \hat{\theta}),
\end{aligned}
\tag{6}
$$

where the success probability is $S = \iint \mathrm{d}\hat{\theta}\,\mathrm{d}\theta\, p(\theta, \hat{\theta}, \mathrm{succ})$. Throughout the rest of the paper we will refer to σ^2 as the uncertainty for brevity. The uncertainty and the probability of success S will fully characterize our probabilistic metrology strategies. In adition, in the asymptotic limit of large number of resources the distribution $p(\hat{\theta}|\theta, \mathrm{succ})$ becomes peaked around the true value θ and the uncertainty (expected loss) approximates the mean-square error (expected loss for quatratic loss function ℓ_{q}).

The set $\{\rho_\theta\}$ is a so-called *covariant family of states* [65], as it is generated by the action of a group of unitaries; $\{U_\theta\}_{\theta \in (-\pi, \pi]}$ in our case. We also note that $\ell_{\mathrm{p}}(\theta, \hat{\theta})$ is invariant under the same group action, namely, $\ell_{\mathrm{p}}(\theta + \theta', \hat{\theta} + \theta') = \ell_{\mathrm{p}}(\theta, \hat{\theta})$ for all $\theta' \in (-\pi, \pi]$. Because of this, there is no loss of generality in choosing the measurement to be covariant [65]. Such *covariant measurements* are defined by $\{M_{\hat{\theta}} = U_{\hat{\theta}}\Omega U_{\hat{\theta}}^\dagger/(2\pi)\}_{\hat{\theta} \in (-\pi, \pi]}$, where Ω is the so-called *seed* of the measurement. In addition, we have the invariant measurement operator $\Pi = \mathbb{1} - \int_{-\pi}^{\pi} \mathrm{d}\hat{\theta}/(2\pi) U_{\hat{\theta}}\Omega U_{\hat{\theta}}^\dagger \leqslant \mathbb{1}$ that corresponds to the abstention event. With this, finding the optimal estimation scheme reduces to finding the operator Ω that mimimizes the uncertainty

$$
\sigma^2(S) = \frac{1}{S} \min_{\Omega} \int_{-\pi}^{\pi} \frac{\mathrm{d}\hat{\theta}}{2\pi} \ell_{\mathrm{p}}(0, \hat{\theta})\, \mathrm{tr}(U_{\hat{\theta}}\Omega U_{\hat{\theta}}^\dagger \rho),
\tag{7}
$$

for a fixed success probability [33]

$$
S = \int_{-\pi}^{\pi} \frac{\mathrm{d}\hat{\theta}}{2\pi} \mathrm{tr}(U_{\hat{\theta}}\Omega U_{\hat{\theta}}^\dagger \rho).
\tag{8}
$$

In deriving equation (7) we have used covariance to fix the value of θ to zero and thereby get rid of the integral over θ in equation (6), and have defined $\rho = \mathscr{D}(\psi)$ accordingly.

3.2. Symmetric probes

We now focus on probe states consisting n-qubits that are initially prepared in a permutation invariant state. This family includes most of the states considered in the literature, our case-study of multiple copies of equatorial-states, and also, as we will show below, the optimal probe-state for probabilistic metrology. The input state is given by

$$
|\psi\rangle = \sum_{m=-J}^{J} c_m |J, m\rangle,
\tag{9}
$$

where $J = n/2$ is the maximum total spin angular momentum (hereafter spin for short) of n qubits and the set of states $\{|J, m\rangle\}_{m=-J}^{J}$ spans the fully symmetric subspace. Given the permutation invariance of the noisy channel, the state $\rho = \mathscr{D}(\psi)$ inherits the symmetry of the probe, and can be conveniently written in a block diagonal form in the total spin bases [66, 67] (see appendix C)

$$
\rho = \sum_j p_j \rho^j \otimes \frac{\mathbb{1}_j}{\nu_j},
\tag{10}
$$

where the state ρ^j has unit trace, p_j is the probability of ρ having spin j, and $\mathbb{1}_j$ stands for the identity in the ν_j-dimensional multiplicity space of the irreducible representation of spin j. The sum over j in equation (10) runs from $j_{\min} = 0$ ($j_{\min} = 1/2$) for n even (odd) to the maximum spin J. Similarly, the measurement operators, can be taken to have the same symmetry and thus be of the form $\Omega = \sum_j |\chi_j\rangle \langle \chi_j| \otimes \mathbb{1}_j$, where $|\chi_j\rangle = \sum_m f_m^j |j, m\rangle$, $0 \leqslant f_m^j \leqslant 1$. The minimum uncertainty $\sigma^2(S)$ for a fixed probability of success S can hence be expressed in terms of the uncertainty $\sigma_j^2(s_j)$ in each irreducible block and its corresponding success probability s_j

$$\sigma^2(S) = \min_{\{s_j\}} \sum_j \frac{p_j s_j}{S} \sigma_j^2(s_j), \quad S = \sum_j p_j s_j, \tag{11}$$

where $\sigma_j^2(s_j) [s_j]$ is defined by equation (7) (equation (8)) with Ω, ρ and $U_{\hat{\theta}}$ projected onto the subspace of total angular momentum j. Recalling that $U_{\hat{\theta}} = \sum_j U_{\hat{\theta}}^j \otimes \mathbb{1}_j$, $[U_{\hat{\theta}}^j]_{m,m'} = \delta_{m,m'} e^{im\hat{\theta}}$ $(-j \leqslant m, m' \leqslant j)$ and $\ell_p(0, \hat{\theta}) = 2 - e^{i\hat{\theta}} - e^{-i\hat{\theta}}$, one can easily integrate $\hat{\theta}$ to obtain

$$\sigma_j^2(s_j) = 2 - \frac{2}{s_j} \max_{0 \leqslant f_m^j \leqslant 1} \sum_m f_m^j \rho_{m,m+1}^j f_{m+1}^j,$$
$$\text{subject to } s_j = \sum_m (f_m^j)^2 \rho_{m,m}^j. \tag{12}$$

This formulation of the problem allows for a natural interpretation of the probabilistic protocol as a two step process: (i) a stochastic filtering channel

$$\mathcal{F}(\rho) = \Phi \rho \Phi, \qquad \Phi = \sum_{j,m} f_m^j |j, m\rangle \langle j, m| \otimes \mathbb{1}_j, \tag{13}$$

that coherently transforms each basis vector as $|j, m\rangle \to f_m^j |j, m\rangle$, so that it modulates the input to a state with enhanced phase-sensitivity, followed by (ii) a canonical covariant measurement with seed $\tilde{\Omega} = \sum_j \sum_{m,m'} |j, m\rangle\langle j, m'| \otimes \mathbb{1}_j$ performed on the transformed state from which the value of the unknown phase is estimated.

By defining the vector ξ^j with components given by $\xi_m^j = f_m^j (\rho_{m,m}^j / s_j)^{1/2}$ and introducing the tridiagonal symmetric matrix H^j, with entries

$$H_{m,m'}^j = 2\delta_{m,m'} - a_m^j \delta_{m,m'-1} - a_{m'}^j \delta_{m-1,m'},$$
$$a_m^j = \frac{\rho_{m,m+1}^j}{\sqrt{\rho_{m,m}^j \rho_{m+1,m+1}^j}}, \tag{14}$$

we can easily recast the former optimization problem as

$$\sigma_j^2 := \min_{|\xi^j\rangle} \langle \xi^j | H^j | \xi^j \rangle, \tag{15}$$

$$\text{subject to } \langle \xi^j | \xi^j \rangle = 1 \text{ and } 0 \leqslant \xi_m^j \leqslant (\rho_{m,m}^j / s_j)^{1/2}. \tag{16}$$

Note that a_m^j, and in turn H^j, depend on the strength of the noise but they take the same values for *all* symmetric probe states, since $\rho_{m,m'}^j \propto c_m c_{m'}$. For deterministic strategies ($S = 1$, i.e., $s_j = 1$ for all j) no minimization is required and one only needs to evaluate the expectation values of H^j for the 'state' $\xi_m^j = (\rho_{m,m}^j)^{1/2}$. For large enough abstention, the problem becomes an unconstrained minimization, so σ_j^2 is the minimal eigenvalue of H^j, and $|\xi^j\rangle$ its corresponding eigenstate. From equation (16), we find that the corresponding filtering operation only succeeds with a probability

$$S^* = \sum_j p_j s_j^*, \qquad s_j^* = \min_m \frac{\rho_{m,m}^j}{\xi_m^{j\,2}}. \tag{17}$$

We will refer to S^* as the critical success probability, since the precision will not improve by decreasing the success probability below this value: $\sigma^2(S) = \sigma^2(S^*)$ for $S \leqslant S^*$.

3.3. Asymptotic scaling: particle in a potential box

In order to compute the scaling of the uncertainty as the number of resources becomes very large we need to solve the above optimization problem in the asymptotic limit of $n \to \infty$. We start be analyzing the uncertainty $\sigma_j^2(s_j)$ for blocks of large j. As shown in appendix G, for each such block we define the ratios $x = m/j$, $m = -j, -j+1, \ldots, j$, that approach a continuous variable as $j \to \infty$. In this limit, $\{\sqrt{j} \xi_m^j\}$ approaches a real function of x, $\sqrt{j} \xi_m^j \to \varphi(x)$, and the expectation value in equation (15) becomes

$$\sigma_j^2 = \frac{1}{j^2} \min_{|\varphi\rangle} \int_{-1}^1 dx \left\{ \left[\frac{d\varphi(x)}{dx} \right]^2 + V^j(x)\varphi(x)^2 \right\}$$
$$:= \frac{1}{j^2} \min_{|\varphi\rangle} \langle \varphi | \mathcal{H}^j | \varphi \rangle, \tag{18}$$

where we have dropped some boundary terms that are irrelevant for this discussion, and where $\mathcal{H}^j := -d^2/dx^2 + V^j(x)$ plays the role of a 'Hamiltonian', with a 'potential'

$$V^j(x) = 2j^2(1 - a_m^j) = j\frac{1 - r^2}{2r\sqrt{1 - (1 - r^2)x^2}}. \tag{19}$$

Furthermore, in equation (18) the function $\varphi(x)$ must be also differentiable and must satisfy the conditions

$$\langle \varphi | \varphi \rangle = \int_{-1}^{1} dx \, [\varphi(x)]^2 = 1, \quad \varphi(x) \leqslant \frac{\tilde{\varphi}(x)}{\sqrt{s_j}}, \tag{20}$$

where for a given large j we define $\tilde{\varphi}(x)$ through

$$\sqrt{j\rho_{mm}^j} \rightarrow \tilde{\varphi}(x), \quad x = \frac{m}{j}. \tag{21}$$

It is now apparent from equations (18) through (21) that our optimization problem is formally equivalent to that of finding the ground-state wave-function of a quantum particle in a box ($-1 \leqslant x \leqslant 1$) for the potential $V^j(x)$ and subject to boundary conditions that are fixed by the probe state, the strength of the noise, and the success probability. Other equivalent variational formulations can be found in [33, 34, 68] for pure states and in [69] for the pointwise approach.

3.4. Multiple-copies
Although our methods apply to general symmetric probes, for the sake of concreteness we study in full detail the paradigmatic case of a probe consisting of n identical copies of equatorial qubits:

$$|\psi_{\mathrm{cop}}\rangle = \frac{1}{\sqrt{2^n}}(|0\rangle + |1\rangle)^{\otimes n}. \tag{22}$$

Decoherence turns this symmetric pure state to a full rank state with a probability of having spin j given by

$$p_j \simeq \frac{e^{-J\frac{(j/J - r)^2}{1 - r^2}}}{\sqrt{\pi J(1 - r^2)}}, \tag{23}$$

where this approximation is valid around its peak, at the typical value $j_0 = rJ$. For each irreducible block and before filtering we have a signal
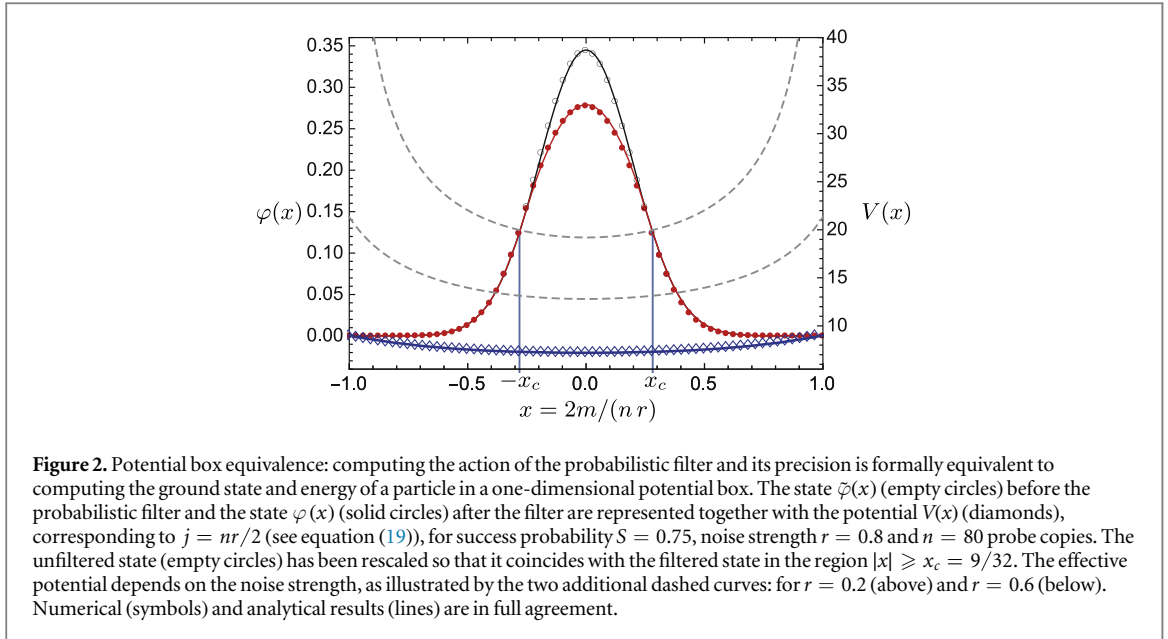
$$\sqrt{j\rho_{mm}^j} \rightarrow \tilde{\varphi}(x) \simeq \left(\frac{jr}{\pi}\right)^{\frac{1}{4}} e^{-\frac{rj}{2}x^2} \tag{24}$$

that peaks at $x = 0$ with variance $\langle x^2 \rangle = (2rj)^{-1}$.

For deterministic protocols ($S = 1$) the constraints completely fix the solution: $\varphi(x) = \tilde{\varphi}(x)$. The corresponding uncertainty is obtained by computing the 'mean energy' $\sigma_j^2 = \langle \mathcal{H}^j \rangle_{\tilde{\varphi}} / j^2$, in equation (18). For large j it is meaningful to use the harmonic approximation $V^j(x) \simeq V_0^j + \omega_j^2 x^2$, where $V_0^j = j(1 - r^2)/(2r)$ and $\omega_j^2 = j(1 - r^2)^2/(4r)$. The leading contribution to σ_j^2 comes from the 'kinetic energy' (i.e., the first term in equation (18)), which gives $\langle p^2 \rangle_{\tilde{\varphi}} = (1/4)\langle x^2 \rangle^{-1} = jr/2$, whereas the harmonic term gives a sub-leading contribution. One easily obtains $\sigma_j^2 = (2jr)^{-1}$. The leading contribution to the uncertainty of the deterministic protocol is given by σ_j^2 at the typical spin j_0: $\sigma_{\mathrm{det}}^2 = (2Jr^2)^{-1} = (nr^2)^{-1}$, in agreement with the previous known (pointwise) bounds.

For unlimited abstention in a block of given spin j (s_j very small) the minimization in equation (18) is effectively unconstrained and the solution (the filtered state) is given by the ground state $\varphi^g(x)$ of the potential $V^j(x)$. Within the harmonic approximation, we notice that the effective frequency of the oscillator grows as $\sqrt{j}$, and the corresponding gaussian ground state is confined around $x = 0$ with variance $\langle x^2 \rangle = (r/j)^{1/2}(1 - r^2)^{-1}$. In this situation both the kinetic and harmonic contributions to the 'energy' are sub-leading, and so are the higher order corrections to $V^j(x)$. Thus, the uncertainty σ_j^2 for spin j is ultimately limited by the constant term V_0^j of the potential. Up to sub-leading order one obtains $\sigma_j^2 = (1 - r^2)(2jr)^{-1}[1 + (r/j)^{1/2}]$. The filtering of $\tilde{\varphi}(x)$ to produce the gaussian ground state $\varphi^g(x)$ succeeds with probability $s_j^* \sim e^{-2j\log(1 + r)}$ (see appendix D). Note that in the absence of noise ($r = 0$) the potential $V^j(x)$ vanishes and the ground state is solely confined by the bounding box $-1 \leqslant x \leqslant 1$. Then, $\varphi^g(x) = \cos(\pi x/2)$, which results in a Heisenberg limited precision (the ultimate pure-state bound): $\sigma^2 = \pi^2/n^2$ [33, 68].

If the optimal filtering is performed on typical blocks, $j \approx j_0$, one obtains $\sigma^2 = (1 - r^2)/(nr^2)$, which coincides with the ultimate deterministic bound found in [17, 69]. This shows that a probabilistic protocol that uses the uncorrelated multi-copy probe state $|\psi_{\mathrm{cop}}\rangle$ can attain the precision bound of a deterministic protocol, for which a highly entangled probe is required. This bound is attained for a critical success probability $S^* \simeq s_{j_0}^* \sim e^{-nr\log(1 + r)}$. More interestingly, we can push the limit further by post-selecting on the block with highest spin (by choosing $f_m^j \propto \delta_{j,J}$) to obtain

Figure 2. Potential box equivalence: computing the action of the probabilistic filter and its precision is formally equivalent to computing the ground state and energy of a particle in a one-dimensional potential box. The state $\tilde{\varphi}(x)$ (empty circles) before the probabilistic filter and the state $\varphi(x)$ (solid circles) after the filter are represented together with the potential $V(x)$ (diamonds), corresponding to $j = nr/2$ (see equation (19)), for success probability $S = 0.75$, noise strength $r = 0.8$ and $n = 80$ probe copies. The unfiltered state (empty circles) has been rescaled so that it coincides with the filtered state in the region $|x| \geqslant x_c = 9/32$. The effective potential depends on the noise strength, as illustrated by the two additional dashed curves: for $r = 0.2$ (above) and $r = 0.6$ (below). Numerical (symbols) and analytical results (lines) are in full agreement.

$$\sigma^2_{\text{ult}} := \sigma^2_{j \approx J} = \frac{1 - r^2}{nr}\left(1 + \sqrt{\frac{2r}{n}}\right),\tag{25}$$

with a critical probability given by $S^* = p_J s_J^* \sim \mathrm{e}^{-n \log 2}$, independently of the noise strength. We note that the leading order is a factor r smaller than the previously established (deterministic) bound, $\sigma^2 = (1 - r^2)/(nr^2)$ [17, 69]. This important enhancement in precision results from post-selection of high-angular momentum, which does not commute with the noise channel. Hence, in contrast to the noiseless scenario, post-selection is not equivalent to a suitable choice of input state.

Having understood the two limiting cases of no abstention (the deterministic protocol) and unlimited abstention, we can now quantify the asymptotic scaling for an arbitrary success probability $\sigma^2(S)$. We use the Karush–Kuhn–Tucker optimization method to minimize equation (18) under the constraints in equation (20). For a given value of s_j, the so-called complementary slackness condition [34, 70] guarantees that the solution $\varphi(x)$ to equation (18) saturates the inequality in equation (20) for x in a certain region called coincidence set, while it coincides with an eigenfunction of the Hamiltonian $\mathcal{H}^j$, defined after equation (18), for x outside this region. The continuity of $\varphi(x)$ and its derivative provide some matching conditions at the border of the coincidence set and a unique solution can be easily found.
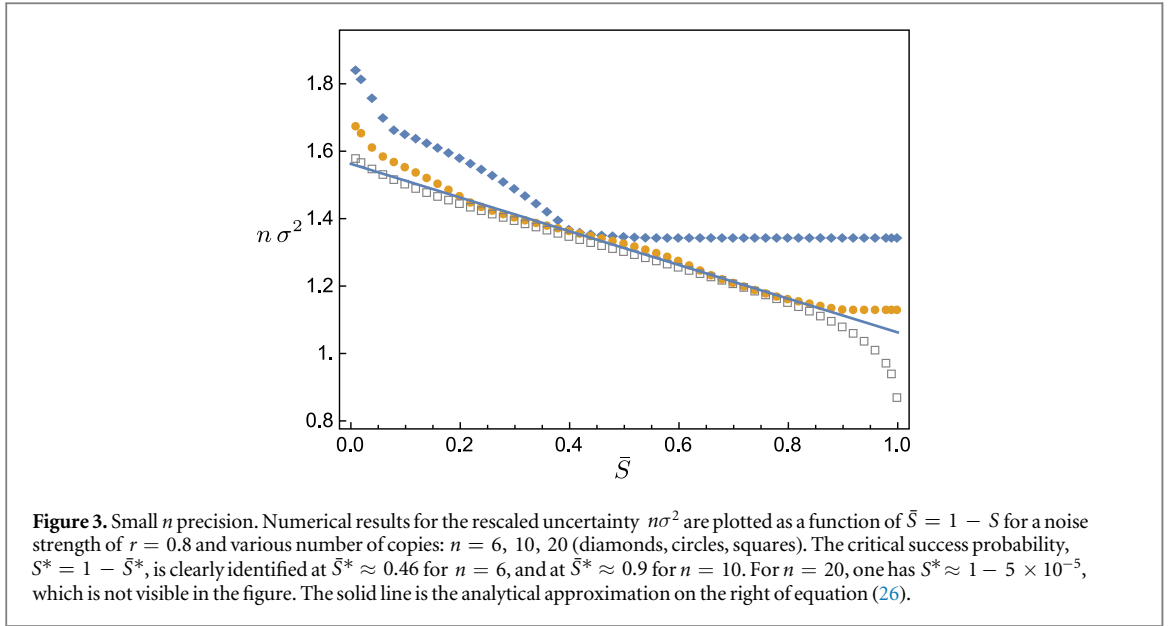
As shown in figure 2, in the case of multiple copies the tails of $\varphi(x)$ coincide with the gaussian profile in equation (24) scaled by the factor $s_j^{-1/2}$ for $|x| > x_c$ (in the coincidence set), while for $|x| < x_c$ the filter takes an active part in reshaping the peak into the optimal profile. Clearly, the wider the filtered region, the higher the precision and the abstention rate. A simple expression for the leading order can be obtained if we notice that with a finite abstention probability one can change the variance of the wave function in equation (24) but not its $1/j$ scaling. Hence, as for the deterministic case, only the kinetic energy and the constant term V_0^j of the potential play a significant role. The solution can then be easily written in terms of the pure-state solution [34], which corresponds to a zero potential inside the box $-1 \leqslant x \leqslant 1$:

$$\sigma^2 \simeq \sigma^2_{j_0} = \frac{1 - r^2}{nr^2} + r\sigma^2_{\text{pure}}(S) \approx \frac{1 - (r^2/2)\bar{S}}{nr^2},\tag{26}$$

where $\bar{S} := 1 - S$ is the probability of abstention, σ^2_{pure} is the uncertainty for pure states ($r = 1$) and for an effective number of qubits $n_{\text{eff}} = 2j_0$. The pre-factor r takes into account the scaling of the variance of the state equation (24) as compared to the pure-state case. The first equality of equation (26) uses the fact that only abstention on blocks about the typical spin j_0 is affordable for finite S. This also fixes the value of S to be approximately s_{j_0}. The simple expression on the right of equation (26) is not an exact bound, but does provide a good approximation for moderate values of $\bar{S}$ (see figure 3). We notice that for low levels of noise ($r \approx 1$) one can have a considerable gain in precision already for finite abstention.

3.5 Finite n

Up to this point, we have given analytical results for asymptotically large n, the number of resources. In order to get exact values for finite n we need to resort on numerical analysis. The main observation here is that our optimization problem can be cast as a semidefinite program:

Figure 3. Small n precision. Numerical results for the rescaled uncertainty $n\sigma^2$ are plotted as a function of $\bar{S} = 1 - S$ for a noise strength of $r = 0.8$ and various number of copies: $n = 6, 10, 20$ (diamonds, circles, squares). The critical success probability, $S^* = 1 - \bar{S}^*$, is clearly identified at $\bar{S}^* \approx 0.46$ for $n = 6$, and at $\bar{S}^* \approx 0.9$ for $n = 10$. For $n = 20$, one has $S^* \approx 1 - 5 \times 10^{-5}$, which is not visible in the figure. The solid line is the analytical approximation on the right of equation (26).

$$\sigma^2 = \min_{\Lambda \,:\, \mathcal{C}} \operatorname{tr} H \Lambda, \qquad (27)$$

subject to a set of linear conditions on the matrix Λ given by $\mathcal{C} := \{\Lambda \geqslant 0,\ \operatorname{tr} \Lambda = 1,\ \Lambda^j_{mm} \leqslant p_j \rho^j_{m,m}/S\}$, where Λ and H have the block diagonal form $\Lambda = \oplus_j \Lambda^j$ and $H = \oplus_j H^j$. Semidefinite programming problems, such as this, can be solved efficiently with arbitrary precision [70].

Figure 3 shows representative results for moderate, experimentally relevant number of qubits n. We plot the uncertainty as a function of the abstention probability $\bar{S}$ and noise strength $r = 0.8$. We observe that for small values of n the precision increases ($n\sigma^2$ decreases) quite rapidly until the critical value S^* is reached. Past this point the precision cannot be improved. For larger n, the initial gain is less dramatic, but the critical point (or plateau) is reached for higher abstention probabilities, hence allowing to reach a higher precision. We see that for moderately large n, abstention can easily provide 60% improvement of the precision. When n is large enough, e.g., $n = 20$ (see the figure), there is a sharp improvement in precision as the success probability approaches the critical value. In the asymptotic limit, $n \to \infty$, it gives rise to a critical behavior that interpolates between the ultimate precision limit, equation (25), and the precision for finite values of $\bar{S}$, equation (26).
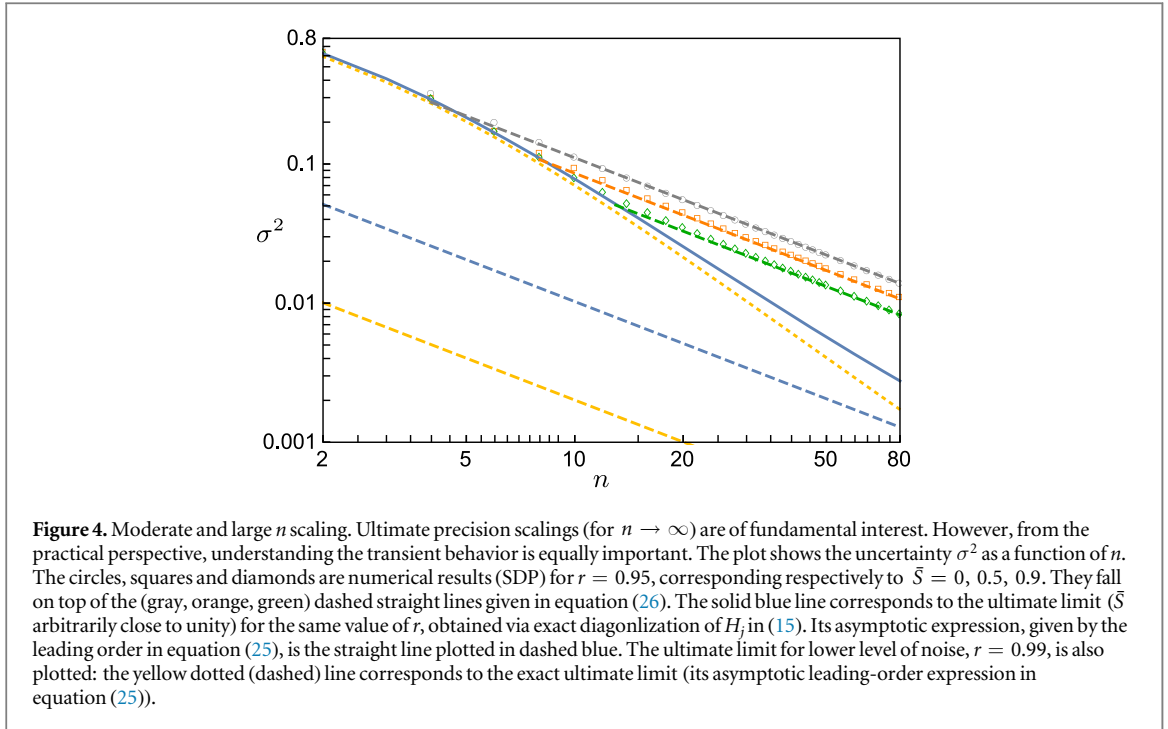
Figure 4 shows the scaling of the uncertainty with the amount of resources, n, for low levels of noise $r = 95\%$ and for different values of the abstention probability $\bar{S}$. For low n all curves exhibit a similar (n^{-1} scaling, i.e., SQL). As we increase n, very soon the curve corresponding to unlimited abstention (solid line) shows a big drop with a quantum-enhanced transient scaling given by $n^{-(\alpha+1)}$, where $\alpha > 0$ depends on the noise strength. For very large n (~ 500) this curve saturates the ultimate asymptotic limit in equation (25) (blue dashed line), which has again SQL scaling. The numerical results for finite S (circles, squares, diamonds) display the optimal scaling up to the point where they meet the asymptotic (dashed) straight lines given by equation (26). Past this point, they fall on top of the corresponding straight lines, which display SQL scaling. The larger the abstention probability, the later this transition takes place. In addition, the figure shows the ultimate scaling for $r = 99\%$ to illustrate the fact that for weaker noise levels the transient is more abrupt (α is larger).

3.6. Ultimate bound for metrology

So far we have studied the best precision bounds that can be attained for a fixed input state. A very relevant question of fundamental and practical interest is whether this bounds can be overcome by an appropriate choice of such state. We answer this question in the negative: the precision bound given by the uncertainty in equation (25) is indeed the ultimate bound for metrology in the presence of local decoherence and can only be attained by a probabilistic strategy.

To this aim, we first show in appendix B that for any probe state and any measurement that attain certain precision (or, equivalently σ^2) with success probability S, we can find a new probe lying in the fully symmetric subspace ($j = J$) and a permutation invariant measurement that attain the very same precision with the very same success probability. This shows that the formulation that we have introduced, with probe states in the fully symmetric subspace, is actually completely general.

We now recall that the Hamiltonian $\mathcal{H}^j$ is independent of the choice of probe state and that such choice determines only the shape of the state $\tilde{\varphi}(x)$ before filtering, and the probability p_j of belonging to the subspace of spin j. Since the bound in equation (25) is attained by the ground-state $\varphi^{\text{g}}(x)$ of the potential $V^J(x)$, the choice

Figure 4. Moderate and large *n* scaling. Ultimate precision scalings (for $n \rightarrow \infty$) are of fundamental interest. However, from the practical perspective, understanding the transient behavior is equally important. The plot shows the uncertainty σ^2 as a function of *n*. The circles, squares and diamonds are numerical results (SDP) for $r = 0.95$, corresponding respectively to $\bar{S} = 0, 0.5, 0.9$. They fall on top of the (gray, orange, green) dashed straight lines given in equation (26). The solid blue line corresponds to the ultimate limit ($\bar{S}$ arbitrarily close to unity) for the same value of *r*, obtained via exact diagonlization of H_j in (15). Its asymptotic expression, given by the leading order in equation (25), is the straight line plotted in dashed blue. The ultimate limit for lower level of noise, $r = 0.99$, is also plotted: the yellow dotted (dashed) line corresponds to the exact ultimate limit (its asymptotic leading-order expression in equation (25)).

of probe cannot further improve the precision, but only change the success probability. In particular one might increase *S* by choosing a probe state that gives rise to a profile $\tilde{\varphi}(x) = \varphi^g(x)$ for $j = J$, without any filtering within the block. In this case the critical success probability becomes $S^* = p_J = e^{-n[\log 2 - \log(1+r)]}$ (see appendix E), which is larger than that attained by $|\psi_{\text{cop}}\rangle$.

At the other extreme, for deterministic strategies, the calculation of $\sigma^2_{\text{opt}}(1)$ can be easily carried out by performing first the sum over *j* and then optimizing over the $(n + 1)$-dimensional probe state. In the continuum limit (large *n*) such calculation can again be cast as a variational problem formally equivalent to that of finding the ground-state of particle in a box with the harmonic potential $V(y) = nr^{-2}(1 - r^2)(1 + y^2)$, $-1 \leqslant y = m/J \leqslant 1$. The corresponding ground state wave function and its energy provide the optimal probe state and uncertainty respectively:

$$\psi_{\text{op}}(y) = \left[\frac{n(1 - r^2)}{(2\pi r)^2} \right]^{\frac{1}{8}} e^{-\frac{\sqrt{n(1-r^2)}}{4r} y^2} \tag{28}$$

and

$$\sigma^2_{\text{op}}(1) = \frac{1 - r^2}{nr^2} + \frac{2\sqrt{1 - r^2}}{n^{3/2}r}. \tag{29}$$

These results agree with their pointwise counterparts in [17, 69]. The presence of noise brings the pointwise and global approaches in agreement, as to both the attainable precision and the optimal probe state are concerned. This agreement between global and pointwise approaches has been recently showed to be a generic feature in noisy scenarios with shot-noise limited precision [21]. This is in stark contrast with the noiseless case, where the probe $\psi(y) = \cos(y\pi/2)$ is optimal for the global approach and gives $\sigma^2_{\text{opt}} = \pi^2/n^2$, while the NOON-type state $|\psi\rangle = 2^{-1/2}(|J, J\rangle + |J, -J\rangle)$ provides the optimal pointwise uncertainty $\sigma^2_{\text{opt}} = 1/n^2$.

It remains an open question to find the optimal probe state given a finite values of $\bar{S}$. As argued above, a finite *S* will only be able to moderately reshape the profile without significantly changing the scaling of its width. Therefore, we expect the optimal state to be fairly independent of the precise (finite) value of *S*, and hence very close to that obtained for the deterministic case (*S* = 1). Numerical evidence (optimizing simultaneously over probes and measurements) suggests that this is indeed the case provided *S* is not too small. With this we are lead to conjecture that the optimal probe state is given by

$$c_m^{\text{opt}} \propto \cos\left(\frac{m\pi}{n + 2} \right) e^{-\sqrt{\frac{1-r^2}{r^2 n^3}} m^2}, \tag{30}$$

independently of *S* (finite), which agrees with equation (28) for asymptotically large *n*. Note that the cosine prefactor guarantees that the solution converges to the optimal one for $r \rightarrow 1$ and it keeps the state confined in the box for all values of *n* and *r*. Such states continue to have a dominant typical value of $j = j_0$ and in those
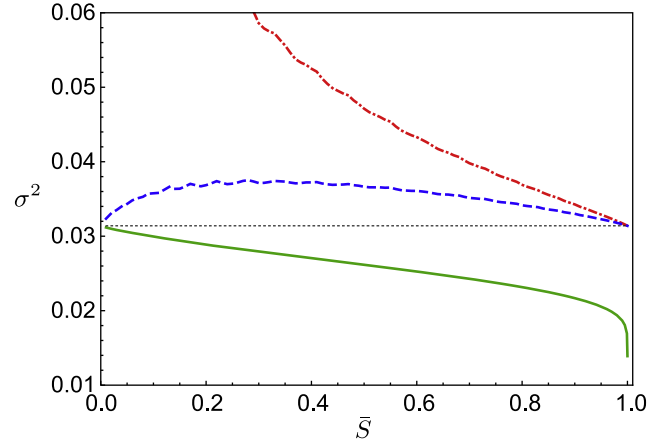
Figure 5. Scavenging information. Uncertainty σ^2 versus probability of abstention $\bar{S}$ from numerical optimization for $n = 50$ and $r = 0.8$. The green solid (red dash-dotted) correspond to σ^2_{opt} ($\bar{\sigma}^2$), where only the favorable (unfavorable) events are taken into account. The dashed curve corresponds to σ^2_{all}, where an estimate is provided on *all* outcomes, favorable or unfavorable. For low success probability ($\bar{S}$ close to unity), both $\bar{\sigma}^2$, and σ^2_{all}, approach the uncertainty of the deterministic protocol σ^2_{det} (dotted line).

blocks both the kinetic and harmonic contributions to the energy are of sub-leading order. Hence, for probes of the form in equation (30) the enhancement due to abstention is very limited, up until very high abstention probabilities where one can afford to post-select high spin states to reach the ultimate limit in equation (25).

3.7. Scavenging information from discarded events

The aim of probabilistic metrology is twofold. First, it should estimate an unknown phase θ encoded in a quantum state with a precision that exceeds the bounds of the deterministic protocols. Second, it should assess the risk of failing to provide an estimate at all (e.g., it should provide the probability of success/abstention). Probabilistic metrology protocols are hence characterized by a precision versus probability of success trade-off curve, or equivalently by $\sigma^2(S)$. As such, no attention is payed to the information on θ that might be available after an unfavorable outcome. Here, we wish to point out that one can attain $\sigma^2_{\text{opt}}(S)$ and still be able to recover, or scavenge, a fairly good estimate from the discarded outcomes (see figure 1).

The optimal scavenging protocol can be easily characterized in terms of the stochastic map $\mathscr{F}$ in equation (13), which describes the state transformation after a favorable event, and that associated to the unfavorable events:

$$\bar{\mathscr{F}}(\rho_\theta) = \bar{\Phi}\rho_\theta\bar{\Phi}, \quad \bar{\Phi} = \sum_{j,m}\bar{f}^j_m|j,m\rangle\langle j,m|\otimes\mathbb{1}_j, \tag{31}$$

where the weights $\bar{f}^j_m$ are defined through the equation $(\bar{f}^j_m)^2 = 1 - (f^j_m)^2$. The addition of the two stochastic channels, $\bar{\mathscr{F}} + \mathscr{F}$, is trace-preserving, i.e., it describes a deterministic operation, with no post-selection. The final measurement is given by the seed $\tilde{\Omega}$ defined after equation (13) for both favorable and unfavorable events. Thus, we can easily compute $\bar{\sigma}^2(S)$ for the the latter, as well as $\sigma^2_{\text{all}}(S)$, where all outcomes are included. Clearly, we must have that $\sigma^2_{\text{all}}(S) \geqslant \sigma^2_{\text{det}}$ [58], as σ^2_{det} refers to the optimal deterministic protocol.

As shown in figure 5, a protocol that is optimized for some probability of abstention $\bar{S}$, performs only slightly worse when forced to provide always a conclusive outcome. In particular, we notice that if such protocol is designed to work at the ultimate limit regime, with uncertainty σ^2_{ult}, which requires a very large abstention probability ($S \to 0$) [58], its performance coincides with that of the optimal deterministic protocol. Actually, this observation follows (see appendix I) from Winter's gentle measurement lemma (lemma 9 in [71]), which states that a measurement with a highly unlikely outcome causes only a little disturbance to the measured quantum state. This is in contrast to the claims in [58], where a random estimate is assigned to the discarded events.

4. Discussion

We have shown that abstention or post-selection can counterbalance the adverse errors in a noisy metrology task. Our results are theoretical and concern abstract systems of n qubits. However, they apply to different quantum metrology implementations, ranging from Ramsey interferometry for frequency standards [11, 12],

atomic magnetometry [5, 6], and quantum photonics (single or multi-mode setups), where the number operator introduced here will play the role of number of photons.

Post-selection is already widely used for preparing quantum information resources, e.g., single photons from weak coherent pulses, heralded down-conversion for EPR-type states, or NOON states for metrology applications. Although some degree of post-selection is common in experiments, its tailored optimized use is not fully exploited. Only recently there have been important developments in this direction in the context of weak value amplification [48, 49, 72]. We note on passing that these schemes can be considered a particular instance of our general set-up, and hence are subject to our bounds.

The optimal probabilistic measurement presented here can be understood as a filtering process selecting the total angular momentum followed by a modulating filter, and a final standard covariant phase measurement. The latter can be implemented by the (almost) optimal adaptive scheme proposed in [55, 73]. The modulation could be implemented by sequential use of amplitude-damping channels taking inspiration from recent experiments in state amplification [42, 44, 45]. In implementations that allow for an individual control of the qubits, such as ion traps, the projection onto the angular momentum basis can be efficiently carried out [74]. For implementations with less degree of control, the projection onto the fully symmetric sub-space can, as a last resort, be implemented by post-selecting outcomes with this symmetry. For instance a simple Stern Gerlach measurement could lead to outcomes ($m = J$) with a precision beyond the deterministic limits.

Regarding the implementation of our conjectured optimal probe state one can use available nonlinear N^2-type two-body interactions to turn the multi-copy gaussian profile to the wider optimal gaussian. Although, our case-study focuses on local dephasing noise, our methods can be adapted and similar, if not greater, benefits are expected for more general and implementation-specific noise models, including correlated noise.

In conclusion, we have shown what are the ultimate limits in precision reachable by any (deterministic or stochastic) quantum metrology protocol in a realistic scenario with local decoherence. We have derived the optimal bounds that can be reached when a certain rate of abstention is allowed and hence provided a full assessment of the risks and benefits of the probabilistic strategy. The benefits are clear for finite and for asymptotically large number of copies, and the precision is strictly better than that attained by deterministic strategies, which include optimal preparation of probe states. The ultimate quantum metrology scaling limit is only reached with a large abstention rate. However, in that case we have shown that it is possible to obtain estimates with standard (deterministic) precision from the discarded events. In this sense, seeking ultra-sensitive measurements is a low-risk endeavour.

Acknowledgments

Appendix A. Notation

Throughout this section we use the following notation. The n-qubit computational basis is denoted by $\{|b\rangle\}_{b=0}^{2^n-1}$, where $b = b_1 b_2 \cdots b_n$ is a binary sequence, i.e., $b_i = 0, 1$ for $i = 1, 2, \ldots, n$. We denote by $|b|$ the sum of the n digits of b, i.e., $|b| := \sum_{j=1}^n b_j$. The digit-wise sum of b and b' modulo 2 will be simply denoted by $b + b'$, hence $|b + b'|$ can be understood as the Hamming distance between b and b', both viewed as binary vectors.

The permutations of n objects, i.e., the elements of the symmetric group S_n, are denoted by π. We define the action of a permutation π over a binary list b as $\pi(b) := b_{\pi(1)} b_{\pi(2)} \cdots b_{\pi(n)}$. This induces a unitary representation of the symmetric group on the Hilbert space $\mathcal{H}^{\otimes n}$ of the n qubits through the definition $U_\pi |b\rangle := |\pi(b)\rangle$. The (fully) symmetric subspace of $\mathcal{H}^{\otimes n}$, which we denote by $\mathcal{H}_+^{\otimes n}$, plays an important role below. An orthornormal basis can be labelled $\beta = |b|$, where $\beta = 0, 1, \ldots, n$:

$$|\beta\rangle = \binom{n}{\beta}^{-1/2} \sum_{b \in B_\beta} |b\rangle \text{ with } \beta = 0, 1, \ldots n \tag{A.1}$$

where $B_\beta = \{b : |b| = \beta\}$. It is well-known that the symmetric subspace $\mathcal{H}_+^{\otimes n}$ carries the irreducible representation of spin $j = J := n/2$ of $SU(2)$. In this language, the magnetic number m is related to β by $m = n/2 - \beta$ (here we are mapping $\beta_i \to m_i = (-1)^{\beta_i}/2$ for qubit i). In other words, we map $|\beta\rangle \to |n/2, n/2 - \beta\rangle$, where we stick to the standard notation $|j, m\rangle$ for the spin angular momentum eigenstates.

We will be concerned with evolution under unitary transformations $U_\theta := u_\theta^{\otimes n}$, where $u_\theta = \exp(i\theta |1\rangle \langle 1|)$, $\theta \in (-\pi, \pi]$. The operator N such that $U_\theta = e^{i\theta N}$ will be referred to as number operator for obvious reasons:

$N|b\rangle = |b||b\rangle$. The effect of noise is taken care of by a CP map $\mathscr{D}$, so the actual evolution of an initial n-qubit state $\psi := |\psi\rangle\langle\psi|$ is $\psi \rightarrow \mathscr{D}(U_\theta \psi U_\theta^\dagger) = \rho_\theta$.

With this notation the uncertainty and success probability in equations (7) and (8) can be written as

$$\sigma^2(S) = 2 - \frac{2}{S} \max_\Omega \sum_{b,b'} \Omega_{b,b'} \rho_{b',b} \delta_{|b'|,|b|+1}, \tag{A.2}$$

$$S = \sum_{b,b'} \Omega_{b,b'} \rho_{b',b} \delta_{|b'|,|b|}, \tag{A.3}$$

where the Kronecker delta tensors result from the integration of $\hat{\theta}$.

Appendix B. Local dephasing: Hadamard channel

In this paper we consider uncorrelated dephasing noise, which can be modeled by phase-flip errors that occur with probability p_f. i.e., at the single qubit level, the effect of the noise is $\varrho \rightarrow (1 - p_f)\varrho + p_f \sigma_z \varrho \sigma_z$, where σ_z is the standard Pauli matrix $\sigma_z = \text{diag}(1, -1)$. For states of n qubits, this, so called dephasing channel $\mathscr{D}$, is most easily characterized through its action on the operator basis $\{|b\rangle\langle b'|\}_{b,b'=1}^{2^n-1}$ as

$$\mathscr{D}(|b\rangle\langle b'|) = r^{|b+b'|}|b\rangle\langle b'|, \tag{B.1}$$

where the parameter r is related to the error probability p_f through $r = 1 - 2p_f$. The effect of $\mathscr{D}$ on a general n-qubit state $\varrho = \sum_{b,b'} \varrho_{b,b'}|b\rangle\langle b'|$ can then be written as the Hadamard (or entrywise) product

$$\mathscr{D}(\varrho) = \sum_{b,b'} r^{|b+b'|}\varrho_{b,b'}|b\rangle\langle b'| := \mathcal{D} \circ \varrho, \tag{B.2}$$

where $\mathcal{D} := \sum_{b,b'} r^{|b+b'|}|b\rangle\langle b'|$ and hereafter we understand that the sums over sequences run over all possible values of b (and b') unless otherwise specified. Note that Hadamard product is basis-dependent.

Appendix C. Symmetric probes

If the probe state is fully symmetric, i.e., $|\psi\rangle \in \mathcal{H}_+^{\otimes n}$, it can be written as $|\psi\rangle = \sum_\beta \psi_\beta|\beta\rangle$, where $|\beta\rangle$ is defined in equation (A.1) and the components are related to those in equation (9) by $c_m = \psi_{J-m}$ and can be taken to be positive with no loss of generality (any phase can be absorbed in the measurement operators).

Then, $\rho = \mathscr{D}(\psi) = \mathcal{D} \circ \psi$ in equations (A.2) and (A.3) becomes

$$\rho = \sum_{\beta,\beta'} \frac{\psi_{\beta'}\psi_\beta}{\binom{n}{\beta}^{1/2}\binom{n}{\beta'}^{1/2}} \sum_{b\in B_\beta}\sum_{b'\in B_{\beta'}} r^{|b+b'|}|b'\rangle\langle b|. \tag{C.1}$$

Since ρ is permutation invariant, Ω can be chosen to be so and we can easily write equations (A.2) and (A.3) in the spin basis. We just need the non-zero Clebsch–Gordan matrix elements $\langle j, m'|b'\rangle\langle b|j, m\rangle$, where implicitly $m = J - \beta, m' = J - \beta'$. If we introduce the shorthand notation $\mathcal{D}_{m',m}^j := \langle j, m'|\mathcal{D}|j, m\rangle$, then using [75], we have

$$\mathcal{D}_{m',m}^j = \sum_{b\in B_\beta}\sum_{b'\in B_{\beta'}} r^{|b+b'|}\langle j, m'|b'\rangle\langle b|j, m\rangle$$

$$= (1 - r^2)^{J-j}r^{m-m'}\sum_k [\Delta_k^{(j)}]_m^{m'} r^{2k}, \tag{C.2}$$

where

$$[\Delta_k^{(j)}]_m^{m'} := \frac{\sqrt{(j-m)!(j+m)!(j-m')!(j+m')!}}{(j-m-k)!(j+m'-k)!(m-m'+k)!k!} \tag{C.3}$$

and the sums run over all integer values for which the factorials make sense. Recalling equation (10), a simple expression, involving just a sum over k in equation (C.3), for $\rho_{m,m'}^j = p_j^{-1}\text{tr}(|j, m\rangle\langle j, m'| \otimes \mathbb{1}_j \rho)$ follows by combining the above results. In short

$$\rho_{m',m}^j = \frac{1}{p_j} \frac{c_{m'}c_m}{\binom{n}{J-m'}^{1/2}\binom{n}{J-m}^{1/2}} \mathcal{D}_{m',m}^j, \tag{C.4}$$

where

$$p_j = \nu_j \sum_m \frac{c_m^2}{\binom{n}{J-m}} \mathcal{D}^j_{m,m} \tag{C.5}$$

and the multiplicity is given by

$$\nu_j = \binom{n}{J-j} \frac{2j+1}{J+j+1} \tag{C.6}$$

and a_m^j in equation (14) becomes

$$a_m^j = \frac{\mathcal{D}^j_{m,m+1}}{\sqrt{\mathcal{D}^j_{m,m} \mathcal{D}^j_{m+1,m+1}}}. \tag{C.7}$$

Appendix D. Relevant expressions for the multi-copy state

If the input state is of the form given in equation (22), the above expressions, equations (C.4) and (C.5), become

$$\rho^j_{m',m} = \frac{\mathcal{D}^j_{m',m}}{\sum_m \mathcal{D}^j_{m,m}} \quad \text{and} \quad p_j = \nu_j \, 2^{-n} \sum_m \mathcal{D}^j_{m,m}, \tag{D.1}$$

where

$$\sum_m \mathcal{D}^j_{m,m} = (1-r^2)^{J-j} \frac{(1+r)^{2j+1} - (1-r)^{2j+1}}{2\,r} \tag{D.2}$$

The probability to find the state in the fully symmetric subspace ($j = J$) is important when assessing the success probability of the the ultimate bounds. Since the multiplicity for the maximum spin J is equal to one, it can be readily seen that p_J scales as

$$p_J \sim e^{-n[\log 2 - \log(1+r)]}, \tag{D.3}$$

The critical probability s_j^* within a block can also be computed in the asymptotic limit $j \gg 1$ from equation (17)

$$s_j^* = \frac{\rho^j_{j,j}}{(\xi^j_j)^2} \sim e^{-2j\log(1+r)}, \tag{D.4}$$

where ξ^j_m is the gaussian ground state, with $(\xi^j_j)^2 \sim \exp(-(1-r^2)\sqrt{j/4r})$. For $m = m' = j$ equation (C.2) gives $D^j_{j,j} = (1-r^2)^{J-j}$ which together with equations (D.1) and (D.2) gives $\rho^j_{j,j} \sim \exp[2j\log(r+1)]$. This scaling dominates over that of ξ^j_m, and hence determines the scaling of s_j^*. From here we obtain critical value for the overall success probability $S^* = p_J \, s_J^* \sim e^{-n\log 2}$.

Appendix E. Ultimate bound without in-block filtering

In the results section we discuss the possibility to prepare a probe state such that after the action of noise becomes an optimal state within the fully symmetric subspace $j = J$. Here we give what its critical success probability, which only entails computing p_J.

For this purpose we first recall that the optimal filtered state ξ^j_m, defined before equation (14), has to fulfil

$$\xi^j_m = f^j_m c_m \sqrt{\frac{\nu_j \mathcal{D}^j_{m,m}}{s_j p_j \binom{n}{J-m}}}. \tag{E.1}$$

The probability of falling in the block of maximum spin J for a given filtered state ξ^j can be easily derived from equation (E.1) recalling that the probe state ψ is normalized, and thus $\sum_m c_m^2 = 1$. Solving equation (E.1) for c_m^2/p_J and summing over m we obtain

$$\frac{1}{p_J} = s_J \sum_{m=-J}^{J} \frac{\binom{n}{J-m}}{\mathcal{D}_{m,m}^J} \left(\frac{\xi_m^J}{f_m^J}\right)^2. \tag{E.2}$$

Now, for our strategy all j but the maximum one, $j = J$, are filtered out, and no further filtering is required within the block J, i.e. we have $f_m^J = 1$, for all $2J + 1$ values of m. Then $s_J = 1$ and

$$p_J = \left\{ \sum_{m=-J}^{J} \frac{\binom{n}{J-m}}{\mathcal{D}_{m,m}^J} (\xi_m^J)^2 \right\}^{-1}. \tag{E.3}$$

In the asymptotic limit the probability p_J can be estimated by noticing that the optimal distribution $(\xi_m^j)^2$ is much wider than $\binom{n}{J-m} / \mathcal{D}_{m,m}^J$ and can be replaced by $(\xi_0^J)^2$. Around $m = 0$, we can use the asymptotic formulas

$$\mathcal{D}_{m,m}^j \sim (1 - r^2)^{J-j} \frac{(1+r)^{2j+1}}{2\sqrt{\pi r j}} e^{-rm^2/j}, \tag{E.4}$$

$$\binom{n}{J-m} \sim \frac{2^n}{\sqrt{\pi J}} e^{-m^2/J}. \tag{E.5}$$

They can be derived using the Stirling approximation and saddle point techniques. Equation (E.4) also requires the EulerMaclaurin approximation to turn the sum over k in equation (C.2) into an integral that can be evaluated using again the saddle point approximation. Retaining only exponential terms, $S^* = p_J \sim (1 + r)^n / 2^n = e^{-n[\log 2 - \log(1+r)]}$.

Appendix F. Equivalence of worst-case and expected loss

Here we give a simple proof that for phase estimation, and assuming a covariant signal, such as $\rho_\theta = \mathcal{D}(U_\theta \psi U_\theta^\dagger)$, and a flat prior $\pi(\theta) = 1/(2\pi)$ (see section 3.1), the worst-case loss in equation (5)

$$L_{wc} = \sup_{\theta \in (-\pi, \pi]} \int d\hat{\theta} \, p(\hat{\theta}|\theta, \text{succ}) \ell(\theta, \hat{\theta}) \tag{F.1}$$

and the expected loss in equation (6)

$$L_s = \frac{1}{S} \iint d\hat{\theta} \, d\theta \, p(\theta, \hat{\theta}, \text{succ}) \ell(\theta, \hat{\theta}), \tag{F.2}$$

take the same value, and so do the corresponding success probabilities.

We first rewrite equation (F.2) as

$$\begin{aligned} L_s &= \frac{1}{S} \iint d\hat{\theta} \, d\theta \, p(\hat{\theta}|\text{succ}, \theta) p(\text{succ}, \theta) \ell(\theta, \hat{\theta}) \\ &= \int d\theta \, p(\theta|\text{succ}) \int d\hat{\theta} \, p(\hat{\theta}|\theta, \text{succ}) \ell(\theta, \hat{\theta}), \end{aligned} \tag{F.3}$$

where in the second equality we have used the fact that $p(\text{succ}, \theta) = p(\theta|\text{succ}) S$. When written in this form, we note the analogy between L_s in equation (F.3) and L_{wc} in equation (F.1), where the average over $p(\theta|\text{succ})$ is replaced by the supremum over θ. Because of this, we obviously have $L_s \leqslant L_{wc}$. We just need to show that the opposite inequality also holds.

We remind the reader that the minimum expected loss, L_s, can be attained by a covariant measurement, $\mathcal{M} = \{M_{\hat{\theta}} = U_{\hat{\theta}} \Omega U_{\hat{\theta}}^\dagger / (2\pi)\}_{\hat{\theta} \in (-\pi, \pi)]}$, where Ω is a suitable seed (the optimal seed). Because of covariance, we note that for any phase θ' we have $p(\hat{\theta}|\theta, \text{succ}) = p(\hat{\theta} + \Delta\theta|\theta', \text{succ})$, where $\Delta\theta = \theta' - \theta$. Likewise, we have $\ell_p(\theta, \hat{\theta}) = \ell_p(\theta', \hat{\theta} + \Delta\theta)$. By shifting variables $\hat{\theta} + \Delta\theta \to \hat{\theta}$ in equation (F.2), the integrant becomes independent of the variable θ, which can be trivially integrated to give

$$L_s = \int d\hat{\theta} \, p(\hat{\theta}|\theta', \text{succ}) \ell_p(\theta', \hat{\theta}) \tag{F.4}$$

for any θ'. It follows that the worst case loss for this particular measurement, $L_{wc}^{\mathcal{M}}$, satisfies $L_{wc}^{\mathcal{M}} = L_s$. Since $\mathcal{M}$ needs not to be the measurement that minimizes equation (F.1), we have $L_{wc} \leqslant L_{wc}^{\mathcal{M}} = L_s$. Combining this with the opposite inequality, derived after equation (F.3), we conclude that $L_s = L_{wc}$.

Proceeding along the same lines, we note that

$$S = \int \frac{d\theta}{2\pi} \int d\hat{\theta}\, p(\hat{\theta}|\theta, \text{succ}) = \int d\hat{\theta}\, p(\hat{\theta}|\theta', \text{succ}) \tag{F.5}$$

for any θ'. Hence $S = S_{wc}$.

Appendix G. The continuum limit: particle in a potential box

Proceeding as in [33, 34, 68], one can easily derive from equations (12), (14) and (15) the following equation:

$$\langle \xi^j | H^j | \xi^j \rangle = \sum_{m=-j}^{j-1} \left\{ (\xi_{m+1} - \xi_m)^2 + \frac{V_m^j}{j^2} \xi_m^2 \right\} + a_{-j} \xi_{-j}^2 + a_j \xi_j^2, \tag{G.1}$$

where $V_m^j = 2j^2(1 - a_m)$ and we have dropped the superscript j in ξ_m^j to simplify the expression. In the asymptotic limit, as j becomes very large, $m/j = x$ approaches a continuum variable that takes values in the interval $[-1, 1]$. Accordingly, the values $\{\sqrt{j}\,\xi_m\}$ approach a real function that we denote by $\varphi(x)$. With this, the former equation becomes

$$\langle \xi^j | H^j | \xi^j \rangle = \frac{1}{j^2} \int_{-1}^{1} dx \left\{ \left[\frac{d\varphi(x)}{dx} \right]^2 + V^j(x)\varphi(x)^2 \right\},$$
$$:= \frac{1}{j^2} \langle \varphi | \mathcal{H}^j | \varphi \rangle, \tag{G.2}$$

where

$$V^j(x) = 2j^2(1 - a_{xj}^j), \quad \mathcal{H}^j := -d^2/dx^2 + V^j(x) \tag{G.3}$$

and we have dropped the boundary term $[\varphi^2(-1) + \varphi^2(-1)]/j$ that stems from the second line in equation (G.1). Minimization of $\langle \xi^j | H^j | \xi^j \rangle$ require the vanishing of this term, and equation (18) readily follows. The formula

$$V^j(x) = j\frac{1 - r^2}{2r\sqrt{1 - (1 - r^2)x^2}} \tag{G.4}$$

(also in equation (19)) follows from the asymptotic expression of a_m^j, defined in equation (14). Our starting point are equations (C.7) and (C.2). The sum over k in the latter can be evaluated using the EulerMaclaurin formula and the saddle point approximation.

Appendix H. Symmetric probe is optimal and no benefit in probe-ancilla entanglement

We next show that permutation invariance enables us to choose with no loss of generality the probe state $|\psi\rangle$ from the symmetric subspace $\mathcal{H}_+^{\otimes n}$ and the seed Ω to be fully symmetric.

We first write equations (A.2) and (A.3) in a more compact form. We define Δ through the relation $\sigma^2(S) = (1 + S^{-1} \min_{\psi,\Omega} \Delta)/2$, where the maximization is performed also over the probe states since here we are concerned with the ultimate precision bound. We also introduce a slight modification of $\mathcal{D}$ that includes the Kronecker delta tensor: $\mathcal{D}_l := \sum_{b,b'} r^{|b+b'|} \delta_{|b'|,|b|+l} |b\rangle\langle b'|, l = 0,1$. Then

$$\Delta = \text{tr}[(\psi \circ \Omega)\mathcal{D}_1], \quad S = \text{tr}[(\psi \circ \Omega)\mathcal{D}_0], \tag{H.1}$$

where we have used that $\text{tr}[A(B \circ C)] = \text{tr}[(C \circ A)B]$ if $B = B^t$. The result we wish to show follows from the invariance of the noise under permutations of the n qubits, namely, from $U_\pi \mathcal{D}_l U_\pi^\dagger = \mathcal{D}_l$, for any $\pi \in S_n$, which implies that the very same value of Δ and S attained by some given measurement seed Ω and some initial state ψ, i.e., attained by $\psi \circ \Omega$, can also be attained by $U_\pi(\psi \circ \Omega)U_\pi^\dagger$, and likewise by the average $(n!)^{-1}\sum_{\pi \in S_n} U_\pi(\psi \circ \Omega)U_\pi^\dagger$.

The proof starts with yet a few more definitions: given a fully general probe state $|\psi\rangle$, we define the $n + 1$ normalized states $|\phi_\beta\rangle = \sum_{b \in B_\beta}(\psi_b/\psi_\beta)|b\rangle$, $\beta = 0, 1, \ldots, n$, where $\psi_\beta^2 = \sum_{b \in B_\beta}|\psi_b|^2$, and write $|\psi\rangle = \sum_{\beta=0}^{n} \psi_\beta|\phi_\beta\rangle$. Additionally, we define

$$|\phi\rangle = \sum_{\beta=0}^{n} \binom{n}{\beta}^{1/2} |\phi_\beta\rangle, \quad \phi = |\phi\rangle\langle\phi|. \tag{H.2}$$

We obviously have $\phi \circ \Omega \geqslant 0$, as the Hadamard product of two positive operators is also a positive operator, and $(n!)^{-1}\sum_{\pi \in S_n} U_\pi (\phi \circ \Omega) U_\pi^\dagger \geqslant 0$, as this expression is a convex combination of positive operators. Similarly, the seed condition $\mathbb{1} - \Omega \geqslant 0$ implies $(n!)^{-1}\sum_{\pi \in S_n} U_\pi [\phi \circ (\mathbb{1} - \Omega)] U_\pi^\dagger \geqslant 0$. But

$$\frac{1}{n!}\sum_{\pi \in S_n} U_\pi(\phi \circ \mathbb{1})U_\pi^\dagger = \sum_{\beta=0}^{n} \frac{\binom{n}{\beta}}{n!}\left(\sum_{\pi \in S_n} U_\pi \phi_\beta U_\pi^\dagger\right) \circ \mathbb{1}, \tag{H.3}$$

since the diagonal entries of ϕ and ϕ_β transform among themselves under permutations. The right-hand side can be written as

$$\sum_{\beta=0}^{n}\sum_{b \in B_\beta} \frac{\binom{n}{\beta}}{n!}\sum_{\pi \in S_n} \frac{|\psi_{\pi^{-1}(b)}|^2}{\psi_\beta^2}|b\rangle\langle b| = \sum_{\beta=0}^{n}\sum_{b \in B_\beta}|b\rangle\langle b| = \mathbb{1}, \tag{H.4}$$

where we have used that, for any $b \in B_\beta$, the set $\{\pi^{-1}(b)\}_{\pi \in S_n}$ contains exactly $\beta!(n-\beta)!$ times each one of the elements of B_β. It follows from equations (H.3) and (H.4) that $\Omega^{\mathrm{sym}} := (n!)^{-1}\sum_{\pi \in S_n} U_\pi (\phi \circ \Omega) U_\pi^\dagger$ satisfies $0 \leqslant \Omega^{\mathrm{sym}} \leqslant \mathbb{1}$ and is invariant under permutations of the n qubits. It is, therefore, a legitimate fully symmetric measurement seed. Moreover

$$\frac{1}{n!}\sum_{\pi \in S_n} U_\pi(\psi \circ \Omega)U_\pi^\dagger = \sum_{\beta,\beta'} \frac{\psi_\beta \psi_{\beta'}}{\binom{n}{\beta}^{1/2}\binom{n}{\beta'}^{1/2}} \mathbb{1}_\beta \Omega^{\mathrm{sym}} \mathbb{1}_{\beta'}, \tag{H.5}$$

where $\mathbb{1}_\beta$ is the projector into the subspace with $|b| = \beta$, namely $\mathbb{1}_\beta := \sum_{b \in B_\beta}|b\rangle\langle b|$. Thus, recalling the definition of $|\beta\rangle$ in equation (A.1), the right-hand side of equation (H.5) can be readily written as

$$\left(\sum_{\beta,\beta'}\psi_\beta \psi_{\beta'}|\beta\rangle\langle\beta'|\right) \circ \Omega^{\mathrm{sym}} = \psi^{\mathrm{sym}} \circ \Omega^{\mathrm{sym}}, \tag{H.6}$$

where $|\psi^{\mathrm{sym}}\rangle := \sum_{\beta=0}^{n}\psi_\beta|\beta\rangle \in \mathcal{H}_+^{\otimes n}$. It follows from these results and equation (H.1) that the very same uncertainty and success probability attained by any pair $(|\psi\rangle, \Omega)$ of probe state and measurement seed is also attained by the state $|\psi^{\mathrm{sym}}\rangle \in \mathcal{H}_+^{\otimes n}$ and the fully symmetric seed Ω^{sym}. This completes the proof.

Now that we have learned that no boost in performance can be achieved by considering probe states more general than those in the symmetric subspace $\mathcal{H}_+^{\otimes n}$ (in the subspace of maximum spin $j = J$), we may wonder if entangling the probe with some ancillary system could enhance the precision. Here we show that this possibility can be immediately ruled out, thus extending the generality of our result. For this purpose we take the general probe-ancilla state $|\Psi_{\mathrm{PA}}\rangle = \sum_b \psi_b |b\rangle |\chi_b\rangle$, where $|\chi_b\rangle$ are normalized states (not necessarily orthogonal) of the ancillary system. The action of the phase evolution and noise on the probe leads to a state of the form $\rho_{\mathrm{PA}}(\theta) = \sum_{b,b'} r^{|b+b'|} e^{i\theta(|b| - |b'|)} \psi_b \psi_{b'}|b\rangle\langle b'| \otimes |\chi_b\rangle\langle\chi_{b'}|$. This state could as well be prepared without the need of an ancillary system by taking instead an initial probe state $|\psi\rangle = \sum_b \psi_b|b\rangle$ and performing the trace-preserving completely positive map defined by $|b\rangle \rightarrow |b\rangle|\chi_b\rangle$ before implementing the measurement. This map can, of course, be interpreted as part of the measurement. It would correspond to a particular Neumark dilation of some measurement performed on the probe system alone, and hence it is included in our analysis.

Appendix I. Scavenging at the ultimate precision limit

The gentle measurement lemma [71] states that if a measurement outcome occurs with very high probability, then the corresponding conditioned state is hardly disturbed. For concreteness, let us assume that some unfavorable event in a probabilistic protocol happens with probability $\bar{S} = \mathrm{tr}[\bar{\mathscr{F}}(\rho_\theta)] = 1 - \epsilon$, then, conditioned to this event, we have $\|\rho_\theta - \bar{\rho}_\theta\|_1 \leqslant \sqrt{2}\,\epsilon$, where $\bar{\rho} = \bar{\mathscr{F}}(\rho_\theta)/\bar{S}$.

Indeed, from equation (7) or (A.2), we find that the uncertainty of the scavenged events, $\bar{\sigma}^2(S)$, rapidly approaches that of the optimal deterministic machine $\sigma^2(S = 0)$:

$$\sigma^2(0) - \bar{\sigma}^2(S) = \max_{0 \leqslant \bar{\Omega} \leqslant \mathbb{1}} \mathrm{tr}\,\bar{W}\bar{\rho}_\theta - \max_{0 \leqslant \Omega \leqslant \mathbb{1}} \mathrm{tr}\,W\rho_\theta$$
$$\leqslant \max_{0 \leqslant \Omega \leqslant \mathbb{1}} \mathrm{tr}[W(\rho_\theta - \bar{\rho}_\theta)] \leqslant \|\rho_\theta - \bar{\rho}_\theta\|_1 \leqslant \sqrt{2}\,S, \tag{I.1}$$

where W (likewise $\bar{W}$) is shorthand for the matrix with entries $W_{b,b'} = 2\Omega_{b,b'}\delta_{|b|,|b'|+1}$.

We recall that, as $\sigma^2(S)$ approaches the ultimate bound $\sigma_{\mathrm{ult}}^2 = 1 - (1 - r^2)/(4nr)$, the success probability S decreases exponentially. Equation (I.1) thus shows that such likely failure is not ruinous since in that event one can still attain the deterministic bound, i.e., one has $\bar{\sigma}^2 = \sigma^2(0) = 1 - (1 - r^2)/(4nr^2)$.

References

[1] Dowling J P and Milburn G J 2003 *Phil. Trans. R. Soc.* A **361** 1655–74

[2] Slavík R *et al* 2010 *Nat. Photon.* **4** 690–5

[3] Chen J, Habif J L, Dutton Z, Lazarus R and Guha S 2012 *Nat. Photon.* **6** 374–9

[4] Inoue K, Waks E and Yamamoto Y 2002 *Phys. Rev. Lett.* **89** 037902

[5] Budker D 2007 *Nat. Phys.* **3** 227–34

[6] Napolitano M *et al* 2011 *Nature* **471** 486–9

[7] Taylor M A *et al* 2013 *Nat. Photon.* **7** 229–33

[8] Crespi A *et al* 2012 *Appl. Phys. Lett.* **100** 233704

[9] Caves C M 1981 *Phys. Rev.* D **23** 1693–708

[10] T L S Collaboration 2011 *Nat. Phys.* **7** 962–5

[11] Bollinger J J, Itano W M, Wineland D J and Heinzen D J 1996 *Phys. Rev.* A **54** R4649–52

[12] Huelga S F *et al* 1997 *Phys. Rev. Lett.* **79** 3865–8

[13] Wineland D J, Bollinger J J, Itano W M, Moore F L and Heinzen D J 1992 *Phys. Rev.* A **46** R6797–800

[14] Kómár P *et al* 2014 *Nat. Phys.* **10** 582-7

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

[16] Giovannetti V, Lloyd S and Maccone L 2011 *Nat. Photon.* **5** 222–9

[17] Demkowicz-Dobrzański R, Kołodyński J and Guţă M 2012 *Nat. Commun.* **3** 1063

[18] Escher B M, de Matos Filho R L and Davidovich L 2011 *Nat. Phys.* **7** 406–11

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

[20] Aspachs M, Calsamiglia J, Muñoz-Tapia R and Bagan E 2009 *Phys. Rev.* A **79** 033834

[21] Jarzyna M and Demkowicz-Dobrzański R 2015 *New J. Phys.* **17** 013010

[22] Chaves R, Brask J B, Markiewicz M, Kołodyński J and Acín A 2013 *Phys. Rev. Lett.* **111** 120401

[23] Chin A W, Huelga S F and Plenio M B 2012 *Phys. Rev. Lett.* **109** 233601

[24] Jeske J, Cole J H and Huelga S 2014 *New. J. Phys.* **16** 073039

[25] Ostermann L, Ritsch H and Genes C 2013 *Phys. Rev. Lett.* **111** 123601

[26] Szankowski P, Trippenbach M and Chwedenczuk J 2014 *Phys. Rev.* A **90** 063619

[27] Boixo S, Flammia S T, Caves C M and Geremia J 2007 *Phys. Rev. Lett.* **98** 090401

[28] Macchiavello C, Huelga S F, Cirac J I, Ekert A K and Plenio M B 2000 *Quantum Communication, Computing, and Measurement 2* ed P Kumar *et al* (Dordrecht: Kluwer)

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

[30] Kessler E, Lovchinsky I, Sushkov A and Lukin M 2014 *Phys. Rev. Lett.* **112** 150802

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

[32] Fiurášek J 2006 *New J. Phys.* **8** 192

[33] Gendra B, Ronco-Bonvehi E, Calsamiglia J, Muñoz-Tapia R and Bagan E 2013 *Phys. Rev. Lett.* **110** 100501

[34] Gendra B, Ronco-Bonvehi E, Calsamiglia J, Muñoz-Tapia R and Bagan E 2013 *Phys. Rev.* A **88** 012128

[35] Marek P 2013 *Phys. Rev.* A **88** 045802

[36] Ivanovic I D 1987 *Phys. Lett.* A **123** 257
Dieks D 1988 *Phys. Lett.* A **126** 303
Peres A 1988 *Phys. Lett.* A **128** 19
Jaeger G and Shimony A 1995 *Phys. Lett.* A **197** 83

[37] Fiurášek J and Ježek M 2003 *Phys. Rev.* A **67** 012321

[38] Bagan E, Muñoz-Tapia R, Olivares-Rentería G A and Bergou J A 2012 *Phys. Rev.* A **86** 040303

[39] Duan L M and Guo G-C 1998 *Phys. Rev. Lett.* **80** 4999
Chefles A and Barnett S M 1999 *Phys. Rev.* A **60** 136

[40] Knill E, Laflamme R and Milburn G J 2001 *Nature* **409** 46–52

[41] Lütkenhaus N, Calsamiglia J and Suominen K-A 1999 *Phys. Rev.* A **59** 3295–300

[42] Ferreyrol F *et al* 2010 *Phys. Rev. Lett.* **104** 123603

[43] Xiang G Y, Higgins B L, Berry D W, Wiseman H M and Pryde G J 2011 *Nat. Photon.* **5** 43–7

[44] Zavatta A, Fiurasek J and Bellini M 2011 *Nat. Photon.* **5** 52–60

[45] Kocsis S, Xiang G Y, Ralph T C and Pryde G J 2013 *Nat. Phys.* **9** 23–8

[46] Chiribella G and Xie J 2013 *Phys. Rev. Lett.* **110** 213602

[47] Pandey S, Jiang Z, Combes J and Caves C M 2013 *Phys. Rev.* A **88** 033852

[48] Dressel J, Malik M, Miatto F M, Jordan A N and Boyd R W 2014 *Rev. Mod. Phys.* **86** 307–16

[49] Hosten O and Kwiat P 2008 *Science* **319** 787–90

[50] Brunner N and Simon C 2010 *Phys. Rev. Lett.* **105** 010405

[51] Zilberberg O, Romito A and Gefen Y 2011 *Phys. Rev. Lett.* **106** 080405

[52] Jordan A N, Martínez-Rincón J and Howell J C 2014 *Phys. Rev.* X **4** 011031

[53] Calsamiglia J 2014 *Nat. Phys.* **10** 91–2

[54] Aaronson S 2005 *Proc. R. Soc.* A **461** 3473–82

[55] Berni A A, Gehring T, Nielsen B M, HŠndchen V, Paris M and Andersen U L 2014 *Nat. Photon.* **9** 577

[56] Benítez N 2000 *Astrophys. J.* **536** 571

[57] Bertsekas D P and Tsitsiklis J N 2008 *Introduction to Probability, Athena Scientific* 2nd ednVideo Lecture 22: Bayesian Statistical Inference II, MIT OpenCourseWare, by John Tsitsiklis.

[58] Combes J, Ferrie C, Jiang Z and Caves C M 2014 *Phys. Rev.* A **89** 052117

[59] Braunstein S L and Caves C M 1994 *Phys. Rev. Lett.* **72** 3439–43

[60] Berry D W, Hall M J W, Zwierz M and Wiseman H M 2012 *Phys. Rev.* A **86** 053813

[61] Giovannetti V and Maccone L 2012 *Phys. Rev. Lett.* **108** 210404

[62] Hall M J W and Wiseman H M 2012 *Phys. Rev.* X **2** 041006

[63] Hayashi M 2011 *Commun. Math. Phys.* **304** 689–709

[64] Chiribella G, Yang Y and Yao A C-C 2013 *Nat. Commun.* **4** 2915

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

[66] Cirac J I, Ekert A K and Macchiavello C 1999 *Phys. Rev. Lett.* **82** 4344
[67] Bagan E, Ballester M A, Gill R D, Monras A and Munoz-Tapia R 2006 *Phys. Rev. A* **73** 032301
[68] Summy G S and Pegg D T 1990 *Opt. Commun.* **77** 75–9
[69] Knysh S I, Chen E H and Durkin G A 2014 arXiv:1402.0495
[70] Boyd S and Vandenberghe L 2004 *Convex Optimization* (New York: Cambridge University Press)
[71] Winter A 1999 *IEEE Trans. Inf. Theor.* **45** 2481–5
[72] Ferrie C and Combes J 2014 *Phys. Rev. Lett.* **112** 040406
[73] Berry D W and Wiseman H M 2000 *Phys. Rev. Lett.* **85** 5098–101
[74] Bacon D, Chuang I L and Harrow A W 2006 *Phys. Rev. Lett.* **97** 170502
[75] Calsamiglia J, de Vicente J I, Muñoz-Tapia R and Bagan E 2010 *Phys. Rev. Lett.* **105** 080504