Quantum scale estimation

View the article online for updates and enhancements.

Quantum scale estimation

Jesús Rubio[*]

Department of Physics and Astronomy, University of Exeter, Stocker Road, Exeter EX4 4QL, UK

(Dated: 25th October 2022)

Quantum scale estimation, as introduced and explored here, establishes the most precise framework for the estimation of scale parameters that is allowed by the laws of quantum mechanics. This addresses an important gap in quantum metrology, since current practice focuses almost exclusively on the estimation of phase and location parameters. For given prior probability and quantum state, and using Bayesian principles, a rule to construct the optimal probability-operator measurement is provided. Furthermore, the corresponding minimum mean logarithmic error is identified. This is then generalised as to accommodate the simultaneous estimation of multiple scale parameters, and a procedure to classify practical measurements into optimal, almost-optimal or sub-optimal is highlighted. As a means of illustration, the new framework is exploited to generalise scale-invariant global thermometry, as well as to address the estimation of the lifetime of an atomic state. On a more conceptual note, the optimal strategy is employed to construct an observable for scale parameters, an approach which may serve as a template for a more systematic search of quantum observables. Quantum scale estimation thus opens a new line of enquire—the precise measurement of scale parameters such as temperatures and rates—within the quantum information sciences.

I. INTRODUCTION

Quantum estimation theory is typically envisioned as a toolbox to design highly precise measurements [1–5]. Discovered over half a century ago [1, 6–13], this framework is currently attracting a great deal of attention due to its central role in 'the second quantum revolution' [14]—an application of the principles of quantum mechanics to sensing, metrology, communication, computation, nano and space technologies [15–21]. But, beyond its unquestionable value as a practical tool, estimation theory may be attributed a more prominent role within the foundations of physics: it enables the means to connect, in a clear and systematic fashion, the variables that we measure in the laboratory with the variables appearing in our theories of nature. Indeed, if x represents one of the possible outcomes that a given measurement could generate, one can often postulate the existence of a map $x \mapsto \tilde{\theta}(x)$ from the outcome x to an estimate $\tilde{\theta}(x)$ for a theoretical variable Θ [1, 22–24]. To find such map in practice, an elegant and neat approach is to minimise a mean error $\bar{\epsilon} = \langle \mathcal{D}[\tilde{\theta}(x), \theta] \rangle$ over the estimator function $\tilde{\theta}(x)$, where θ is a hypothesis about the true value of Θ, $\mathcal{D}$ gauges the deviation of $\tilde{\theta}(x)$ from θ, and the brackets indicate a weighted average over (x, θ). An instance of $\mathcal{D}$ is the familiar square error $\mathcal{D}(\tilde{\theta}, \theta) = (\tilde{\theta} - \theta)^2$ [1, 24].

Estimation-theoretic quantities can, in addition, transcend their purely statistical role to acquire a new function within physics. For example, mean errors are known to enable the derivation of generalised uncertainty relations [1, 22, 25–29] such as those for phase-number [1, 25, 30], time-energy [25, 31] and temperature-energy [32, 33] variables, and they appear to lead to a more detailed account of the notion of compatibility in quantum mechanics [4, 34–40]. Moreover, estimators provide the means to construct quantum observables such as those for phase [9, 41] and time [22] parameters. Perhaps more importantly, estimation theory offers the possibil-

ity of treating dynamical variables and parameters alike in an unified manner [1, 22]; this is demonstrated by the estimation-theoretic study of position variables in Ref. [1, Ch. 8].

In light of this fundamental outlook, selecting a deviation function $\mathcal{D}(\tilde{\theta}, \theta)$ cannot be purely a matter of choice or convenience, but should instead be informed by physical reasoning. This is apparent in phase estimation [3, 42], where the periodic nature of optical phases imposes certain constraints on which functions $\mathcal{D}$ are consistent in that context [3, 22]. The sine error $\mathcal{D}(\tilde{\theta}, \theta) = 4 \sin^2[(\tilde{\theta} - \theta)/2]$, in particular, is often employed for this purpose [1, 43–46], and while there are other functions $\mathcal{D}$ that are appropriate for phase estimation [22, 43], it is instructive to appreciate that the square error $\mathcal{D}(\tilde{\theta}, \theta) = (\tilde{\theta} - \theta)^2$ is generally not one of them.[1]

In thermometry [33, 48–57], the unknown parameter, temperature, can sometimes be seen as establishing the relevant energy scale [48, 58]. Such a scenario calls for an estimation technique that preserves scale invariance [24, 59, 60], and this idea has be shown to lead to a logarithmic deviation function $\mathcal{D}(\tilde{\theta}, \theta) = \log^2(\tilde{\theta}/\theta)$ [33]. The implications and scope of this tool within quantum thermometry are currently under analysis [55–57], including the role of adaptive measurements [55] and its connection with the notion of thermodynamic length [56]. However, the logarithmic error transcends its thermometric origin, in the sense that it is also valid for other quantities playing the role of a scale [33]. For example, to estimate Poisson rates [24, 59], kinetic parameters [61–66], and decay rates [67]. We may then ask: how can quantum mechanics enable, on the basis of logarithmic errors, the most precise estimation of scale parameters?

The present work provides a definite answer by establishing a quantum framework for the estimation scale parameters. Once Sec. II formulates the problem, Sec. III derives an analytical expression for the true minimum mean logarithmic error. Not only does this recover the optimal estim-

---

[*] J.Rubio-Jimenez@exeter.ac.uk

[1] The square error can be used as an *approximation* to the sine error when the estimator $\tilde{\theta}$ and the values for the hypothesis θ are close [3, 47].

ator found in Ref. [33], but, notably, it further provides a rule to calculate the associated probability-operator measurement (POM)[2] that is optimal for a given quantum state. Moreover, this framework can incorporate prior knowledge, as it is built on Bayesian principles [24, 60].

To illustrate these results, Sec. IV considers two applications: thermometry, and the estimation of the lifetime of an atomic state. Within the context of equilibrium thermometry [51, 69], and assuming discrete but otherwise arbitrary spectra, the question of whether energy measurements are generally optimal—i.e., for an arbitrarily large prior temperature range—is answered in the affirmative. This is not only of independent interest,[3] but it also shows that the abstract POM which is optimal for quantum scale estimation can correspond to a practical measurement strategy.

Sec. V proceeds then to explore some implications of the new toolbox. Sec. V A and V B construct a quantum observable for scale parameters, and the procedure leading to it is highlighted as a systematic method to construct other observables. Sec. V C rewrites the true minimum as an inequality statement to assess whether a given practical POM is optimal, almost-optimal or sub-optimal. Sec. V D applies the new framework to the simultaneous estimation of multiple scale parameters, identifying a lower bound on a multiparameter mean logarithmic error. Sec. V E closes with a discussion of invariance arguments in estimation theory.

To conclude, Sec. VI argues that quantum scale estimation, when taken together with phase estimation and those results based on the mean square error, completes a trio of estimation theories for three of the most elementary quantities that one could possibly measure: phases, locations and scales.

II. FORMULATION OF THE PROBLEM

Let an experiment be described by the following quantities: a measurand x, an unknown parameter Θ, and a collection of known parameters $\boldsymbol{y} = (y_1, y_2, \dots)$. Construct the vector $\boldsymbol{z} = (x, \boldsymbol{y})$. We say that Θ scales z_i—the i-th component of $\boldsymbol{z}$—if, for fixed Θ, z_i is considered 'large' when $z_i/\Theta \gg 1$ and 'small' when $z_i/\Theta \ll 1$, provided that Θ and z_i are positive and measured in the same units. The key aspect of this definition is its invariance under transformations $\{z_i \mapsto z_i' = \gamma z_i, \Theta \mapsto \Theta' = \gamma\Theta\}$, with positive γ, since $z_i/\Theta = z_i'/\Theta'$. This is the notion of *scale parameter*—referring to Θ—employed in this work. We shall then call *scale estimation* to any procedure aimed at retrieving an unknown scale parameter.

Now suppose we wish to estimate Θ, under the following assumptions: (i) knowledge of $\boldsymbol{y}$ does not inform the plausibility of the possible values for Θ; (ii) Θ scales all the paramet-

ers $\boldsymbol{y}$, but not the measurand x, which is dimensionless; and (iii) Θ is encoded within a quantum state.

To proceed, we first construct a prior probability $p(\theta|\boldsymbol{y})$ encoding what is known about Θ irrespective of the measurement process, where it is recalled that θ denotes a hypothesis about the true value of Θ. Assumption (i) allows us to take $p(\theta|\boldsymbol{y}) \mapsto p(\theta)$ [24]. The form of $p(\theta)$ is determined either by context-specific prior knowledge or, should this not be available, by imposing maximum ignorance on the basis of assumption (ii) [4, 24], which leads to Jeffreys's prior $p(\theta) \propto 1/\theta$ [59, 70].

The next step is to find a likelihood model $p(x|\theta, \boldsymbol{y})$ linking the measurand x with the unknown parameter Θ. Assumption (ii) can be enforced by imposing $p(x|\theta, \boldsymbol{y}) = p(x|\theta', \boldsymbol{y}')$. This leads to $p(x|\theta, \boldsymbol{y}) = p(x|\gamma\theta, \gamma\boldsymbol{y})$, which is a functional equation with solution $p(x|\theta, \boldsymbol{y}) = h(x, \boldsymbol{y}/\theta)$ [71, Ch. 3], for some function h. That is, the likelihood for a given measurement outcome is an arbitrary function of the ratios $(y_1/\Theta, y_2/\Theta, \dots)$.

At the same time, assumption (iii), together with the Born rule, implies $p(x|\theta, \boldsymbol{y}) = \mathrm{Tr}[M_{\boldsymbol{y}}(x)\rho_{\boldsymbol{y}}(\theta)]$, where $\rho_{\boldsymbol{y}}(\theta)$ is a density operator and $M_{\boldsymbol{y}}(x)$ is a POM. Hence, the likelihood function is

$$p(x|\theta, \boldsymbol{y}) = \mathrm{Tr}[M_{\boldsymbol{y}}(x)\rho_{\boldsymbol{y}}(\theta)] = h\left(x, \frac{\boldsymbol{y}}{\theta}\right), \qquad (1)$$

with the specific shape of h given by detailed knowledge of the experimental design. As we can see, Eq. (1) imposes a constraint on both the state and the POM, and it defines a broad class of models for which the framework of quantum scale estimation in this work applies. For a discussion of other scale estimation problems that can arise when assumptions (ii) and (iii) are relaxed, see Sec. IV A 5.

Equipped with the prior probability and the likelihood function above, we can invoke Bayes theorem [24, 60] to calculate the posterior probability

$$p(\theta|x, \boldsymbol{y}) \propto p(\theta)\mathrm{Tr}[M_{\boldsymbol{y}}(x)\rho_{\boldsymbol{y}}(\theta)] = p(\theta)\, h\left(x, \frac{\boldsymbol{y}}{\theta}\right), \quad (2)$$

which fully solves the problem of estimating Θ. Although we could stop here, it is often useful to summarise the key features of Eq. (2) via a point estimator $\tilde{\theta}_{\boldsymbol{y}}(x)$ with measurement-dependent uncertainty[4] [57, 72–74]

$$\bar{\epsilon}_{\boldsymbol{y}}(x) = \int d\theta\, p(\theta|x, \boldsymbol{y})\, \mathcal{D}[\tilde{\theta}_{\boldsymbol{y}}(x), \theta], \qquad (3)$$

for some deviation function $\mathcal{D}$. In this way, the unknown parameter Θ can, for fixed estimator and POM, be assigned a single empirical number which quantifies its value.[5] We thus need a procedure to identify which estimator best captures the essential information of Eq. (2).

---

[2] Also known as positive operator-valued measure (POVM) [68].
[3] In particular, this endows the error bound for energy measurements on a gas of spin-$1/2$ particles in Ref. [33] with a fundamental character, provided that adaptive measurements are excluded [55].

---

[4] For a discussion of this type of errors in quantum metrology, see Ref. [72].
[5] Nevertheless, note that failing to consider the full posterior density [Eq. (2)] might leave potentially valuable information unused [60].

Since such a task is to be carried out without knowledge of a specific measurement outcome, one proceeds by demanding the overall uncertainty

$$\bar{\epsilon}_{\boldsymbol{y}} = \int dx \, p(x|\boldsymbol{y}) \, \bar{\epsilon}_{\boldsymbol{y}}(x), \tag{4}$$

as opposed to using just $\bar{\epsilon}_{\boldsymbol{y}}(x)$, to be minimal [1, 3, 4, 24], where $p(x|\boldsymbol{y}) = \int d\theta \, p(\theta) p(x|\theta, \boldsymbol{y})$. This implies that the notion of 'optimal strategy' crucially depends on the chosen measure of uncertainty [75, 76], thus reinforcing the idea that deviation functions must be carefully justified.

The fact that Θ is a scale parameter leads to the logarithmic deviation function $\mathcal{D}(\tilde{\theta}, \theta) = \log^2(\tilde{\theta}/\theta)$ [33]. This is a type of relative error which is suitable for scale estimation as it satisfies the following properties: symmetry, $\mathcal{D}(\tilde{\theta}, \theta) = \mathcal{D}(\theta, \tilde{\theta})$; scale invariance, $\mathcal{D}(\gamma\tilde{\theta}, \gamma\theta) = \mathcal{D}(\tilde{\theta}, \theta)$; and having its absolute minimum at $\tilde{\theta} = \theta$, from (towards) where it grows (decreases) monotonically when $\tilde{\theta} > \theta$ ($\tilde{\theta} < \theta$). Its derivation, available in Ref. [33], will be revisited in Sec. V D while addressing the multiparameter case. By inserting $\mathcal{D}(\tilde{\theta}, \theta) = \log^2(\tilde{\theta}/\theta)$ into Eq. (4), and using the identity $p(x|\boldsymbol{y}) \, p(\theta|x, \boldsymbol{y}) = p(x, \theta|\boldsymbol{y})$, one finds the mean logarithmic error

$$\bar{\epsilon}_{\boldsymbol{y}} \mapsto \bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}} = \int dx \, d\theta \, p(x, \theta|\boldsymbol{y}) \log^2 \left[\frac{\tilde{\theta}_{\boldsymbol{y}}(x)}{\theta} \right]; \tag{5}$$

this is the uncertainty quantifier that we will use.

If an experimental platform is already built, the associated joint probability $p(x, \theta|\boldsymbol{y})$ can be fixed, so that the search for a good estimate amounts to minimising Eq. (5) with respect to $\tilde{\theta}_{\boldsymbol{y}}(x)$ [33]. However, quantum-mechanical probabilities allow for this minimisation to be performed while the experimental design is itself optimised. Specifically, using $p(x, \theta|\boldsymbol{y}) = p(\theta) \, p(x|\theta, \boldsymbol{y})$, together with the first equality in Eq. (1), we can rewrite Eq. (5) as

$$\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}} = \mathrm{Tr} \left\{ \int dx \, M_{\boldsymbol{y}}(x) \, W_{\boldsymbol{y}}[\tilde{\theta}_{\boldsymbol{y}}(x)] \right\}, \tag{6}$$

where the operator $W_{\boldsymbol{y}}[\tilde{\theta}_{\boldsymbol{y}}(x)]$ reads [1]

$$W_{\boldsymbol{y}}[\tilde{\theta}_{\boldsymbol{y}}(x)] = \int d\theta \, p(\theta) \, \rho_{\boldsymbol{y}}(\theta) \log^2 \left[\frac{\tilde{\theta}_{\boldsymbol{y}}(x)}{\theta} \right]. \tag{7}$$

Provided that $p(\theta)$ and $\rho_{\boldsymbol{y}}(\theta)$ are known, we can then search for the estimator *and* the POM that together minimise Eq. (6). In other words, we seek the solution to the optimisation problem

$$\min_{M_{\boldsymbol{y}}(x), \tilde{\theta}_{\boldsymbol{y}}(x)} \mathrm{Tr} \left\{ \int dx \, M_{\boldsymbol{y}}(x) \, W_{\boldsymbol{y}}[\tilde{\theta}_{\boldsymbol{y}}(x)] \right\} = \bar{\epsilon}_{\boldsymbol{y},\min}. \tag{8}$$

Once found, the optimal POM informs how the measurement protocol should be designed as to extract maximum information about Θ, while the associated estimator provides a rule to optimally process measurement data into an estimate.

III. QUANTUM ESTIMATION OF SCALE PARAMETERS

A. Fundamental minimum

Our first task is to calculate the minimum in Eq. (8). Start by expanding $W_{\boldsymbol{y}}[\tilde{\theta}_{\boldsymbol{y}}(x)]$ in Eq. (7) as

$$W_{\boldsymbol{y}}[\tilde{\theta}_{\boldsymbol{y}}(x)] = \varrho_{\boldsymbol{y},2} + \varrho_{\boldsymbol{y},0} \log^2 \left[\frac{\tilde{\theta}_{\boldsymbol{y}}(x)}{\theta_u} \right] - 2\varrho_{\boldsymbol{y},1} \log \left[\frac{\tilde{\theta}_{\boldsymbol{y}}(x)}{\theta_u} \right], \tag{9}$$

where the operator $\varrho_{\boldsymbol{y},k}$ is defined as

$$\varrho_{\boldsymbol{y},k} := \int d\theta \, p(\theta) \rho_{\boldsymbol{y}}(\theta) \log^k \left(\frac{\theta}{\theta_u} \right) \tag{10}$$

and θ_u is an arbitrary constant, with the same units as θ, introduced to guarantee that the argument of logarithmic functions is dimensionless [77], i.e.,

$$\log \left[\frac{\tilde{\theta}_{\boldsymbol{y}}(x)}{\theta} \right] = \log \left[\frac{\tilde{\theta}_{\boldsymbol{y}}(x)}{\theta_u} \right] - \log \left(\frac{\theta}{\theta_u} \right). \tag{11}$$

Inserting Eq. (9) into Eq. (6) further gives

$$\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}} = \int d\theta \, p(\theta) \log^2 \left(\frac{\theta}{\theta_u} \right) + \mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{A}_{\boldsymbol{y},2} - 2\varrho_{\boldsymbol{y},1}\mathcal{A}_{\boldsymbol{y},1}), \tag{12}$$

where we have introduced the operator

$$\mathcal{A}_{\boldsymbol{y},k} := \int dx \, M_{\boldsymbol{y}}(x) \log^k \left[\frac{\tilde{\theta}_{\boldsymbol{y}}(x)}{\theta_u} \right]. \tag{13}$$

Since the first term of $\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}$ in Eq. (12) is estimator- and POM-independent, we only need to minimise the trace term.

Next, define a new estimator function $\tilde{\omega}_{\boldsymbol{y}}(x)$ as

$$\tilde{\omega}_{\boldsymbol{y}}(x) := \log \left[\frac{\tilde{\theta}_{\boldsymbol{y}}(x)}{\theta_u} \right], \tag{14}$$

so that Eq. (13) becomes

$$\mathcal{A}_{\boldsymbol{y},k} = \int dx \, M_{\boldsymbol{y}}(x) \, \tilde{\omega}_{\boldsymbol{y}}(x)^k. \tag{15}$$

One then sees that

$$\mathcal{A}_{\boldsymbol{y},2} - \mathcal{A}_{\boldsymbol{y},1}^2 = \int dx \, M_{\boldsymbol{y}}(x) \, \tilde{\omega}_{\boldsymbol{y}}(x)^2 - \left[\int dx \, M_{\boldsymbol{y}}(x) \, \tilde{\omega}_{\boldsymbol{y}}(x) \right]^2 \geq 0 \tag{16}$$

due to Jensen's operator inequality [78, 79], which in turn implies the inequality

$$\mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{A}_{\boldsymbol{y},2} - 2\varrho_{\boldsymbol{y},1}\mathcal{A}_{\boldsymbol{y},1}) \geq \mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{A}_{\boldsymbol{y},1}^2 - 2\varrho_{\boldsymbol{y},1}\mathcal{A}_{\boldsymbol{y},1}) \tag{17}$$

for the trace in Eq. (12). The inequality in Eq. (16)—and thus that in Eq. (17)—is saturated when the POM is projective, i.e.,

$$M_{\boldsymbol{y}}(x) = \mathcal{P}_{\boldsymbol{y}}(x), \tag{18}$$

with $\mathcal{P}_{\boldsymbol{y}}(x)\,\mathcal{P}_{\boldsymbol{y}}(x') \mapsto \delta(x - x')\,\mathcal{P}_{\boldsymbol{y}}(x')$. Therefore, we can restrict our calculation to the set of projective measurements without loss of optimality.[6]

From these considerations we see that minimising the mean logarithmic error $\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}$ amounts to minimising the right hand side of Eq. (17). Given that both the estimator and the POM appear inside the operator $\mathcal{A}_{\boldsymbol{y},1}$, we can proceed as

$$\min_{\mathcal{A}_{\boldsymbol{y},1}} \mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{A}_{1,\boldsymbol{y}}^2 - 2\varrho_{\boldsymbol{y},1}\mathcal{A}_{\boldsymbol{y},1}) = -\mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}}^2), \tag{19}$$

where the operator $\mathcal{S}_{\boldsymbol{y}}$ is solution to the Lyaponuv equation

$$\mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0} + \varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}} = 2\varrho_{\boldsymbol{y},1} \tag{20}$$

and the minimum is achieved when $\mathcal{A}_{\boldsymbol{y},1} = \mathcal{S}_{\boldsymbol{y}}$. The details of this calculation, which is based on the calculus of variations in operator form [9, 80, 81], can be found in Appendix A.

By combining Eq. (12) with the equality in Eq. (17) and the minimum in Eq. (19), we finally arrive at the minimum mean logarithmic error

$$\bar{\epsilon}_{\boldsymbol{y},\min} = \int d\theta\, p(\theta) \log^2\left(\frac{\theta}{\theta_u}\right) - \mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}}^2). \tag{21}$$

The practical importance of this result is apparent: for given prior density $p(\theta)$ and density operator $\rho_{\boldsymbol{y}}(\theta)$, it provides the means to calculate fundamental limits to the precision in quantum scale estimation.[7]

B. Optimal quantum strategy

The next step is to identify the estimator function and the POM reaching the minimum in Eq. (21). That is, we seek the optimal quantum strategy. To find it, first recall that Eq. (21) relies on fulfilling Eq. (18) and the condition $\mathcal{A}_{\boldsymbol{y},1} = \mathcal{S}_{\boldsymbol{y}}$. Using Eq. (15) and the eigendecomposition $\mathcal{S}_{\boldsymbol{y}} = \int ds\, \mathcal{P}_{\boldsymbol{y}}(s)\, s$, these can be combined into a single condition as

$$\int dx\, \mathcal{P}_{\boldsymbol{y}}(x)\, \tilde{\omega}_{\boldsymbol{y}}(x) = \int ds\, \mathcal{P}_{\boldsymbol{y}}(s)\, s, \tag{22}$$

which leads to $\mathcal{P}_{\boldsymbol{y}}(x)\, dx = \mathcal{P}_{\boldsymbol{y}}(s)\, ds$ and $\tilde{\omega}_{\boldsymbol{y}}(x = s) = s$.

According to the first of these constraints, the optimal POM is given by the projectors associated with $\mathcal{S}_{\boldsymbol{y}}$, i.e.,

$$M_{\boldsymbol{y}}(x) = \mathcal{P}_{\boldsymbol{y}}(x) \mapsto \mathcal{M}_{\boldsymbol{y}}(s) = \mathcal{P}_{\boldsymbol{y}}(s). \tag{23}$$

---

[6] An analogous argument for the minimisation of the mean square error has been considered before [79, Appendix A].

[7] Note that since θ_u does not appear in Eq. (6), subsequent manipulations of the mean logarithmic error—including the minimum in Eq. (21)—are also independent of its specific value.

To identify the meaning of the second constraint, we need to revert the transformation in Eq. (14), i.e.,

$$\tilde{\theta}_{\boldsymbol{y}}(x) = \theta_u \exp\left[\tilde{\omega}_{\boldsymbol{y}}(x)\right]. \tag{24}$$

The optimal estimator function is then found to be

$$\tilde{\theta}_{\boldsymbol{y}}(x) \mapsto \tilde{\vartheta}_{\boldsymbol{y}}(s) = \theta_u \exp(s). \tag{25}$$

Two consistency checks are in order. First, one may confirm [Appendix B] that Eqs. (23) and (25) satisfy the conditions for the optimal quantum strategy laid out by Holevo [12, 13] and Helstrom [1, 10].

Secondly, Eq. (25) can be shown to recover the probability-based version of the optimal estimator that was derived in Ref. [33], as follows. Consider the quantity $\mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(s)\varrho_{\boldsymbol{y},1}]$ and insert $\varrho_{\boldsymbol{y},1} = (\mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0} + \varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}})/2$, which is true by virtue of Eq. (20); then

$$\begin{aligned}
\mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(s)\varrho_{\boldsymbol{y},1}] &= \mathrm{Re}\{\mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(s)\,\mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0}]\} \\
&= \int d\breve{s}\, \breve{s}\, \mathrm{Re}\{\mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(s)\mathcal{P}_{\boldsymbol{y}}(\breve{s})\varrho_{\boldsymbol{y},0}]\} \\
&= \int d\breve{s}\, \breve{s}\, \delta(s - \breve{s})\, \mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(\breve{s})\varrho_{\boldsymbol{y},0}] \\
&= s\, \mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(s)\varrho_{\boldsymbol{y},0}], \tag{26}
\end{aligned}$$

so that

$$s = \frac{\mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(s)\varrho_{\boldsymbol{y},1}]}{\mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(s)\varrho_{\boldsymbol{y},0}]}. \tag{27}$$

By introducing the expressions for $\rho_{\boldsymbol{y},0}$ and $\rho_{\boldsymbol{y},1}$ [Eq. (10)] in Eq. (27), we see that

$$\mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(s)\varrho_{\boldsymbol{y},0}] = \int d\theta\, p(\theta)\, \mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(s)\rho_{\boldsymbol{y}}(\theta)] = p(s|\boldsymbol{y}) \tag{28}$$

and

$$\begin{aligned}
s &= \frac{1}{p(s|\boldsymbol{y})} \int d\theta\, p(\theta)\, \mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(s)\rho_{\boldsymbol{y}}(\theta)] \log\left(\frac{\theta}{\theta_u}\right) \\
&= \int d\theta\, \frac{p(\theta)\, p(s|\theta,\boldsymbol{y})}{p(s|\boldsymbol{y})} \log\left(\frac{\theta}{\theta_u}\right) \\
&= \int d\theta\, p(\theta|s,\boldsymbol{y}) \log\left(\frac{\theta}{\theta_u}\right). \tag{29}
\end{aligned}$$

Finally, using Eq. (25) leads to

$$\tilde{\vartheta}_{\boldsymbol{y}}(s) = \theta_u \exp\left[\int d\theta\, p(\theta|s,\boldsymbol{y}) \log\left(\frac{\theta}{\theta_u}\right)\right], \tag{30}$$

in agreement with the result in Ref. [33].

We have thus identified a transparent and straightforward procedure—solving Eq. (20) for $\mathcal{S}_{\boldsymbol{y}}$—to construct the optimal POM [Eq. (23)] and estimator [Eqs. (25) and (30)] for any prior probability and density operator, when the error is logarithmic and squared. This, together with the minimum in Eq. (21), fully solves the optimisation problem in Eq. (8). For an alternative derivation using the method in Ref. [30], see Appendix C.

IV. EXAMPLES

Two applications of quantum scale estimation are illustrated in the following: thermometry, and the estimation of the lifetime of an atomic state.

A. Global quantum thermometry

In Ref. [33], global quantum thermometry was formulated for a given likelihood model. Within this framework, optimising the design of a thermometric experiment can only be achieved by assessing, one by one, the uncertainty associated with different likelihood functions. However, such a procedure does not guarantee that the optimal design will be found. To address this problem, one would ideally optimise the estimation error with respect to either the state, the POM, or both. In Secs. IV A 1 - IV A 4, quantum scale estimation is shown to enable such a possibility for Bayesian thermometry. Additionally, even if the likelihood model is assumed at the outset, the formalism in this work still provides a useful language to describe different thermometric protocols in a unified way; we will see this in Sec. IV A 5. The following example thus demonstrates that quantum scale estimation, when applied to thermometry, generalises the framework in Ref. [33].

1. Probability law for thermal states

Let $H = \sum_n |n\rangle\langle n| \, \varepsilon_n$ be the Hamiltonian for a quantum probe—the 'thermometer'—which is in weak thermal contact with another system whose temperature $\Theta = T$ we wish to estimate [69, 82]. When the probe is fully equilibrated, it can be described by a thermal state

$$\rho_{\boldsymbol{y}}(\theta) = \frac{\exp[-H/(k_B\theta)]}{\mathrm{Tr}\{\exp[-H/(k_B\theta)]\}} = \sum_n |n\rangle\langle n| \, h_n\left(\frac{\boldsymbol{y}}{\theta}\right), \quad (31)$$

where k_B is Boltzmann's constant,

$$h_n\left(\frac{\boldsymbol{y}}{\theta}\right) = \frac{\exp(-y_n/\theta)}{\sum_m \exp(-y_m/\theta)}, \quad (32)$$

and $\boldsymbol{y} = (\varepsilon_0, \varepsilon_1, \dots)/k_B$ is a vector of temperatures. That is, the energy spectrum gives rise to the set of known parameters.

Consider, in addition, an arbitrary POM

$$M_{\boldsymbol{y}}(x) = \sum_{nm} |n\rangle\langle m| \, M_{\boldsymbol{y},nm}(x) \quad (33)$$

where both $M_{\boldsymbol{y},nm}(x) = \langle n| M_{\boldsymbol{y}}(x) |m\rangle$ and x are dimensionless. Then, the Born rule reads

$$p(x|\boldsymbol{y},\theta) = \sum_n M_{\boldsymbol{y},nn}(x) \, h_n\left(\frac{\boldsymbol{y}}{\theta}\right). \quad (34)$$

Here, h_n is trivially invariant under transformations $\{\boldsymbol{y}' \mapsto \gamma\boldsymbol{y}, \, \theta' \mapsto \gamma\theta\}$. But the same must hold true for $M_{\boldsymbol{y},nn}$ due to dimensional consistency, since $\boldsymbol{y}$ is dimensioned. Consequently, $p(x|\boldsymbol{y},\theta) = h(x|\boldsymbol{y}/\theta)$ [Sec. II], and so the formalism for quantum scale estimation in Sec. III applies.

2. Measurement strategy

For an arbitrary prior density $p(\theta)$, we wish to find the POM that is optimal given the state in Eq. (31). The first step is to insert Eqs. (31) and (32) into Eq. (10), which gives

$$\varrho_{\boldsymbol{y},k} = \sum_n |n\rangle\langle n| \chi_{\boldsymbol{y}}^{n,k}, \quad (35)$$

with each coefficient $\chi_{\boldsymbol{y}}^{n,k}$ defined as

$$\chi_{\boldsymbol{y}}^{n,k} = \int d\theta \, p(\theta) \, \frac{\exp(-y_n/\theta) \log^k(\theta/\theta_u)}{\sum_m \exp(-y_m/\theta)}. \quad (36)$$

Next, using $\mathcal{S}_{\boldsymbol{y}} = \sum_{nm} |n\rangle\langle m| \, \mathcal{S}_{\boldsymbol{y},nm}$, where $\mathcal{S}_{\boldsymbol{y},nm} = \langle n| \mathcal{S}_{\boldsymbol{y}} |m\rangle$, and taking $\varrho_{\boldsymbol{y},0}$ and $\varrho_{\boldsymbol{y},1}$ from Eq. (35), the Lyaponuv equation (20) becomes

$$\sum_{nm} |n\rangle\langle m| \, [(\chi_{\boldsymbol{y}}^{n,0} + \chi_{\boldsymbol{y}}^{m,0}) \, \mathcal{S}_{\boldsymbol{y},nm} - 2\chi_{\boldsymbol{y}}^{n,1}\delta_{nm}] = 0. \quad (37)$$

Solving for each component $\mathcal{S}_{\boldsymbol{y},nm}$, we find

$$\mathcal{S}_{\boldsymbol{y}} = \sum_n |n\rangle\langle n| \, \frac{\chi_{\boldsymbol{y}}^{n,1}}{\chi_{\boldsymbol{y}}^{n,0}}, \quad (38)$$

which is diagonal in the energy basis $\{|n\rangle\langle n|\}$. We can further rewrite Eq. (38) in the language of continuous variables as $\mathcal{S}_{\boldsymbol{y}} = \int ds \, \mathcal{P}_{\boldsymbol{y}}(s) \, s$ by using the projective measurement

$$\mathcal{P}_{\boldsymbol{y}}(s) = \sum_n |n\rangle\langle n| \, \delta(s - s_{\boldsymbol{y},n}), \quad (39)$$

where we have defined $s_{\boldsymbol{y},n} := \chi_{\boldsymbol{y}}^{n,1}/\chi_{\boldsymbol{y}}^{n,0}$. Hence, the optimal POM is $M_{\boldsymbol{y}}(x) \mapsto \mathcal{M}_{\boldsymbol{y}}(s) = \mathcal{P}_{\boldsymbol{y}}(s)$.

Let us now examine the probability density associated with this POM. By inserting the latter into Eq. (34), and noticing that $h_n(\boldsymbol{y}/\theta) = p(n|\theta,\boldsymbol{y})$ is a discrete probability, we have

$$p(s|\theta,\boldsymbol{y}) = \sum_n \delta(s - s_{\boldsymbol{y},n}) \, p(n|\theta,\boldsymbol{y}). \quad (40)$$

Then, the optimal strategy effectively consists in measuring energy. Note that Eq. (40) is invariant under changes in the scale of temperature—since both $s_{\boldsymbol{y},n}$ and $p(n|\theta,\boldsymbol{y})$ are—in consistency with the initial premise.

From a practical perspective, this result teaches us two important lessons. First, it demonstrates that $\mathcal{S}_{\boldsymbol{y}}$ can give rise to a realisable measurement protocol—in this case, energy measurements. This is crucial as it implies that quantum scale estimation will be applicable at least in some experiments.

Secondly, this is valid 'globally' [2], in the sense that it holds for an arbitrarily large prior temperature range. In contrast, 'local' thermometry assumes—often implicitly—a prior temperature range $[\theta_{\min}, \theta_{\max}]$ such that $\theta_{\max}/\theta_{\min} \sim 1$. Performing energy measurements on the state in Eq. (31) was known to be optimal in local estimation [49, 51, 69]; our discussion shows this to hold *also* in scale-invariant global thermometry [33], which was an open question.

3. Estimate and experimental error

We next identify the optimal estimator. From Eq. (30),

$$p(s|\boldsymbol{y}) \log \left[\frac{\tilde{\vartheta}_{\boldsymbol{y}}(s)}{\theta_u} \right] = \int d\theta \, p(\theta) \, p(s|\theta, \boldsymbol{y}) \log \left(\frac{\theta}{\theta_u} \right).$$
(41)

Using Eq. (40), we see that

$$\begin{aligned} p(s|\boldsymbol{y}) &= \int d\theta \, p(\theta) \, p(s|\theta, \boldsymbol{y}) \\ &= \sum_n \delta(s - s_{\boldsymbol{y},n}) \int d\theta \, p(\theta) \, p(n|\theta, \boldsymbol{y}) \\ &= \sum_n \delta(s - s_{\boldsymbol{y},n}) \, p(n|\boldsymbol{y}). \end{aligned}$$
(42)

Inserting Eqs. (40) and (42) into Eq. (41), integrating over s, and rearranging, we get $\sum_n F_n = 0$, where

$$\begin{aligned} F_n &:= p(n|\boldsymbol{y}) \log \left[\frac{\tilde{\vartheta}_{\boldsymbol{y}}(s_{\boldsymbol{y},n})}{\theta_u} \right] \\ &\quad - \int d\theta \, p(\theta) \, p(n|\theta, \boldsymbol{y}) \log \left(\frac{\theta}{\theta_u} \right). \end{aligned}$$
(43)

Taking each addend as $F_n = 0$, and solving for $\tilde{\vartheta}_{\boldsymbol{y}}(s_{\boldsymbol{y},n})$, finally renders

$$\tilde{\vartheta}_{\boldsymbol{y}}(s_{\boldsymbol{y},n}) = \theta_u \exp \left[\int d\theta \, p(\theta|n, \boldsymbol{y}) \log \left(\frac{\theta}{\theta_u} \right) \right],$$
(44)

where $p(\theta|n, \boldsymbol{y}) \propto p(\theta) \exp(-y_n/\theta)/[\sum_m \exp(-y_m/\theta)]$. This is the optimal estimator for thermal quantum states.

To calculate the measurement-dependent error associated with this estimator, we first introduce the optimal POM $\mathcal{P}_{\boldsymbol{y}}(s)$ and the deviation function $\mathcal{D}(\tilde{\theta}, \theta) = \log^2(\tilde{\theta}/\theta)$ in Eq. (3), that is,

$$\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}(s) = \int d\theta \, p(\theta|s, \boldsymbol{y}) \log^2 \left[\frac{\tilde{\vartheta}_{\boldsymbol{y}}(s)}{\theta} \right].$$
(45)

Next, we rewrite this expression as

$$p(s|\boldsymbol{y}) \, \bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}(s) = \int d\theta \, p(\theta) \, p(s|\theta, \boldsymbol{y}) \log^2 \left[\frac{\tilde{\vartheta}_{\boldsymbol{y}}(s)}{\theta} \right].$$
(46)

We can then proceed, *mutatis mutandis*, as with Eq. (41), finding

$$\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}(s_{\boldsymbol{y},n}) = \int d\theta \, p(\theta|n, \boldsymbol{y}) \log^2 \left[\frac{\tilde{\vartheta}_{\boldsymbol{y}}(s_{\boldsymbol{y},n})}{\theta} \right].$$
(47)

By inserting Eq. (44), we finally arrive at

$$\begin{aligned} \bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}(s_{\boldsymbol{y},n}) &= \int d\theta \, p(\theta|n, \boldsymbol{y}) \log^2 \left(\frac{\theta}{\theta_u} \right) \\ &\quad - \left[\int d\theta \, p(\theta|n, \boldsymbol{y}) \log \left(\frac{\theta}{\theta_u} \right) \right]^2. \end{aligned}$$
(48)

Given a thermometric experiment described by Eq. (31)—even if only effectively—Eqs. (44) and (48) allow us to report optimal temperature estimates as

$$\tilde{\vartheta}_{\boldsymbol{y}}(s_{\boldsymbol{y},n}) \pm \Delta \tilde{\vartheta}_{\boldsymbol{y}}(s_{\boldsymbol{y},n}),$$
(49)

where $\Delta \tilde{\vartheta}_{\boldsymbol{y}}(s_{\boldsymbol{y},n}) := \tilde{\vartheta}_{\boldsymbol{y}}(s_{\boldsymbol{y},n})\sqrt{\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}(s_{\boldsymbol{y},n})}$ plays the role of an error bar [33].

4. Theoretical optimum and energy spectrum

We have so far assumed that the energy spectrum in Eq. (31) is given *a priori*. However, one may also be interested in searching for the energy spectrum which leads to the greatest precision on average. In local thermometry [69, 82], this is done by maximising the quantum Fisher information [2, 3]. Here, the task is instead to minimise Eq. (21) over $\boldsymbol{y}$. The key advantage of this approach is the possibility of including prior information—via $p(\theta)$—in an exact and controlled fashion.

By inserting Eq. (38) into Eq. (21), we have

$$\bar{\epsilon}_{\boldsymbol{y},\mathrm{min}} = \int d\theta \, p(\theta) \log^2 \left(\frac{\theta}{\theta_u} \right) - \sum_n \frac{\left(\chi_{\boldsymbol{y}}^{n,1} \right)^2}{\chi_{\boldsymbol{y}}^{n,0}}.$$
(50)

Since the first term is spectrum-independent, minimising this uncertainty amounts to maximising the quantity

$$\sum_n \frac{\left(\chi_{\boldsymbol{y}}^{n,1} \right)^2}{\chi_{\boldsymbol{y}}^{n,0}}$$
(51)

over $\boldsymbol{y}$, for given $p(\theta)$.

According to Eq. (51), the form of the optimal energy spectrum will generally depend on the prior temperature range. One may thus expect an emergence of different phases for the associated probes as the temperature range is increased—this was first noted by Mok *et al.* [53] by using a global version of the Cramér-Rao bound [83]. Given that quantum scale estimation offers a more accurate control of which prior information enters our temperature estimates, the search of optimal energy spectra for thermal states should be revisited via Eq. (51). This is left for future work.

5. Other probability laws in thermometry

Previous sections have relied on the class of probability models in Eq. (2). We have seen that Eq. (2) emerges from assumptions (i - iii) in Sec. II, and that these provide one way of formulating scale estimation at the level of quantum operators. Nevertheless, in cases where working with probability functions suffices—e.g., when both the state and the measurement have already been chosen in a given experiment—one can relax these assumptions to find other scale-invariant laws in thermometry. For the sake completeness, this is now illustrated. The functional equations in this section are solved as indicated in Ref. [71, Ch. 3].

Suppose we admit assumption (i), so that the prior probability is still constructed as in Sec. II. We also keep assumption (ii) for a single parameter y, but we drop assumption (iii). Unlike in Sec. II, here we only need to impose $p(x|\theta, y) = p(x|\gamma\theta, \gamma y)$, which leads to

$$p(x|\theta, y) = h\left(x, \frac{y}{\theta}\right), \qquad (52)$$

where h is a free function. Eq. (52) was recognised by Jaynes [59] as a scale-invariant likelihood in the context of Poisson processes.

The following thermometric models belong to the family in Eq. (52). If we have a noninteracting gas of n spin-$1/2$ particles, the probability that x of them are found in the excited state is[8] [33, 50, 55],

$$h\left(x, \frac{y}{\theta}\right) = \binom{n}{x}\frac{\exp(-xy/\theta)}{Z(y/\theta)}, \qquad (53)$$

where $y = \hbar\omega/k_B$, $\hbar$ is the reduced Planck's constant, ω is an angular frequency, and $Z(y/\theta) = [1 + \exp(-y/\theta)]^n$. Similarly, the probability that, for a quantum harmonic oscillator in thermal equilibrium, the dimensionless position coordinate x lies between x and $x + dx$ is written as [33, 84]

$$h\left(x, \frac{y}{\theta}\right) dx = \frac{\exp\{-x^2/[2\,\sigma(y/\theta)^2]\}}{\sqrt{2\pi\sigma(y/\theta)^2}}dx, \qquad (54)$$

where $y = \hbar\omega/(2k_B)$ and $2\sigma(y/\theta)^2 = \coth(y/\theta)$. Analogous considerations apply to the likelihood model for qubit thermometry in Ref. [57].

The class of models (52) can be generalised by considering a collection of dimensionless measurands $\boldsymbol{x} = (x_1, \ldots, x_n)$, so that

$$p(\boldsymbol{x}|\theta, y) = h\left(\boldsymbol{x}, \frac{y}{\theta}\right). \qquad (55)$$

Collisional thermometry, as formulated in Ref. [54], falls under this category. This is because the probability that (locally) measuring a block of n ancillas generates a particular string of zeros and ones—i.e., $x_i = 0$ or 1 for the i-th ancilla— has the form of Eq. (55), with $y = \hbar\Omega/k_B$. Here, Ω is the resonant frequency associated with both the ancillas and an auxiliary system which connects the former with the system whose temperature is to be found; see Ref. [54] for details.

Now imagine that we further drop assumption (ii). Instead, let x be a dimensioned measurand which is scaled by T, and assume that no other dimensioned parameters are involved. The condition to construct the likelihood function reads, in this case, $p(x|\theta)\,dx = p(x'|\theta')\,dx' = p(\gamma x|\gamma\theta)\gamma dx$. This leads to the functional equation $p(x|\theta) = \gamma p(\gamma x|\gamma\theta)$, whose solution is

$$p(x|\theta) = h\left(\frac{x}{\theta}\right)\frac{1}{\theta}, \qquad (56)$$

---

[8] In denoting likelihood functions, we shall omit the dependency on dimensionless parameters which, as n here, hold no relevance from the point of view of scale invariance.

for some function h. Eq. (56) formalises the notion of scale parameter traditionally employed in statistics [85]. Yet, we have seen that Eq. (56) is not the only probability type leading to scale estimation problems.

Thermal states with a continuous energy spectrum can be expressed as in Eq. (56). Indeed, the probability that the energy E given in units of temperature as $x = E/k_B$ lies between x and $x + dx$ can be written as [33, 48],

$$\begin{aligned}
p(x|\theta)\,dx &= f\left(\frac{x}{\theta}\right)\left[\int_0^\infty d\breve{x}\,f\left(\frac{\breve{x}}{\theta}\right)\right]^{-1}dx \\
&= f\left(\frac{x}{\theta}\right)\left[\theta\int_0^\infty dt\,f(t)\right]^{-1}dx \\
&= h\left(\frac{x}{\theta}\right)\frac{dx}{\theta},
\end{aligned}$$

where f is an arbitrary function,

$$h\left(\frac{x}{\theta}\right) = f\left(\frac{x}{\theta}\right)\left[\int_0^\infty dt\,f(t)\right]^{-1}, \qquad (57)$$

and we have employed the change of variables $t = \breve{x}/\theta$.

While seemingly different, all these models are invariant under transformations $\theta \mapsto \theta' = \gamma\theta$ and either $y \mapsto y' = \gamma y$ or $x \mapsto x' = \gamma x$, and so temperature can here be interpreted as a scale parameter [Sec. II]. Therefore, even when the operator-based results in Sec. III do not apply for fixed likelihoods, temperature estimation based on either of these probability functions should arguably employ the formalism of global quantum thermometry in Ref. [33], since the latter is a probability-based version of scale estimation theory.

B. Quantum estimation of a lifetime

Now we turn to a different example illustrating the potential of quantum scale estimation beyond thermometry. Consider a two-level atom initially prepared in a superposition $|\psi\rangle = \sqrt{1-a}\,|g\rangle + \sqrt{a}\,|e\rangle$, where $|g\rangle$ and $|e\rangle$ denote its ground and excited states, respectively. Due to the spontaneous emission of a photon, the excited state will decay to the ground state, a process whose statistics can be described by the density matrix [67, Ch. 4]

$$\begin{aligned}
\rho_t(\tau) = [1 - a\,\eta_t(\tau)]\,|g\rangle\langle g| + a\,\eta_t(\tau)\,|e\rangle\langle e| \\
+ [a(1-a)\,\eta_t(\tau)]^{\frac{1}{2}}\,(|g\rangle\langle e| + |e\rangle\langle g|). \qquad (58)
\end{aligned}$$

Here, $\eta_t(\tau) \coloneqq \exp(-t/\tau)$, t is the elapsed time, and τ is the lifetime. That is, the decay takes place at a rate $1/\tau$.

Imagine that the lifetime is unknown, i.e., $\Theta = \tau$, and that we wish to estimate it by observing whether or not a photon is emitted. Following Ref. [67, Ch. 4], this is captured by a POM with elements $M_{t,\tau_0}^Y = [1 - \eta_t(\tau_0)]\,|e\rangle\langle e|$ ('Yes') and $M_{t,\tau_0}^N = |g\rangle\langle g| + \eta_t(\tau_0)\,|e\rangle\langle e|$ ('No'). To use this POM, we need an initial 'hint', denoted by τ_o, at the true lifetime τ [3, 33]. By combining such a physically-motivated measurement scheme with the state in Eq. (58) via the Born rule, one arrives at the likelihood model $p(Y|\tau, \tau_0, t) = a\eta_t(\tau)[1 - \eta_t(\tau_0)]$ and

$p(N|\tau,\tau_0,t) = 1 - p(Y|\tau,\tau_0,t)$. We then see that this is a scale estimation problem, since $p(Y|\tau,\tau_0,t) \mapsto p(Y|\theta,\boldsymbol{y}) = h_Y(\boldsymbol{y}/\theta)$, with $\boldsymbol{y} = (\tau_0,t)$ and θ playing the role of a hypothesis about the true value of τ. As such, if the photon is detected, the optimal lifetime estimate is given by

$$\tilde{\vartheta}_{\boldsymbol{y}}^Y = \theta_u \exp\left[\int d\theta\, p(\theta|Y,\boldsymbol{y}) \log\left(\frac{\theta}{\theta_u}\right)\right], \qquad (59)$$

where $p(\theta|Y,\boldsymbol{y}) \propto p(\theta)\,p(Y|\theta,\boldsymbol{y})$. Otherwise, it is given by $\tilde{\vartheta}_{\boldsymbol{y}}^N$, with expression analogous to that in Eq. (59). Furthermore, the overall uncertainty is calculated as

$$\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}} = \int d\theta\, p(Y,\theta|\boldsymbol{y}) \log^2\left(\frac{\tilde{\theta}_{\boldsymbol{y}}^Y}{\theta}\right)$$

$$+ \int d\theta\, p(N,\theta|\boldsymbol{y}) \log^2\left(\frac{\tilde{\theta}_{\boldsymbol{y}}^N}{\theta}\right). \qquad (60)$$

To assess the quality of this strategy, Eq. (60) must be compared against two benchmarks: the uncertainty prior to performing the measurement, which we may denote as $\bar{\epsilon}_p$, and the fundamental limit to the precision associated with the state $\rho_t(\tau)$ in Eq. (58). The latter can be found by evaluating the quantum minimum $\bar{\epsilon}_{\boldsymbol{y},\min}$ in Eq. (21) using $\rho_t(\tau) \mapsto \rho_y(\theta)$, with $y = t$. As for the prior uncertainty, this is given as [33]

$$\bar{\epsilon}_p = \int d\theta\, p(\theta) \log^2\left(\frac{\theta}{\tilde{\vartheta}_p}\right), \qquad (61)$$

where

$$\tilde{\vartheta}_p = \theta_u \exp\left[\int d\theta\, p(\theta) \log\left(\frac{\theta}{\theta_u}\right)\right] \qquad (62)$$

is the optimal prior estimate.

For the sake of example, suppose we prepare the atom such that $a = 0.9$. Furthermore, consider the prior lifetime range $\theta/t \in [0.01, 10]$, so that the maximum-ignorance prior probability reads $p(\theta) = 0.145/\theta$. Given the dependence of the chosen POM on the initial hint τ_0, it is natural to examine how the error changes as the ratio τ_0/t takes different values within the prior range. We shall perform this calculation numerically. The evaluation of Eqs. (61), (60) and (21) lead, respectively, to the dotted purple curve in Fig. 1 for the prior uncertainty, the dashed blue line for the error associated with the physical POM, and the solid green line giving the fundamental limit to the precision for the atomic state in Eq. (58). Note that both the prior uncertainty and the fundamental limit are independent of τ_0, which is only relevant for the physical measurement, and so they appear as horizontal lines.

Fig. 1 shows that the physical POM is informative to some extent; indeed, the prior uncertainty (dotted purple) is above its error (dashed blue) for the range of τ_0/t. However, we also see that the scheme extracts more information when $\tau_0/t \ll 1$ than when $\tau_0/t \gg 1$. That is, it is easier to determine the lifetime in a regime where the decay is likely to have already happened, in consistency with our intuition. At the same time,
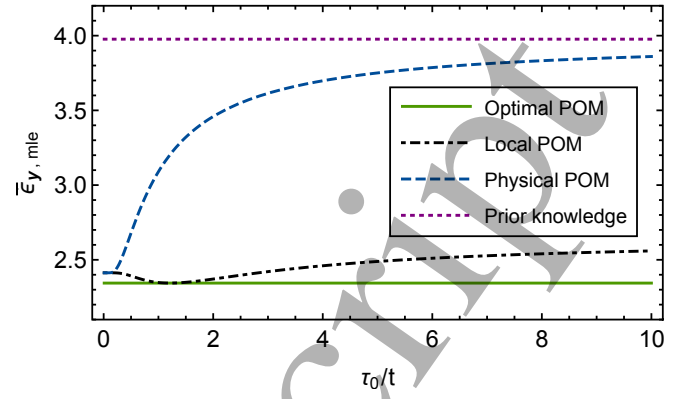

Figure 1. Quantum estimation of the lifetime of an atomic state. The mean logarithmic errors of three measurement schemes are compared: one based on the optimal POM as predicted by quantum scale estimation (solid green), one using a physical POM representing whether or not a photon is detected (dashed blue), and one based on the POM predicted by *local* quantum estimation theory [2, 3], hence referred to as 'local POM' (dash-dotted black). The latter two require an initial hint τ_0 at the true value of the unknown lifetime [3, 33], and this motivates examining the error as a function the ratio τ_0/t, where t is the time at which the measurement is performed. The dotted purple line indicates the amount of knowledge available prior to performing the measurement. As can be seen, the physical POM is informative in the sense that its associated error is below the prior uncertainty for the range of τ_0/t, and it even becomes close to the fundamental limit when $\tau_0/t \ll 1$. Yet, this is no longer the case when $\tau_0/t \gg 1$, implying that it is *not* optimal in general. The local POM leads to a substantially smaller uncertainty in this regime, but it is still above the fundamental limit. See Sec. IV B for the details of this calculation.

the error for the physical POM becomes close to the fundamental limit (solid green) when $\tau_0/t \ll 1$, implying that the physical POM is almost-optimal in this regime, but this ceases to be the case as τ_0/t grows. We thus conclude that the physical POM, while useful, is generally sub-optimal.

We have seen that the framework in this work allows us to further determine which POM reaches the fundamental limit. Specifically, by calculating the eigenstates of the operator $\mathcal{S}_y$. Solving the Lyaponuv equation (20) with the numerical values of this example leads to $\mathcal{S}_y = -1.887|g\rangle\langle g| + 0.967|e\rangle\langle e| + 0.271(|g\rangle\langle e| + |e\rangle\langle g|)$. Therefore, the optimal POM reaching the solid green line in Fig. 1 is given by $\mathcal{M}_y^+ = |\psi_+\rangle\langle\psi_+|$ and $\mathcal{M}_y^- = |\psi_-\rangle\langle\psi_-|$, where $|\psi_+\rangle = 0.094|g\rangle + 0.996|e\rangle$ and $|\psi_-\rangle = 0.996|g\rangle - 0.094|e\rangle$. This demonstrates that quantum scale estimation can inform the design of optimal metrology schemes whenever the quantity to be estimated—in this case, a lifetime—is a scale parameter.

Finally, it is natural to ask whether the fundamental limit in this section can be reached via a POM given by the more familiar framework of *local* quantum estimation theory. Specifically, by choosing the POM elements $M_{t,\tau_0}^i = |\lambda_{t,\tau_0}^i\rangle\langle\lambda_{t,\tau_0}^i|$, $i = 1, 2$, where $|\lambda_{t,\tau_0}^i\rangle$ is the i-th eigenstate of the symmetric logarithmic derivative $L_t(\tau_0)$ which is given by the Lyaponuv equation $L_t(\tau)\rho_t(\tau) + \rho_t(\tau)L_t(\tau) = 2\partial_\tau\rho_t(\tau)$ [2, 3]. Using this local POM, a numerical calculation based on

an expression analogous to that in Eq. (60) renders the black dash-dotted line in Fig. 1. As can be seen, the uncertainty associated with the local POM approaches the fundamental limit (solid green) as $\tau_0/t \ll 1$, and it actually reaches it when $\tau_0/t \sim 1$. However, it becomes sub-optimal for larger values of τ_0/t, indicating that only quantum scale estimation can provide generally optimal POMs in the presence of finite prior knowledge.

V. DISCUSSION

A. The meaning of the operator $\mathcal{S}_{\boldsymbol{y}}$

We have seen that, in quantum scale estimation, the operator $\mathcal{S}_{\boldsymbol{y}}$ is the key to find both the optimal quantum strategy and the associated minimum uncertainty. It is thus desirable to identify its physical meaning.

Motivated by the exponential map in Eq. (25), which connects the spectrum of $\mathcal{S}_{\boldsymbol{y}}$ with the optimal estimates for Θ, consider the related construction

$$
\begin{aligned}
\hat{\Theta}_{\boldsymbol{y}} &:= \theta_u \exp(\mathcal{S}_{\boldsymbol{y}}) \\
&= \theta_u \sum_{m=0}^{\infty} \frac{1}{m!} \left[\int ds\, \mathcal{P}_{\boldsymbol{y}}(s)\, s \right]^m \\
&= \theta_u \sum_{m=0}^{\infty} \frac{1}{m!} \int ds\, \mathcal{P}_{\boldsymbol{y}}(s)\, s^m \\
&= \theta_u \int ds\, \mathcal{P}_{\boldsymbol{y}}(s) \exp(s) \\
&= \int ds\, \mathcal{M}_{\boldsymbol{y}}(s)\, \tilde{\vartheta}_{\boldsymbol{y}}(s),
\end{aligned}
\tag{63}
$$

where $\{\tilde{\vartheta}_{\boldsymbol{y}}(s), \mathcal{M}_{\boldsymbol{y}}(s)\}$ denotes the estimation strategy identified as optimal in Sec. III B. We then recognise $\hat{\Theta}_{\boldsymbol{y}}$ as the optimal operator-valued estimator for the scale parameter Θ, and $\mathcal{S}_{\boldsymbol{y}}$ is simply its log-transformed version $\mathcal{S}_{\boldsymbol{y}} = \log(\hat{\Theta}_{\boldsymbol{y}}/\theta_u)$.

B. A quantum observable for scale parameters

The operator-valued estimator $\hat{\Theta}_{\boldsymbol{y}}$ plays, in addition, a more suggestive role. Let an ideal measurement process be represented by the POM $\mathcal{P}_{\boldsymbol{y}}(s)$. The operator $\mathcal{S}_{\boldsymbol{y}}$ is, in that case, a valid quantum observable [9, 22]. Since one can construct new observables by taking functions over existing ones, the fact that $\mathcal{S}_{\boldsymbol{y}}$ is an observable implies that so is $\hat{\Theta}_{\boldsymbol{y}}$. Recalling that the eigenvalues of $\hat{\Theta}_{\boldsymbol{y}}$ are the values that can be experimentally assigned to the unknown parameter Θ—indeed, they are estimates—it is then reasonable to treat $\hat{\Theta}_{\boldsymbol{y}}$ as a quantum observable for scale parameters.

This interpretation is in line with standard practice in quantum estimation theory [1, 22]. Nevertheless, the physical implications of estimation-theoretic observables are not always fully appreciated. Recall that $\mathcal{S}_{\boldsymbol{y}}$—and thus $\hat{\Theta}_{\boldsymbol{y}}$—depends on: (i) the initial state and the parameter encoding,

both captured by $\rho_{\boldsymbol{y}}(\theta)$; (ii) the prior information as represented by $p(\theta)$; and (iii) the fact that scale uncertainties are quantified using the mean logarithmic errors. Then, the scale observable $\hat{\Theta}_{\boldsymbol{y}}$ serves to model scenarios where what we can observe may depend not only on the preparation of the experimental arrangement, as indicated by $\rho_{\boldsymbol{y}}(\theta)$, but also on what information is either available to us *a priori* or simply deemed logically possible, which is what both $p(\theta)$ and $\bar{\epsilon}_{\boldsymbol{y},\text{mle}}$ allow to quantify in a precise manner. The logarithmic uncertainty $\bar{\epsilon}_{\boldsymbol{y},\text{mle}}$, in particular, is a type of noise-to-signal ratio mimicking the physics of scale parameters [33].

An appealing aspect of observables such as $\hat{\Theta}_{\boldsymbol{y}}$ is that a well-defined procedure for their construction exists. Namely, the optimisation of error functionals such as $\bar{\epsilon}_{\boldsymbol{y},\text{mle}}$. One could then take the notion of error functional as a primitive, search for the form of the deviation function which best suits the type of parameter of interest, find the estimation strategy which optimises such functional, and use it construct an observable for the aforementioned parameter. This procedure could in principle be applied to any quantity in physics, provided that one can frame it within an estimation problem.

C. Fundamental error bounds

We now focus on the fundamental minimum in Eq. (21). While knowledge of such a minimum renders a discussion of bounds on $\bar{\epsilon}_{\boldsymbol{y},\text{mle}}$ superfluous, rephrasing Eq. (21) as an inequality statement is useful to quantify the relative performance of measurements which are feasible in practice but different from the optimal POM in Eq. (23). For example, as illustrated in Sec. IV B for the estimation of an atomic lifetime. More generally, consider the chain of inequalities

$$
\bar{\epsilon}_{\boldsymbol{y},\text{mle}} \geq \min_{\tilde{\theta}_{\boldsymbol{y}}} \mathcal{E}(\tilde{\theta}_{\boldsymbol{y}}, M_{\boldsymbol{y}}) \geq \min_{M_{\boldsymbol{y}}, \tilde{\theta}_{\boldsymbol{y}}} \mathcal{E}(\tilde{\theta}_{\boldsymbol{y}}, M_{\boldsymbol{y}}),
\tag{64}
$$

where

$$
\mathcal{E}[\tilde{\theta}_{\boldsymbol{y}}(x), M_{\boldsymbol{y}}(x)] := \text{Tr}\left\{ \int dx\, M_{\boldsymbol{y}}(x)\, W_{\boldsymbol{y}}[\tilde{\theta}_{\boldsymbol{y}}(x)] \right\}.
\tag{65}
$$

From Ref. [33] we know that

$$
\min_{\tilde{\theta}_{\boldsymbol{y}}} \mathcal{E}(\tilde{\theta}_{\boldsymbol{y}}, M_{\boldsymbol{y}}) = \bar{\epsilon}_p - \mathcal{K}_{\boldsymbol{y}},
\tag{66}
$$

where $\bar{\epsilon}_p$ is given in Eq. (61),

$$
\mathcal{K}_{\boldsymbol{y}} = \int dx\, p(x|\boldsymbol{y}) \log^2 \left[\frac{\tilde{\vartheta}_{\boldsymbol{y}}(x)}{\tilde{\vartheta}_p} \right]
\tag{67}
$$

represents the information gained via the (not necessarily optimal) measurement $M_{\boldsymbol{y}}(x)$, $\tilde{\vartheta}_p$ and $\tilde{\vartheta}_{\boldsymbol{y}}(x)$ are given in Eqs. (62) and Eq. (30), respectively.

Similarly, the minimum uncertainty found in this work,

Eq. (21), can be expressed as

$$\bar{\epsilon}_{\boldsymbol{y},\min} = \int d\theta\, p(\theta) \log^2\left(\frac{\theta}{\theta_u}\right) - \mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}}^2)$$

$$= \int d\theta\, p(\theta) \log^2\left(\frac{\theta}{\theta_u}\right) - \mathrm{Tr}\left[\varrho_{\boldsymbol{y},0} \log^2\left(\frac{\hat{\Theta}_{\boldsymbol{y}}}{\theta_u}\right)\right]$$

$$= \bar{\epsilon}_p - \mathrm{Tr}\left[\varrho_{\boldsymbol{y},0} \log^2\left(\frac{\hat{\Theta}_{\boldsymbol{y}}}{\tilde{\vartheta}_p}\right)\right] := \bar{\epsilon}_p - \mathcal{J}_{\boldsymbol{y}}; \qquad (68)$$

this stems from the operator-valued estimator $\hat{\Theta}_{\boldsymbol{y}}$ in Eq. (63) plus choosing, without loss of generality, $\theta_u = \tilde{\vartheta}_p$.[9] Hence,

$$\min_{M_{\boldsymbol{y}},\,\tilde{\theta}_{\boldsymbol{y}}} \mathcal{E}(\tilde{\theta}_{\boldsymbol{y}}, M_{\boldsymbol{y}}) = \bar{\epsilon}_p - \mathcal{J}_{\boldsymbol{y}}. \qquad (69)$$

Given Eqs. (66) and (69), the chain of inequalities in Eq. (64) can finally be written as

$$\bar{\epsilon}_{\mathrm{mle}} \geq \bar{\epsilon}_p - \mathcal{K}_{\boldsymbol{y}} \geq \bar{\epsilon}_p - \mathcal{J}_{\boldsymbol{y}}. \qquad (70)$$

The first inequality is saturated by processing the outcome x with the optimal estimator $\tilde{\vartheta}_{\boldsymbol{y}}(x)$, and in principle there is no reason to choose any other estimator.[10] The real advantage of Eq. (70) lies in its second inequality: given $p(\theta)$, $\rho_{\boldsymbol{y}}(\theta)$ and a practical POM $M_{\boldsymbol{y}}(x)$, we can determine whether $M_{\boldsymbol{y}}(x)$ is optimal, almost-optimal [43] or manifestly sub-optimal by evaluating and comparing $\mathcal{K}_{\boldsymbol{y}}$ and $\mathcal{J}_{\boldsymbol{y}}$. Importantly, such a procedure offers a computationally similar but more widely applicable alternative to the usual practice of comparing the classical Fisher information with its quantum counterpart.

It must be noted that this comparative study, while practical in scope, is generally theoretical in nature. This is because $\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}$ does not depend on a specific measurement outcome [Sec. II], and so results based on it are to be understood as general statements about the class of schemes under analysis, rather than taking $\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}$ as an error that must be measured in the laboratory. $\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}$ can of course be 'measured' if *all* the probabilities involved in its construction happen to correspond to known distributions of relative frequencies; Li *et al.* [86] examines this for errors of the same nature as $\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}$. However, by no means needs the role of $\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}$ be restricted in such a way, neither does an empirical $\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}$ negate the necessity of reporting data processing errors via Eq. (3) [24].

D. Multiparameter schemes

This work has focused on single-parameter estimation, but realistic sensing schemes typically involve several unknown

---

[9] Note that choosing $\theta_u = \tilde{\vartheta}_p$ for the expression of the minimum is consistent with the definition of $\tilde{\vartheta}_p$. Indeed, inserting $\theta_u = \tilde{\vartheta}_p$ into Eq. (62) gives $0 = \int d\theta\, p(\theta) \log(\theta/\tilde{\vartheta}_p) = -\log(\tilde{\vartheta}_p/\check{\theta}_u) + \int d\theta\, p(\theta) \log(\theta/\check{\theta}_u)$, with arbitrary $\check{\theta}_u$, from where Eq. (62) reemerges.

[10] Admittedly, there may be cases where the numerical computation of $\tilde{\vartheta}_{\boldsymbol{y}}(x)$ is challenging. If so, then it might be reasonable to work with a simpler estimator, provided that we also assess how close to the first bound in Eq. (70) the associated uncertainty $\bar{\epsilon}_{\mathrm{mle}}$ is.

quantities [4, 87]. In our context, this type of scenario is illustrated, for example, by the family of probability models

$$p(x|\boldsymbol{\theta},\boldsymbol{y}) = \mathrm{Tr}[M_{\boldsymbol{y}}(x)\rho_{\boldsymbol{y}}(\boldsymbol{\theta})] = h\left(x, \frac{y_1}{\theta_1}, \ldots, \frac{y_d}{\theta_d}\right), \quad (71)$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)$ denote hypotheses for the unknown parameters $\boldsymbol{\Theta} = (\Theta_1, \ldots, \Theta_d)$, and each y_i is scaled by Θ_i. While a general formulation is left for future work, it is interesting to explore—using Eq. (71)—a first approach to multiparameter scale estimation. As in Sec. II, we assume that knowledge of $\boldsymbol{y}$ does not inform the plausibility of the possible values for $\boldsymbol{\Theta}$.

First, one should arguably be sceptical about the existence of relationships between different parameters unless there is a reason (i.e., further information) to think otherwise. If we accept this viewpoint, any form of maximum ignorance in multiparameter estimation may be represented by a separable prior probability, i.e., $p(\boldsymbol{\theta}) = \prod_i p(\theta_i)$.

We further see that, as per Eq. (71), maximum ignorance about each parameter amounts to imposing invariance under transformations $y_i \mapsto y_i' = \gamma_i y_i$ and $\theta_i \mapsto \theta_i' = \gamma_i \theta_i$ [59, 60], with $0 < \gamma_i < \infty$. Then, the condition $p(\boldsymbol{\theta})d\boldsymbol{\theta} = p(\boldsymbol{\theta}')d\boldsymbol{\theta}'$ leads to the functional equation $\prod_i p(\theta_i) = \prod_i \gamma_i\, p(\gamma_i \theta_i)$. Its solution, $p(\boldsymbol{\theta}) \propto \prod_i (1/\theta_i)$ [Appendix D], is simply the product of Jeffreys's prior [70] for each parameter.

Next, the transformation $\phi_i = \alpha_i \log(\theta_i/\theta_{u,i})$, with constant α_i and $\theta_{u,i}$, connects the estimation of a set of location parameters $\boldsymbol{\Phi} = (\Phi_1, \ldots, \Phi_d)$, for which maximum ignorance is represented as $p(\boldsymbol{\phi}) = \prod_i p(\phi_i) \propto 1$, to our multiparameter scale estimation problem, in the sense that $p(\boldsymbol{\phi})d\boldsymbol{\phi} = p(\boldsymbol{\theta})d\boldsymbol{\theta}$ implies $p(\boldsymbol{\phi}) \propto 1 \mapsto p(\boldsymbol{\theta}) \propto \prod_i(1/\theta_i)$. Then, given a figure of merit for $\boldsymbol{\Phi}$, we may construct the figure of merit for $\boldsymbol{\Theta}$ by transforming the former.

For a single location parameter ϕ_i, the correct deviation function is the translation-invariant distance $|\tilde{\phi}_i(x) - \phi_i|^{k_i}$ [24, 33]. For multiple location parameters, current practice considers the sum of distances $\mathcal{D}[\tilde{\boldsymbol{\phi}}(x), \boldsymbol{\phi}] = \sum_{i=1}^d |\tilde{\phi}_i(x) - \phi_i|^{k_i}/d$ [88, 89], where the $1/d$ factor indicates that all of them are taken to be equally important [89, 90]. By transforming $\mathcal{D}(\tilde{\boldsymbol{\phi}}, \boldsymbol{\phi})$ we thus find

$$\mathcal{D}(\tilde{\boldsymbol{\phi}}, \boldsymbol{\phi}) \mapsto \mathcal{D}(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \frac{1}{d}\sum_i \left|\alpha_i \log\left(\frac{\tilde{\theta}_i}{\theta_i}\right)\right|^{k_i}, \qquad (72)$$

which is scale invariant. To recover a sum of noise-to-signal ratios in the limit of local prior knowledge, we take $\alpha_i = 1$ and $k_i = 2$, for all i [33].

Having found a deviation function for several scale parameters, we can now use it to upgrade the notion of mean logarithmic error as

$$\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}} = \frac{1}{d}\sum_i \bar{\epsilon}_{\boldsymbol{y},\mathrm{mle},i}, \qquad (73)$$

where

$$\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle},i} = \int dx\, d\boldsymbol{\theta}\, p(x, \boldsymbol{\theta}|\boldsymbol{y}) \log^2\left[\frac{\tilde{\theta}_{\boldsymbol{y},i}(x)}{\theta_i}\right] \qquad (74)$$

and $p(x, \boldsymbol{\theta} | \boldsymbol{y}) = p(\boldsymbol{\theta}) \text{Tr}[M_{\boldsymbol{y}}(x)\rho_{\boldsymbol{y}}(\boldsymbol{\theta})]$. Note that this is consistent with the definition in Ref. [33], as the new $\bar{\epsilon}_{\boldsymbol{y},\text{mle}}$ reduces to Eq. (5) for $d = 1$.

An immediate consequence of the new definition for $\bar{\epsilon}_{\boldsymbol{y},\text{mle}}$ is that we can lower bound this error by finding the minimum value of each individual error $\bar{\epsilon}_{\boldsymbol{y},\text{mle},i}$, since $\bar{\epsilon}_{\boldsymbol{y},\text{mle},i} \geq 0$. Start by upgrading the definitions in Eqs. (10) and (13) to

$$\varrho_{\boldsymbol{y},k,i} = \int d\boldsymbol{\theta} \, p(\boldsymbol{\theta}) \rho_{\boldsymbol{y}}(\boldsymbol{\theta}) \log^k \left(\frac{\theta_i}{\theta_{u,i}} \right) \qquad (75)$$

and

$$\mathcal{A}_{\boldsymbol{y},k,i} = \int dx \, M_{\boldsymbol{y}}(x) \log^k \left[\frac{\tilde{\theta}_{\boldsymbol{y},i}(x)}{\theta_{u,i}} \right], \qquad (76)$$

respectively. Note that $\varrho_{\boldsymbol{y},0,i} = \int d\boldsymbol{\theta} \, p(\boldsymbol{\theta}) \rho_{\boldsymbol{y}}(\boldsymbol{\theta})$ for all i. Each error $\bar{\epsilon}_{\boldsymbol{y},\text{mle},i}$ can then be rewritten as

$$\bar{\epsilon}_{\boldsymbol{y},\text{mle},i} = \int d\boldsymbol{\theta} \, p(\boldsymbol{\theta}) \log^2 \left(\frac{\theta_i}{\theta_{u,i}} \right) + \text{Tr}(\rho_{\boldsymbol{y},0,i} \mathcal{A}_{\boldsymbol{y},2,i} - 2\rho_{\boldsymbol{y},1,i} \mathcal{A}_{\boldsymbol{y},1,i}). \qquad (77)$$

We see that minimising Eq. (77) with respect to $\tilde{\theta}_{\boldsymbol{y},i}(x)$ and $M_{\boldsymbol{y}}(x)$ is formally identical to the minimisation performed in Sec. III A. Consequently, $\bar{\epsilon}_{\boldsymbol{y},\text{mle}}$ is lower-bounded as

$$\bar{\epsilon}_{\boldsymbol{y},\text{mle}} \geq \frac{1}{d} \sum_i \left[\int d\boldsymbol{\theta} \, p(\boldsymbol{\theta}) \log^2 \left(\frac{\theta_i}{\theta_{u,i}} \right) - \text{Tr}(\rho_{\boldsymbol{y},0,i} \mathcal{S}_{\boldsymbol{y},i}^2) \right], \qquad (78)$$

where each operator $\mathcal{S}_{\boldsymbol{y},i}$ solves the associated Lyapunov equation $\mathcal{S}_{\boldsymbol{y},i}\rho_{\boldsymbol{y},0,i} + \rho_{\boldsymbol{y},0,i}\mathcal{S}_{\boldsymbol{y},i} = 2\rho_{\boldsymbol{y},1,i}$.

Eq. (78) is a useful starting point to assess the overall uncertainty in the estimation of multiple scale parameters. An analogous result for the mean square error has in fact been proven informative in sensing networks and imaging [30, 91]. Yet, one can anticipate the inability to saturate the bound in Eq. (78) whenever $[\mathcal{S}_{\boldsymbol{y},i}, \mathcal{S}_{\boldsymbol{y},j}] \neq 0$ for $i \neq j$ [4, 30]. How is this incompatibility to be assessed and quantified is precisely the focus of current efforts in multiparameter quantum estimation [4, 34–39], as this is expected to shed new light on our understanding of the quantum-to-classical transition [4]. In that sense, bounds such as Eq. (78) and that in Ref. [30] may help to frame the question of quantum compatibility such that its dependency on the prior information and the measure of uncertainty—which has been noted before [92]—is manifest.

E.   The role of invariance arguments

We close on a conceptual note. In this work, probability theory is understood as an extension of the propositional calculus [93–95] capable of relating a set of propositions with the evidence that supports them [96]. This approach, which we may call *objective inference*, is one of the so-called Bayesian varieties of probability theory [24, 96]. Within physics, this view gives rise to a notion of consistency such that two observers holding the same information must assign the same

probability [24, 59]. In turn, this leads to constraints that our probability functions must satisfy [59, 60].

Invariance arguments [22, 97–99] offer one way of finding these constraints. The estimation problem of interest is first enunciated as done at the beginning of Sec. II. The set of propositions involved in this step constitute our *state of information* [24, 95]. The task is then to identify which parameter transformations leave this state of information unaltered. Once a relevant symmetry is identified—in our case, scale invariance—it can be used to construct, or at least constrain, the required probabilities. This is the origin of the functional equations in Secs. II, IV A 5 and V D.

If a likelihood model is available, its functional form may reveal which parameter transformations are important. These can then be used to calculate ignorance priors [59, 60]. Such a procedure—which has been exploited in temperature estimation [33, 48] as well as in Sec. V D—is often interpreted as if invariance arguments could only be made with reference to a pre-existing likelihood model [98, Ch. 10]. Yet, Sec. II challenges this view by defining scale invariance *before* a likelihood model—and even a quantum state—has been specified. In general, our state of information, and thus its associated symmetries, precedes the probability functions employed to represent it [24, Ch. 12].

States of information may sometimes satisfy no relevant symmetry. In that case, all potential reparametrisations—i.e., estimating some function $f(\Theta)$ rather than Θ—could be seen as equally legitimate, motivating a demand for invariance under *all* reparametrisations. That is, a principle of indifference [24]. This lack of symmetry might emerge, for instance, in temperature estimation with some non-equilibrium states; the generally-invariant version of global thermometry put forward by Jørgensen *et al.* [56] could provide a suitable framework there. Notwithstanding this, note that enforcing such a strong requirement on thermometric protocols which explicitly display simple symmetries—e.g., scale invariance in protocols based on Eq. (31)—would overconstrain our estimates, and this should be avoided for the sake of economy.

It is also noted that the desideratum of invariance under all reparametrisations further leads to Jeffreys's general rule, upon which the prior probability $p(\theta)$ is taken as proportional to the square root of the Fisher information for a likelihood model [97, 100]. Using this rule has been suggested, for example, as a way to address 'uninformed' thermometry [57]. While superficially appealing, an immediate difficulty with this rule is that it violates the so-called likelihood principle [24, 97]. Even if this is ignored, or bypassed by using the *quantum* Fisher information—which only depends on the density operator [3]—Jeffreys's general rule still has a strong and potentially unrealistic implication: that we know, *a priori*, that the most likely values for the unknown parameter happen to be those for which the experimental setup is more sensitive. This may be the case for exceedingly well-calibrated protocols, which *are* realisable as per having abundant measurement data [2, 3]. Yet, for limited-data protocols [47, 73, 74], the aforementioned approach of analysing which specific symmetries are at play [59, 97], so that $p(\theta)$ is based on minimal assumptions, appears as a safer path to calculate

truly uninformative priors.

The reasoning in this section can be challenged by switching to a different system of probability; the alternatives include subjective inference [96, 101, 102], pure frequency interpretations [96], and the abstract random-variable formulation [103]. Yet, objective inference offers two advantages that together justify having advocated it in this work.

First, not only does objective inference accommodate probabilities without a frequency interpretation, but it also recovers the usual correspondence with statistical frequencies whenever large numbers limits hold [24, 103]. Hence, use of this system of probability renders the widespread 'Bayesian vs frequentist' divide unnecessary, both practically and conceptually. Secondly, the logical—in the sense of taking propositions as a primitive—and impersonal language associated with this approach is arguably the most appropriate to search for mathematical principles that, on the basis of empirical data, account for natural phenomena and their laws—the aim of physics.

VI. CONCLUDING REMARKS: PHASES, LOCATIONS AND SCALES

This work has addressed an important gap in quantum metrology. Namely, current efforts gravitate around the estimation of phase and location parameters [3, 4], using either periodic or square errors, but these are not appropriate when dealing with scale parameters [33]. In contrast, the framework of quantum scale estimation allows for the most precise measurement of scale parameters—as defined in Sec. II—in a consistent manner, by following the optimal strategy in Eqs. (23) and (25). Moreover, the minimum in Eq. (21) enables the search of fundamental limits to this precision, while the inequalities in Eq. (70) give the means to assess how close to these limits can practical POMs be.

Since this framework has been formulated in a single-shot fashion, the next natural step is to consider multi-shot protocols and explore their asymptotic limits. Some results in this direction have been reported in Refs. [33, 55] within the context of thermometry. In Ref. [33], for example, the mean logarithmic error was shown to recover the quantum noise-to-signal ratio [2] in the limit of many measurement trials, establishing a connection with the classical Fisher information. One would thus expect a link between the new framework and the *quantum* Fisher information to exist. On a related note, it would be interesting to explore the potential connection between scale estimation as formulated here and the minimax approach [104].

The most immediate use of the results in this work is to bridge the gap between the Bayesian approach to quantum thermometry currently under development [33, 54–57] and the full power of quantum estimation theory. To that purpose, the study of thermal states in Sec. IV A, together with the discussion of invariance principles in Sec. V E, provides a starting point. This could help, for instance, to clarify the link between Bayesian thermometry and other approaches to temperature estimation [105–109] that may rely on non-equilibrium states.

Nonetheless, we have seen that quantum scale estimation is not restricted to thermometry. Namely, Sec. IV B has shown how the framework can be exploited to improve the measurement of the lifetime of an atomic state [67]. Furthermore, the multiparameter extension in Sec. V D may prove to be useful for the simultaneous estimation of multiple kinetic parameters—key in the biosciences [61–64]—when these are set to be measured using quantum states of light [65, 66].

Beyond its applicability, quantum scale estimation offers a wider theoretical perspective. When the focus is mainly—and perhaps inadvertently—placed on square errors, one is tempted to believe that results based on $\mathcal{D}(\tilde{\theta}, \theta) = (\tilde{\theta} - \theta)^2$ are applicable to any possible parameter of interest. Although consideration of phases alleviates this to some extent, the approximation $\mathcal{D}(\tilde{\theta}, \theta) = 4\sin^2[(\tilde{\theta} - \theta)/2] \approx (\tilde{\theta} - \theta)^2$ [3] encourages the idea that square errors may still suffice in a somehow wide range of cases. The need for relative errors, as is $\mathcal{D}(\tilde{\theta}, \theta) = \log^2(\tilde{\theta}/\theta)$ [33, 56], manifestly reduces such range, and it suggests a more attractive viewpoint: that three of the most elementary concepts in physics—phase, location and scale—should each have a dedicated quantum estimation theory, in accordance with Tab. I.

Appendix A: Minimisation of $\mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{A}_{\boldsymbol{y},1}^2 - 2\varrho_{\boldsymbol{y},1}\mathcal{A}_{\boldsymbol{y},1})$ over $\mathcal{A}_{\boldsymbol{y},1}$

This appendix uses the calculus of variations in operator form to find

$$\min_{\mathcal{A}_{\boldsymbol{y},1}} f(\mathcal{A}_{\boldsymbol{y},1}), \tag{A1}$$

where $f(\mathcal{A}_{\boldsymbol{y},1}) = \mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{A}_{\boldsymbol{y},1}^2 - 2\varrho_{\boldsymbol{y},1}\mathcal{A}_{\boldsymbol{y},1})$.

The condition for a given $\mathcal{A}_{\boldsymbol{y},1}$ to give rise to an extremum is [9, 80]

$$\left.\frac{df(\mathcal{A}_{\boldsymbol{y},1} + \alpha\Gamma)}{d\alpha}\right|_{\alpha=0} = 0, \tag{A2}$$

where Γ denotes a Hermitian but otherwise arbitrary operator variation. Given that the first variation of $f(\mathcal{A}_{\boldsymbol{y},1})$ reads

$$\left.\frac{df(\mathcal{A}_{\boldsymbol{y},1} + \alpha\Gamma)}{d\alpha}\right|_{\alpha=0} = \mathrm{Tr}[(\mathcal{A}_{\boldsymbol{y},1}\varrho_{\boldsymbol{y},0} + \varrho_{\boldsymbol{y},0}\mathcal{A}_{\boldsymbol{y},1} - 2\varrho_{\boldsymbol{y},1})\Gamma], \tag{A3}$$

Type of parameter	phase	location	scale
General support	$0 \leq \theta < 2\pi$	$-\infty < \theta < \infty$	$0 < \theta < \infty$
Symmetry	$\theta \mapsto \theta' = \theta + 2\gamma\pi,\ \gamma \in \mathbb{Z}$	$\theta \mapsto \theta' = \theta + \gamma,\ \gamma \in \mathbb{R}$	$\theta \mapsto \theta' = \gamma\theta,\ \gamma \in \mathbb{R}_{++}$
Maximum ignorance	$p(\theta) = 1/2\pi$	$p(\theta) \propto 1$	$p(\theta) \propto 1/\theta$
Deviation function $\mathcal{D}(\tilde{\theta}, \theta)$	$4\sin^2[(\tilde{\theta} - \theta)/2]$	$(\tilde{\theta} - \theta)^2$	$\log^2(\tilde{\theta}/\theta)$

Table I. A basic prescription to develop quantum estimation theories for three of the most elementary notions in physics: phase, location and scale. The justification may be found, e.g., in Refs. [1, 3, 22] for phases, Refs. [24, 97] for locations, and Refs. [33, 48, 59] for scales.

Eq. (A2) amounts to impose

$$\mathrm{Tr}[(\mathcal{A}_{\boldsymbol{y},1}\varrho_{\boldsymbol{y},0} + \varrho_{\boldsymbol{y},0}\mathcal{A}_{\boldsymbol{y},1} - 2\varrho_{\boldsymbol{y},1})\Gamma] = 0 \quad\text{(A4)}$$

for all Γ. This is satisfied when

$$\mathcal{A}_{\boldsymbol{y},1}\varrho_{\boldsymbol{y},0} + \varrho_{\boldsymbol{y},0}\mathcal{A}_{\boldsymbol{y},1} = 2\varrho_{\boldsymbol{y},1}, \quad\text{(A5)}$$

which is a Lyaponuv equation for $\mathcal{A}_{\boldsymbol{y},1}$. For the sake of clarity, we may denote the operator $\mathcal{A}_{\boldsymbol{y},1}$ that is solution to such equation by $\mathcal{S}_{\boldsymbol{y}}$, i.e.,

$$\mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0} + \varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}} = 2\varrho_{\boldsymbol{y},1}. \quad\text{(A6)}$$

So far we have that $\mathcal{A}_{\boldsymbol{y},1} = \mathcal{S}$ makes the quantity $f(\mathcal{A}_{\boldsymbol{y},1})$ extremal. Given that

$$\left.\frac{d^2 f(\mathcal{A}_{\boldsymbol{y},1} + \alpha\Gamma)}{d\alpha^2}\right|_{\alpha=0} = 2\,\mathrm{Tr}(\varrho_{\boldsymbol{y},0}\Gamma^2) > 0, \quad\text{(A7)}$$

we can conclude that, in fact, $\mathcal{A}_{\boldsymbol{y},1} = \mathcal{S}_{\boldsymbol{y}}$ gives rise to a minimum. To see the validity of Eq. (A7), first consider the eigendecomposition $\Gamma = \int d\gamma\, \mathcal{P}(\gamma)\,\gamma$. In that case,

$$\begin{aligned}
\mathrm{Tr}(\varrho_{\boldsymbol{y},0}\Gamma^2) &= \int d\gamma d\breve{\gamma}\, \mathrm{Tr}[\varrho_{\boldsymbol{y},0}\mathcal{P}(\gamma)\mathcal{P}(\breve{\gamma})]\,\gamma\,\breve{\gamma} \\
&= \int d\gamma d\breve{\gamma}\, \delta(\gamma - \breve{\gamma})\, \mathrm{Tr}[\mathcal{P}(\breve{\gamma})\varrho_{\boldsymbol{y},0}]\,\gamma\,\breve{\gamma} \\
&= \int d\gamma\, \mathrm{Tr}[\mathcal{P}(\gamma)\varrho_{\boldsymbol{y},0}]\,\gamma^2 \\
&= \int d\gamma\, p(\gamma|\boldsymbol{y})\,\gamma^2, \quad\text{(A8)}
\end{aligned}$$

so that Eq. (A7) holds. The last equality in Eq. (A8) is found as in Eq. (28).

Finally, by inserting $\mathcal{A}_{\boldsymbol{y},1} = \mathcal{S}_{\boldsymbol{y}}$ into $f(\mathcal{A}_{\boldsymbol{y},1})$, and using the fact that $\mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0} + \varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}} = 2\varrho_{\boldsymbol{y},1}$ implies $\mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}}^2) = \mathrm{Tr}(\varrho_{\boldsymbol{y},1}\mathcal{S}_{\boldsymbol{y}})$, we find the minimum to be

$$\min_{\mathcal{A}_{\boldsymbol{y},1}} f(\mathcal{A}_{\boldsymbol{y},1}) = f(\mathcal{S}_{\boldsymbol{y}}) = -\mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}}^2), \quad\text{(A9)}$$

as stated in the main text.

Appendix B: The fundamental equations of the optimal quantum strategy

Upon adapting it to scale estimation, a fundamental result given by Holevo [12, 13] and Helstrom [1, 10] is as follows:

if $\mathcal{M}_{\boldsymbol{y}}(z)$ is the POM that minimises the overall uncertainty $\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}}$—hence called optimal—and $\tilde{\vartheta}_{\boldsymbol{y}}(z)$ is the associated estimator, then there exists an operator

$$\Upsilon_{\boldsymbol{y}} = \int dz\, \mathcal{M}_{\boldsymbol{y}}(z)\, W_{\boldsymbol{y}}[\tilde{\vartheta}_{\boldsymbol{y}}(z)] = \int dz\, W_{\boldsymbol{y}}[\tilde{\vartheta}_{\boldsymbol{y}}(z)]\, \mathcal{M}_{\boldsymbol{y}}(z) \quad\text{(B1)}$$

satisfying

$$W_{\boldsymbol{y}}[\tilde{\vartheta}_{\boldsymbol{y}}(z)] - \Upsilon_{\boldsymbol{y}} \geq 0 \quad\text{(B2)}$$

and leading to the minimum uncertainty $\bar{\epsilon}_{\boldsymbol{y},\mathrm{min}} = \mathrm{Tr}(\Upsilon_{\boldsymbol{y}})$.

Sec. III B showed that the optimal strategy for quantum scale estimation is given by

$$\mathcal{M}_{\boldsymbol{y}}(z) \mapsto \mathcal{M}_{\boldsymbol{y}}(s) = \mathcal{P}_{\boldsymbol{y}}(s) \quad\text{(B3)}$$

and

$$\tilde{\vartheta}(z) \mapsto \tilde{\vartheta}(s) = \theta_u \exp(s), \quad\text{(B4)}$$

where $\mathcal{P}_{\boldsymbol{y}}(s)$ is the projector associated with the eigendecomposition $\mathcal{S}_{\boldsymbol{y}} = \int ds\, \mathcal{P}_{\boldsymbol{y}}(s)\, s$. To see that this strategy does indeed verify Eq. (B2), we now follow the steps of the related proof for the mean square error in Ref. [1, Ch. 8].

By inserting Eqs. (B3) and (B4) into Eq. (7), we find that $W_{\boldsymbol{y}}[\tilde{\vartheta}_{\boldsymbol{y}}(s)] = W_{\boldsymbol{y}}[\theta_u \exp(s)] = \varrho_{\boldsymbol{y},2} + \varrho_{\boldsymbol{y},0}s^2 - 2\varrho_{\boldsymbol{y},1}s$, and so

$$\Upsilon_{\boldsymbol{y}} = \varrho_{\boldsymbol{y},2} + \mathcal{S}_{\boldsymbol{y}}^2\varrho_{\boldsymbol{y},0} - 2\mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},1} = \varrho_{\boldsymbol{y},2} + \varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}}^2 - 2\varrho_{\boldsymbol{y},1}\mathcal{S}_{\boldsymbol{y}}. \quad\text{(B5)}$$

Using Eq. (20) and the two forms of $\Upsilon_{\boldsymbol{y}}$ in Eq. (B5), we can write

$$W_{\boldsymbol{y}}[\tilde{\theta}_{\boldsymbol{y}}(s)] = \varrho_{\boldsymbol{y},2} + s\varrho_{\boldsymbol{y},0}s - \mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0}s - s\varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}} \quad\text{(B6)}$$

and

$$\Upsilon_{\boldsymbol{y}} = \varrho_{\boldsymbol{y},2} - \mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}}, \quad\text{(B7)}$$

so that $W_{\boldsymbol{y}}[\tilde{\theta}_{\boldsymbol{y}}(s)] - \Upsilon_{\boldsymbol{y}} = (\mathcal{S}_{\boldsymbol{y}} - s)\varrho_{\boldsymbol{y},0}(\mathcal{S}_{\boldsymbol{y}} - s)$. By using this and the fact that $\mathcal{S}_{\boldsymbol{y}}^\dagger = \mathcal{S}_{\boldsymbol{y}}$, we find that

$$\begin{aligned}
\langle v|\left\{W[\tilde{\theta}_{\boldsymbol{y}}(s)] - \Upsilon_{\boldsymbol{y}}\right\}|v\rangle &= \langle v|\,(\mathcal{S}_{\boldsymbol{y}} - s)\varrho_{\boldsymbol{y},0}(\mathcal{S}_{\boldsymbol{y}} - s)\,|v\rangle \\
&= \langle v'|\,\varrho_{\boldsymbol{y},0}\,|v'\rangle \geq 0, \quad\text{(B8)}
\end{aligned}$$

where $|v'\rangle = (\mathcal{S}_{\boldsymbol{y}} - s)\,|v\rangle$ and $|v\rangle$ is an arbitrary ket. Eq. (B8) is equivalent to Eq. (B2), as required. Moreover, note that $\mathrm{Tr}(\Upsilon_{\boldsymbol{y}}) = \mathrm{Tr}(\varrho_{\boldsymbol{y},2} - \mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}})$ leads to the minimum mean square error in Eq. (21).

Appendix C: Alternative proof of Eqs. (21), (23) **and** (25)

This appendix recovers the results in Sec. III using a different method—that in Ref. [30]. Start by minimising $\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}} = \mathcal{E}[\tilde{\theta}_{\boldsymbol{y}}(x), M_{\boldsymbol{y}}(x)]$ over the estimator $\tilde{\theta}_{\boldsymbol{y}}(x)$ while keeping the POM $M_{\boldsymbol{y}}(x)$ fixed. Here, $\mathcal{E}$ is defined as in Eq. (65). As per Ref. [33], this gives

$$\min_{\tilde{\theta}_{\boldsymbol{y}}(x)} \mathcal{E}[\tilde{\theta}_{\boldsymbol{y}}(x), M_{\boldsymbol{y}}(x)] = A - B_{\boldsymbol{y}}[M_{\boldsymbol{y}}(x)], \qquad \text{(C1)}$$

where

$$A = \int d\theta\, p(\theta) \log^2\left(\frac{\theta}{\theta_u}\right) \qquad \text{(C2)}$$

is a measurement-independent quantity, the POM enters in $B[M_{\boldsymbol{y}}(x)]$ as

$$B[M_{\boldsymbol{y}}(x)] = \int dx\, p(x|\boldsymbol{y}) \log^2\left[\frac{\tilde{\vartheta}_{\boldsymbol{y}}(x)}{\theta_u}\right], \qquad \text{(C3)}$$

and $\tilde{\vartheta}_{\boldsymbol{y}}(x)$ is the optimal estimator given by

$$\tilde{\vartheta}_{\boldsymbol{y}}(x) = \theta_u \exp\left[\int d\theta\, p(\theta|x, \boldsymbol{y}) \log\left(\frac{\theta}{\theta_u}\right)\right]. \qquad \text{(C4)}$$

Next, use the Born rule $p(x|\theta, \boldsymbol{y}) = \mathrm{Tr}[M_{\boldsymbol{y}}(x)\rho_{\boldsymbol{y}}(\theta)]$ to write $p(x|\boldsymbol{y}) = \mathrm{Tr}[M_{\boldsymbol{y}}(x)\varrho_{\boldsymbol{y},0}]$,

$$\tilde{\vartheta}_{\boldsymbol{y}}(x) = \theta_u \exp\left\{\frac{\mathrm{Tr}[M_{\boldsymbol{y}}(x)\varrho_{\boldsymbol{y},1}]}{\mathrm{Tr}[M_{\boldsymbol{y}}(x)\varrho_{\boldsymbol{y},0}]}\right\} \qquad \text{(C5)}$$

and, as a consequence,

$$B_{\boldsymbol{y}}[M_{\boldsymbol{y}}(x)] = \int dx\, \frac{\mathrm{Tr}[M_{\boldsymbol{y}}(x)\varrho_{\boldsymbol{y},1}]^2}{\mathrm{Tr}[M_{\boldsymbol{y}}(x)\varrho_{\boldsymbol{y},0}]}, \qquad \text{(C6)}$$

where $\varrho_{\boldsymbol{y},0}$ and $\varrho_{\boldsymbol{y},1}$ have been defined in Sec. III A.

Eq. (C6) is formally equivalent to the starting point for the proof of the celebrated Braunstein-Caves inequality [110, 111], as well as of its Bayesian analogue [30, 73]. Following the structure of such proofs, *redefine* the operator $\varrho_{\boldsymbol{y},1}$ as $\varrho_{\boldsymbol{y},1} = (\mathcal{X}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0} + \varrho_{\boldsymbol{y},0}\mathcal{X}_{\boldsymbol{y}})/2$. At this point, the meaning of $\mathcal{X}_{\boldsymbol{y}}$ is unknown. However, to simplify the notation we may rename such $\mathcal{X}_{\boldsymbol{y}}$ as $\mathcal{S}_{\boldsymbol{y}}$ in anticipation to the label employed in the main text. If we use this to rewrite Eq. (C6), then we can upper-bound $B_{\boldsymbol{y}}[M_{\boldsymbol{y}}(x)]$, for any $p(\theta)$ and $\rho_{\boldsymbol{y}}(\theta)$, as

$$\begin{aligned} B_{\boldsymbol{y}}[M_{\boldsymbol{y}}(x)] &= \int dx \left(\frac{\mathrm{Re}\{\mathrm{Tr}[M_{\boldsymbol{y}}(x)\mathcal{S}_{\boldsymbol{y}}\rho_{\boldsymbol{y},0}]\}}{\sqrt{\mathrm{Tr}[M_{\boldsymbol{y}}(x)\rho_{\boldsymbol{y},0}]}}\right)^2 \\ &\leq \int dx \left|\frac{\mathrm{Tr}[M_{\boldsymbol{y}}(x)\mathcal{S}_{\boldsymbol{y}}\rho_{\boldsymbol{y},0}]}{\sqrt{\mathrm{Tr}[M_{\boldsymbol{y}}(x)\rho_{\boldsymbol{y},0}]}}\right|^2 \\ &= \int dx \left|\mathrm{Tr}\left[\frac{\varrho_{\boldsymbol{y},0}^{\frac{1}{2}}M_{\boldsymbol{y}}(x)^{\frac{1}{2}}}{\sqrt{\mathrm{Tr}[M_{\boldsymbol{y}}(x)\rho_{\boldsymbol{y},0}]}}M_{\boldsymbol{y}}(x)^{\frac{1}{2}}\mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0}^{\frac{1}{2}}\right]\right|^2 \\ &\leq \int dx\, \mathrm{Tr}[M_{\boldsymbol{y}}(x)\mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}}] \\ &= \mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}}^2), \end{aligned} \qquad \text{(C7)}$$

where we have used the Cauchy-Schwarz inequality [7]

$$|\mathrm{Tr}(Y^\dagger Z)|^2 \leq \mathrm{Tr}(Y^\dagger Y)\,\mathrm{Tr}(Z^\dagger Z), \qquad \text{(C8)}$$

with

$$Y = \frac{M_{\boldsymbol{y}}(x)^{\frac{1}{2}}\varrho_{\boldsymbol{y},0}^{\frac{1}{2}}}{\sqrt{\mathrm{Tr}[M_{\boldsymbol{y}}(x)\rho_{\boldsymbol{y},0}]}}, \quad Z = M_{\boldsymbol{y}}(x)^{\frac{1}{2}}\mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0}^{\frac{1}{2}}. \qquad \text{(C9)}$$

The combination of Eqs. (C1), (C2) and (C7) thus leads to

$$\bar{\epsilon}_{\boldsymbol{y},\mathrm{mle}} \geq \int d\theta\, p(\theta) \log^2\left(\frac{\theta}{\theta_u}\right) - \mathrm{Tr}(\varrho_{\boldsymbol{y},0}\mathcal{S}_{\boldsymbol{y}}^2). \qquad \text{(C10)}$$

A real-valued $\mathrm{Tr}[M_{\boldsymbol{y}}(x)\mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0}]$ saturates the first inequality in Eq. (C7). The Cauchy-Schwarz inequality is saturated when $Y \propto Z$ [7], which implies that

$$\frac{M_{\boldsymbol{y}}(x)^{\frac{1}{2}}\varrho_{\boldsymbol{y},0}^{\frac{1}{2}}}{\mathrm{Tr}[M_{\boldsymbol{y}}(x)\varrho_{\boldsymbol{y},0}]} = \frac{M_{\boldsymbol{y}}(x)^{\frac{1}{2}}\mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0}^{\frac{1}{2}}}{\mathrm{Tr}[M_{\boldsymbol{y}}(x)\mathcal{S}_{\boldsymbol{y}}\varrho_{\boldsymbol{y},0}]}. \qquad \text{(C11)}$$

By choosing

$$M_{\boldsymbol{y}}(x) \mapsto \mathcal{M}_{\boldsymbol{y}}(s) = \mathcal{P}_{\boldsymbol{y}}(s), \qquad \text{(C12)}$$

where $\mathcal{P}_{\boldsymbol{y}}(s)$ is the projector associated with $\mathcal{S}_{\boldsymbol{y}}$, both saturation conditions are fulfilled at once [30, 111]. Furthermore, inserting Eq. (C12) into Eq. (C5) allows us to write the optimal estimator as

$$\tilde{\vartheta}_{\boldsymbol{y}}(s) = \theta_u \exp\left\{\frac{\mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(s)\varrho_{\boldsymbol{y},1}]}{\mathrm{Tr}[\mathcal{P}_{\boldsymbol{y}}(s)\varrho_{\boldsymbol{y},0}]}\right\} = \theta_u \exp(s), \qquad \text{(C13)}$$

where the second inequality stems from Eq. (27).

Given that the equality in Eq. (C10) can be reached by using the measurement in Eq. (C12) and the estimator in Eq. (C13), we identify it as the minimum associated with our original optimisation problem in Eq. (8). As expected, it is identical to that in Eq. (21), as are Eqs. (C12) and (C13) to Eqs. (23) and (25) in the main text, respectively.

Appendix D: A multiparameter prior density

To have a notion of maximum ignorance in multiparameter scale estimation, Sec. V D identifies the constraint $\prod_i p(\theta_i) = \prod_i \gamma_i\, p(\gamma_i\theta_i)$. This functional equation may be solved as follows. Rewrite it as

$$\gamma_1 p(\gamma_1\theta_1) = p(\theta_1) \prod_{i\neq 1} \frac{p(\theta_i)}{\gamma_i\, p(\gamma_i\theta_i)} \qquad \text{(D1)}$$

and take the derivative with respect to γ_1 [24]; this transforms Eq. (D1) into $p(\gamma_1\theta_1) + \gamma_1\,\theta_1\,\dot{p}(\gamma_1\theta_1) = 0$, where $\dot{p}(z) := dp(z)/dz$. In turn, this leads to the ordinary differential equation

$$p(z) + z\,\dot{p}(z) = 0 \qquad \text{(D2)}$$

whose solution can be found straightforwardly as

$$\frac{dp(z)}{p(z)} = -\frac{dz}{z} \implies p(z) = \frac{\text{const.}}{z} \propto \frac{1}{z}. \quad \text{(D3)}$$

Choosing $z \mapsto \theta_1$ leads to Jeffreys's prior $p(\theta_1) \propto 1/\theta_1$ for the first parameter. Since the same reasoning applies to the other parameters, we recover the result in the main text, i.e., $p(\boldsymbol{\theta}) \propto \prod_i (1/\theta_i)$.

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

[2] M. G. A. paris, Quantum estimation for quantum technology, International Journal of Quantum Information **07**, 125 (2009).

[3] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, Quantum Limits in Optical Interferometry, Progress in Optics **60**, 345 (2015).

[4] R. Demkowicz-Dobrzański, W. Górecki, and M. Guţă, Multiparameter estimation beyond quantum Fisher information, Journal of Physics A: Mathematical and Theoretical **53**, 363001 (2020).

[5] J. S. Sidhu and P. Kok, Geometric perspective on quantum parameter estimation, AVS Quantum Science **2**, 014701 (2020).

[6] C. W. Helstrom, Minimum mean-squared error of estimates in quantum statistics, Physics Letters A **25**, 101 (1967).

[7] C. Helstrom, The minimum variance of estimates in quantum signal detection, IEEE Transactions on Information Theory **14**, 234 (1968).

[8] S. Personick, *Efficient analog communication over quantum channels*, Ph.D. thesis, Dep. Elec. Eng., M.I.T., Cambridge, Mass. (1969).

[9] S. Personick, Application of quantum estimation theory to analog communication over quantum channels, IEEE Transactions on Information Theory **17**, 240 (1971).

[10] C. Helstrom and R. Kennedy, Noncommuting observables in quantum detection and estimation theory, IEEE Transactions on Information Theory **20**, 16 (1974).

[11] H. Yuen and M. Lax, Multiple-parameter quantum estimation and measurement of nonselfadjoint observables, IEEE Transactions on Information Theory **19**, 740 (1973).

[12] A. S. Holevo, Statistical decision theory for quantum systems, Journal of Multivariate Analysis **3**, 337 (1973).

[13] A. S. Holevo, Statistical problems in quantum physics, in *Proceedings of the Second Japan-USSR Symposium on Probability Theory*, edited by G. Maruyama and Y. V. Prokhorov (Springer Berlin Heidelberg, Berlin, Heidelberg, 1973) pp. 104–119.

[14] J. P. Dowling and G. J. Milburn, Quantum technology: the second quantum revolution, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **361**, 1655 (2003).

[15] J. A. Dunningham, Using quantum theory to improve measurement precision, Contemporary Physics **47**, 257 (2006).

[16] S. Vinjanampathy and J. Anders, Quantum thermodynamics, Contemporary Physics **57**, 545 (2016).

[17] S. M. Barnett, A. Beige, A. Ekert, B. M. Garraway, C. H. Keitel, V. Kendon, M. Lein, G. J. Milburn, H. M. Moya-Cessa, M. Murao, J. K. Pachos, G. M. Palma, E. Paspalakis, S. J. D. Phoenix, B. Piraux, M. B. Plenio, B. C. Sanders, J. Twamley, A. Vidiella-Barranco, and M. S. Kim, Journeys from Quantum Optics to Quantum Technology, Progress in Quantum Electronics **54**, 19 (2017).

[18] D. E. Browne, S. Bose, F. Mintert, and M. S. Kim, From Quantum Optics to Quantum Technologies, Progress in Quantum Electronics **54**, 2 (2017).

[19] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Reviews of Modern Physics **89**, 035002 (2017).

[20] A. Acín, I. Bloch, H. Buhrman, T. Calarco, C. Eichler, J. Eisert, D. Esteve, N. Gisin, S. J. Glaser, F. Jelezko, S. Kuhr, M. Lewenstein, M. F. Riedel, P. O. Schmidt, R. Thew, A. Wallraff, I. Walmsley, and F. K. Wilhelm, The quantum technologies roadmap: a European community view, New Journal of Physics **20**, 080201 (2018).

[21] A. Belenchia, M. Carlesso, Ömer Bayraktar, D. Dequal, I. Derkach, G. Gasbarri, W. Herr, Y. L. Li, M. Rademacher, J. Sidhu, D. K. Oi, S. T. Seidel, R. Kaltenbaek, C. Marquardt, H. Ulbricht, V. C. Usenko, L. Wörner, A. Xuereb, M. Paternostro, and A. Bassi, Quantum physics in space, Physics Reports **951**, 1 (2022).

[22] A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (Edizioni della Normale, Springer Basel, 2011).

[23] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice Hall, 1993).

[24] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, 2003).

[25] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Generalized Uncertainty Relations: Theory, Examples, and Lorentz Invariance, Annals of Physics **247**, 135 (1996).

[26] D. C. Brody and L. P. Hughston, Geometry of Quantum Statistical Inference, Phys. Rev. Lett. **77**, 2851 (1996).

[27] Y. Watanabe, T. Sagawa, and M. Ueda, Uncertainty relation revisited from quantum estimation theory, Phys. Rev. A **84**, 042121 (2011).

[28] M. Tsang, H. M. Wiseman, and C. M. Caves, Fundamental quantum limit to waveform estimation, Phys. Rev. Lett. **106**, 090401 (2011).

[29] T. G. Downes, J. R. van Meter, E. Knill, G. J. Milburn, and C. M. Caves, Quantum estimation of parameters of classical spacetimes, Phys. Rev. D **96**, 105004 (2017).

[30] J. Rubio and J. Dunningham, Bayesian multiparameter quantum metrology with limited data, Phys. Rev. A **101**, 032114 (2020).

[31] D. C. Brody and L. P. Hughston, Generalised Heisenberg relations for quantum statistical estimation, Physics Letters A **236**, 257 (1997).

[32] H. J. D. Miller and J. Anders, Energy-temperature uncertainty relation in quantum thermodynamics, Nature Communications **9**, 2203 (2018).

[33] J. Rubio, J. Anders, and L. A. Correa, Global quantum thermometry, Phys. Rev. Lett. **127**, 190402 (2021).

[34] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, Compatibility in multiparameter quantum metrology, Physical Review A **94**, 052108 (2016).

[35] A. Carollo, B. Spagnolo, A. A. Dubkov, and D. Valenti, On quantumness in multi-parameter quantum estimation, Journal of Statistical Mechanics: Theory and Experiment **2019**, 094010 (2019).

[36] F. Albarelli, J. F. Friel, and A. Datta, Evaluating the Holevo Cramér-Rao Bound for Multiparameter Quantum Metrology,

Phys. Rev. Lett. **123**, 200503 (2019).

[37] F. Albarelli, M. Barbieri, M. G. Genoni, and I. Gianani, A perspective on multiparameter quantum metrology: From theoretical tools to applications in quantum imaging, Physics Letters A , 126311 (2020).

[38] M. Tsang, F. Albarelli, and A. Datta, Quantum semiparametric estimation, Phys. Rev. X **10**, 031023 (2020).

[39] J. S. Sidhu, Y. Ouyang, E. T. Campbell, and P. Kok, Tight bounds on the simultaneous estimation of incompatible parameters, Phys. Rev. X **11**, 011028 (2021).

[40] A. Candeloro, M. G. A. Paris, and M. G. Genoni, On the properties of the asymptotic incompatibility measure in multiparameter quantum estimation, Journal of Physics A: Mathematical and Theoretical **54**, 485301 (2021).

[41] J. H. Shapiro and S. R. Shepard, Quantum phase measurement: A system-theory perspective, Phys. Rev. A **43**, 3795 (1991).

[42] L. Pezzè and A. Smerzi, Quantum theory of phase estimation, in *Atom Interferometry, Proceedings of the International School of Physics Enrico Fermi, Course 188, Varenna*, edited by G. Tino and M. Kasevich (Springer Berlin Heidelberg, IOS Press, Amsterdam, 2014) p. 691.

[43] D. W. Berry and H. M. Wiseman, Optimal states and almost optimal adaptive measurements for quantum interferometry, Physical Review Letters **85**, 5098 (2000).

[44] C. Macchiavello, Optimal estimation of multiple phases, Phys. Rev. A **67**, 062302 (2003).

[45] R. Demkowicz-Dobrzański, Optimal phase estimation with arbitrary a priori knowledge, Physical Review A **83**, 061802 (2011).

[46] V. Gebhart, A. Smerzi, and L. Pezzè, Bayesian quantum multiphase estimation algorithm, Phys. Rev. Applied **16**, 014035 (2021).

[47] J. Rubio and J. A. Dunningham, Quantum metrology in the presence of limited data, New Journal of Physics **21**, 043037 (2019).

[48] H. B. Prosper, Temperature fluctuations in a heat bath., American Journal of Physics **61**, 54 (1993).

[49] T. M. Stace, Quantum limits of thermometry, Phys. Rev. A **82**, 011611 (2010).

[50] T. Jahnke, S. Lanéry, and G. Mahler, Operational approach to fluctuations of thermodynamic variables in finite quantum systems, Phys. Rev. E **83**, 011109 (2011).

[51] A. De Pasquale and T. M. Stace, Quantum thermometry, in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*, edited by F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso (Springer International Publishing, Cham, 2018) pp. 503–527.

[52] P. P. Potts, J. B. Brask, and N. Brunner, Fundamental limits on low-temperature quantum thermometry with finite resolution, Quantum **3**, 161 (2019).

[53] W.-K. Mok, K. Bharti, L.-C. Kwek, and A. Bayat, Optimal probes for global quantum thermometry, Communications physics **4**, 1 (2021).

[54] G. O. Alves and G. T. Landi, Bayesian estimation for collisional thermometry, Phys. Rev. A **105**, 012212 (2022).

[55] M. Mehboudi, M. R. Jørgensen, S. Seah, J. B. Brask, J. Kołodyński, and M. Perarnau-Llobet, Fundamental Limits in Bayesian Thermometry and Attainability via Adaptive Strategies, Phys. Rev. Lett. **128**, 130502 (2022).

[56] M. R. Jørgensen, J. Kołodyński, M. Mehboudi, M. Perarnau-Llobet, and J. B. Brask, Bayesian quantum thermometry based on thermodynamic length, Phys. Rev. A **105**, 042601 (2022).

[57] J. Boeyens, S. Seah, and S. Nimmrichter, Uninformed Bayesian quantum thermometry, Phys. Rev. A **104**, 052214 (2021).

[58] P. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica Support* (Cambridge University Press, 2005).

[59] E. T. Jaynes, Prior probabilities, IEEE Transactions on Systems and Cybernetics **4**, 227 (1968).

[60] U. von Toussaint, Bayesian inference in physics, Reviews of Modern Physics **83**, 943 (2011).

[61] M. D. Baaske, M. R. Foreman, and F. Vollmer, Single-molecule nucleic acid interactions monitored on a label-free microcavity biosensor platform, Nature nanotechnology **9**, 933 (2014).

[62] S. Subramanian, S. Frustaci, and F. Vollmer, Microsecond single-molecule enzymology using plasmonically enhanced optical resonators, in *Frontiers in Biological Detection: From Nanosensors to Systems XII*, Vol. 11258, edited by A. Danielli, B. L. Miller, and S. M. Weiss, International Society for Optics and Photonics (SPIE, 2020) pp. 23 – 30.

[63] S. Subramanian, H. Jones, S. Frustaci, S. Winter, M. van der Kamp, V. Arcus, C. Pudney, and F. Vollmer, Sensing enzyme activation heat capacity at the single-molecule level using gold-nanorod-based optical whispering gallery modes, ACS Applied Nano Materials **4**, 4576–4583 (2021).

[64] N. Eerqing, S. Subramanian, J. Rubio, T. Lutz, H.-Y. Wu, J. Anders, C. Soeller, and F. Vollmer, Comparing Transient Oligonucleotide Hybridization Kinetics Using DNA-PAINT and Optoplasmonic Single-Molecule Sensing on Gold Nanorods, ACS Photonics **8**, 2882 (2021).

[65] K. T. Mpofu, C. Lee, G. E. M. Maguire, H. G. Kruger, and M. S. Tame, Measuring kinetic parameters using quantum plasmonic sensing, Phys. Rev. A **105**, 032619 (2022).

[66] K. T. Mpofu, C. Lee, G. E. M. Maguire, H. G. Kruger, and M. S. Tame, Experimental measurement of kinetic parameters using quantum plasmonic sensing, Journal of Applied Physics **131**, 084402 (2022).

[67] S. M. S. M. Barnett, *Quantum information*, Oxford master series in physics. Atomic, optical and laser physics; 16 (Oxford University Press, Oxford, 2009).

[68] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, 10th ed. (Cambridge University Press, Cambridge, 2010).

[69] M. Mehboudi, A. Sanpera, and L. A. Correa, Thermometry in the quantum regime: recent theoretical progress, Journal of Physics A: Mathematical and Theoretical **52**, 303001 (2019).

[70] H. Jeffreys, *Theory of Probability*, 3rd ed. (Clarendon P., 1961).

[71] J. Aczél, *A Short Course on Functional Equations* (Springer Netherlands, 1987).

[72] F. Hanamura, W. Asavanant, K. Fukui, S. Konno, and A. Furusawa, Estimation of Gaussian random displacement using non-Gaussian states, Phys. Rev. A **104**, 062601 (2021).

[73] J. Rubio Jiménez, *Non-asymptotic quantum metrology: extracting maximum information from limited data*, Ph.D. thesis, University of Sussex (2020).

[74] S. Morelli, A. Usui, E. Agudelo, and N. Friis, Bayesian parameter estimation using Gaussian states and measurements, Quantum Science and Technology **6**, 25018 (2021).

[75] I. Bialynicki-Birula, M. Freyberger, and W. Schleich, Various measures of quantum phase uncertainty: a comparative study, Physica Scripta **T48**, 113 (1993).

[76] A. Luis and A. Rodil, Alternative measures of uncertainty in quantum metrology: Contradictions and limits, Phys. Rev. A

87, 034101 (2013).

[77] C. F. Matta, L. Massa, A. V. Gubskaya, and E. Knoll, Can One Take the Logarithm or the Sine of a Dimensioned Quantity or a Unit? Dimensional Analysis Involving Transcendental Functions, Journal of Chemical Education 88, 67 (2011).

[78] F. Hansen and G. K. Pedersen, Jensen's Operator Inequality, Bulletin of the London Mathematical Society 35, 553 (2003).

[79] K. Macieszczak, M. Fraas, and R. Demkowicz-Dobrzański, Bayesian quantum frequency estimation in presence of collective dephasing, New Journal of Physics 16, 113002 (2014).

[80] K. F. Riley, M. P. Hobson, and S. J. Bence, *Mathematical methods for physics and engineering* (Cambridge University Press, 2004).

[81] J. Z. Bernád, C. Sanavio, and A. Xuereb, Optimal estimation of the optomechanical coupling strength, Phys. Rev. A 97, 063821 (2018).

[82] L. A. Correa, M. Mehboudi, G. Adesso, and A. Sanpera, Individual quantum probes for optimal thermometry, Phys. Rev. Lett. 114, 220405 (2015).

[83] M. E. Pearce, E. T. Campbell, and P. Kok, Optimal quantum metrology of distant black bodies, Quantum 1, 21 (2017).

[84] R. K. Pathria, *Statistical mechanics*, 3rd ed. (Elsevier, Amsterdam, 2011).

[85] A. M. Mood, *Introduction to the theory of statistics*, 3rd ed. (McGraw Hill, New York, 1974).

[86] Y. Li, L. Pezzè, M. Gessner, Z. Ren, W. Li, and A. Smerzi, Frequentist and Bayesian Quantum Phase Estimation, Entropy 20, 628 (2018).

[87] M. Szczykulska, T. Baumgratz, and A. Datta, Multi-parameter quantum metrology, Advances in Physics: X 1, 621 (2016).

[88] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, Quantum enhanced multiple phase estimation, Phys. Rev. Lett. 111, 070403 (2013).

[89] T. J. Proctor, P. A. Knott, and J. A. Dunningham, Networked quantum sensing, arXiv:1702.04271 (2017).

[90] T. J. Proctor, P. A. Knott, and J. A. Dunningham, Multiparameter estimation in networked quantum sensors, Phys. Rev. Lett. 120, 080501 (2018).

[91] K.-K. Lee, C. Gagatsos, S. Guha, and A. Ashok, Quantum Multi-Parameter Adaptive Bayesian Estimation and Application to Super-Resolution Imaging, arXiv:2202.09980 (2022).

[92] A. Luis, Complementarity for Generalized Observables, Phys. Rev. Lett. 88, 230401 (2002).

[93] R. T. Cox, Probability, Frequency and Reasonable Expectation, American Journal of Physics 14, 1 (1946).

[94] J. B. Paris, *The Uncertain Reasoner's Companion: A Mathematical Perspective* (Cambridge University Press, 1994).

[95] K. S. Van Horn, Constructing a logic of plausible inference: a guide to Cox's theorem, International Journal of Approximate Reasoning 34, 3 (2003).

[96] L. E. Ballentine, Propensity, Probability, and Quantum Theory, Foundations of physics 46, 973 (2016).

[97] R. E. Kass and L. Wasserman, The Selection of Prior Distributions by Formal Rules, Journal of the American Statistical Association 91, 1343 (1996).

[98] W. v. d. Linden, V. Dose, and U. v. Toussaint, *Bayesian Probability Theory: Applications in the Physical Sciences* (Cambridge University Press, 2014).

[99] M. Larocca, F. Sauvage, F. M. Sbahi, G. Verdon, P. J. Coles, and M. Cerezo, Group-Invariant Quantum Machine Learning, arXiv:2205.02261 (2022).

[100] H. Jeffreys, An invariant form for the prior probability in estimation problems, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences 186, 453 (1946).

[101] B. de Finetti, *Theory of Probability: a Critical introductory Treatment*, Wiley Classics Library ed. (John Wiley & Sons Ltd., 1990).

[102] J. M. Bernardo, *Bayesian theory* (Wiley, Chichester, 1994).

[103] J. S. Rosenthal, *A first look at rigorous probability theory*, 2nd ed. (Singapore; Hackensack, N. J.: World Scientific, 2006).

[104] W. Górecki and R. Demkowicz-Dobrzański, Multiple-phase quantum interferometry: Real and apparent gains of measuring all the phases simultaneously, Phys. Rev. Lett. 128, 040504 (2022).

[105] M. Jarzyna and M. Zwierz, Quantum interferometric measurements of temperature, Phys. Rev. A 92, 032112 (2015).

[106] A. De Pasquale, D. Rossini, R. Fazio, and V. Giovannetti, Local quantum thermal susceptibility, Nature communications 7, 1 (2016).

[107] A. De Pasquale, K. Yuasa, and V. Giovannetti, Estimating temperature via sequential measurements, Phys. Rev. A 96, 012316 (2017).

[108] A. H. Kiilerich, A. De Pasquale, and V. Giovannetti, Dynamical approach to ancilla-assisted quantum thermometry, Phys. Rev. A 98, 042124 (2018).

[109] S. Razavian, C. Benedetti, M. Bina, Y. Akbari-Kourbolagh, and M. G. A. Paris, Quantum thermometry by single-qubit dephasing, The European Physical Journal Plus 134, 284 (2019).

[110] S. L. Braunstein and C. M. Caves, Statistical distance and the geometry of quantum states, Phys. Rev. Lett. 72, 3439 (1994).

[111] M. G. Genoni, P. Giorda, and M. G. A. Paris, Optimal estimation of entanglement, Phys. Rev. A 78, 032303 (2008).