Measurement incompatibility in Bayesian multiparameter quantum estimation

Francesco Albarelli[1,2], Dominic Branford[3], and Jesús Rubio[4,5]

[1]Università di Parma, Dipartimento di Scienze Matematiche, Fisiche e Informatiche, I-43124 Parma, Italy
[2]INFN—Sezione di Milano-Bicocca, gruppo collegato di Parma, I-43124 Parma, Italy
[3]Dipartimento di Fisica e Astronomia, Università di Firenze, via G. Sansone 1, I-50019 Sesto Fiorentino (FI), Italy
[4]School of Mathematics and Physics, University of Surrey, Guildford GU2 7XH, United Kingdom
[5]Department of Physics and Astronomy, University of Exeter, Stocker Road, Exeter EX4 4QL, United Kingdom
3rd August 2026

arXiv:2511.16645v4 [quant-ph] 3 Aug 2026

We present a comprehensive and pedagogical formulation of Bayesian multiparameter quantum estimation. Within this framework, we analyse the role of measurement incompatibility and establish its quantitative effect on attainable precision. We achieve this by deriving upper bounds based on the pretty good measurement—a notion from hypothesis testing—combined with the evaluation of the Nagaoka–Hayashi lower bound. In general, we prove that, as in the many-copy regime of local estimation theory, incompatibility can at most double the minimum loss relative to the idealised scenario in which individually optimal measurements are assumed jointly implementable. Therefore, in practical situations, the latter may provide a sufficient and computationally efficient benchmark without solving the full optimisation problem. Our results, which we illustrate through applications of discrete phase imaging, phase and dephasing estimation, and qubit sensing, provide analytical and numerical tools for assessing ultimate precision limits and the role of measurement incompatibility in Bayesian multiparameter quantum metrology, including an open-source package for all the bounds discussed here.

1 Introduction

Quantum metrology explores the fundamental and practical impact of quantum mechanical phenomena, such as superposition, squeezing, and entanglement, on high-precision measurements of physical parameters [1–5]. Its central aim is to surpass the precision limits of classical strategies that do not rely on these quantum effects. Even if such advantage is not always feasible, a proper understanding of the fundamental precision limits imposed by quantum mechanics is crucial to develop new technologies and guide resource allocation [6–8].

Francesco Albarelli: francesco.albarelli@gmail.com
Dominic Branford: dominic.branford@unifi.it
Jesús Rubio: jesus@rubiojimenez.com

Quantum estimation theory (QET) provides the mathematical framework for analysing the precision limits achievable with quantum systems. Its most celebrated result is the quantum Cramér–Rao bound (CRB) [9]. While many results in quantum metrology rely on the CRB, this approach has important limitations. Roughly speaking, the CRB is a meaningful figure of merit only when a large number of experiment repetitions is performed, with a significant caveat that it is often hard to understand a priori how many repetitions are needed. This flavour of QET is known as *local*, meaning that its predictions are generally valid when the discrepancy between the estimates and the true value of the parameters is sufficiently small, which *eventually* happens as samples accumulate.

Global estimation theory [10, 11] represents a complementary approach that is especially relevant when resources are constrained and only a limited number of experimental trials can be performed [12–14]. While several global approaches have been proposed [15–23], the most established and extensively developed framework for going beyond the local estimation paradigm is Bayesian quantum estimation. Dating back to the early days of the field [9, 24–29], it is nowadays being increasingly applied to a wide range of multiparameter sensing experiments involving both photons [30–35], atoms [36, 37], and solid-state systems [38]. This is the framework adopted in the present work.

Bayesian estimation is not only practically significant in regimes with limited experimental data but also holds deep fundamental implications. For example, it clarifies impractical claims arising from applying local QET beyond its valid domain, such as the alleged effectiveness of super-Heisenberg-limited schemes [12, 39, 40] or the apparent enhancement in multiple phase estimation [41–44], and it reveals the need to correct the Heisenberg limit in phase estima-

tion by a factor of π [45]. In general, the strategies predicted by local QET can become suboptimal in the non-asymptotic regime [20], where the Bayesian framework consistently enables efficient information extraction [13, 46]. This is unsurprising given that Bayesian inference constitutes the most basic form of *learning* [47], which is the most comprehensive approach for characterising quantum systems [48–56].

One of the fundamental aspects of quantum theory is the existence of incompatible measurements that cannot be implemented jointly [57, 58]. In the context of quantum metrology, this phenomenon becomes crucial when multiple parameters must be estimated simultaneously: measurement incompatibility can impose fundamental trade-offs on the achievable precision of different parameters [59–62], which have been quantified in various ways [63–66]. Besides this concrete consequence, incompatibility makes multiparameter theory mathematically challenging [10]. While the fundamental precision limits in single-parameter quantum metrology are relatively well-understood, fundamental multiparameter limits remain a subject of active research [67–69].

It is instructive to review the impact of such incompatibility effects on the development of local QET, before drawing a parallel with the Bayesian approach. A multiparameter version of the quantum CRB was first derived by Helstrom at the dawn of the field, based on the notion of the symmetric logarithmic derivative (SLD) [9, 70]. We shall call this bound the Helstrom (SLD) CRB. While this bound proved attainable for single-parameter estimation [71, 72], it does not account for the effects of incompatibility and is not generally attainable for multiparameter estimation. In the following decades, further quantum generalisations of the CRB were developed, providing a better account of measurement incompatibility, and culminating in the discovery that the ultimate precision limit attainable with collective (entangling) measurements over many copies is given by the Holevo–Cramér–Rao bound (HCRB) [10, 73–77]. It was eventually understood that, in this asymptotic scenario where many copies can be measured collectively, the effect of measurement incompatibility is relatively mild: the HCRB can be at most twice the original bound obtained by Helstrom [78, 79] (see also Refs. [10, 64, 67]).

By contrast, the fundamental precision achievable with separable measurements on each copy of the state is not efficiently computable in local QET, even numerically (although collective measurements are unnecessary for pure states [80]). Nonetheless, several bounds for this setting have been derived [81–86], among which we highlight the Nagaoka–Hayashi (NH) bound [87–94]. Most recently, the problem has been studied using conic programming [95], providing a unifying framework and new tools that have already been applied to probe state optimisation [96].

Despite significant advances in local QET, the role of measurement incompatibility in Bayesian estimation has received much less attention. The general conditions that a Bayesian estimation strategy must satisfy to be optimal were already identified in the aforementioned early works [9, 26, 27, 29], including a discrete analogue for hypothesis testing [97, 98]. A generalisation to convex functions has also been proposed [99]. Ref. [9] provided examples of how these conditions can be used to verify whether a proposed strategy is indeed optimal. Reversing this procedure—namely, using the conditions to derive the optimal strategy—is, however, exceedingly challenging, with only a limited number of problems where this can be done analytically. For instance, this procedure can be carried out in the simple case of two parameters [100].

For single-parameter estimation, the class of location-isomorphic estimation problems can be optimised exactly [101], and the solution can be shown to satisfy the corresponding optimality conditions. This family includes, as a particular case, the optimisation of the standard mean square error—called mean square loss (MSL) in this work—a problem originally solved by Personick [25]. While Personick made an early attempt to extend this reasoning to the estimation of two parameters [24], closed-form solutions based on indirect measurements were subsequently obtained [102], and a solution for the estimation of two displacement parameters in Gaussian states with Gaussian priors is known [103], a general and practical procedure for determining optimal solutions in the Bayesian estimation of multiple parameters has remained elusive for decades.

Only recently has a multiparameter bound, analogous to the Helstrom (SLD) CRB—and thus neglecting incompatibility effects—been derived [43]. As motivated later, we propose referring to this as the symmetric posterior mean (SPM) bound. Bayesian analogues of several other local bounds, including the NH bound, have since been established, incorporating measurement incompatibility to a certain extent [104, 105]. In addition, numerical algorithms for identifying optimal estimation strategies are now under active investigation [106, 107].

Continuing the parallel between the local and Bayesian frameworks, this work addresses a key question: how strongly does measurement incompatibility affect Bayesian multiparameter QET? In other words, what is the maximal discrepancy between the achievable precision and the bound that neglect incompatibility?

Contributions.—We show that, in analogy with local QET, the minimum MSL can be at most twice the SPM benchmark that is obtained when measurement incompatibility is ignored. Key differences, however, are that this upper bound now holds at the single-shot level, but can also be trivial, i.e., larger than the MSL derived solely from *a priori* informa-

tion. This effect stems from the central role of prior information in Bayesian estimation, which, as illustrated through examples in quantum imaging, phase and dephasing estimation, and qubit sensing, may significantly modify the apparent incompatibility of the estimation protocol. Consequently, measurement incompatibility in Bayesian multiparameter QET is not only a relatively mild factor in the overall estimation precision, but it can also be concealed when the influence of the prior is strong. We further derive a tighter upper bound that is guaranteed to be non-trivial, by combining measurements introduced in hypothesis testing [108–111] with optimal classical Bayes estimators [47]. Our upper bounds are complementary to existing lower bounds tighter than the SPM one, where the Bayesian NH bound [104] represents the state of the art. In addition, we derive a class of lower bounds based on generalised posterior mean operators beyond the symmetric version that, while less tight than the NH bound, may be of independent interest, since they are obtained via the theory of monotone metrics, analogously to established results in local QET [112]. An open-source code for the numerical evaluation of all relevant bounds is available on GitHub [113]. Finally, beyond deriving bounds on the minimum MSL, we determine explicit conditions for the measurement strategies that achieve it. Combined with the pedagogical and self-contained introduction to Bayesian multiparameter quantum metrology presented in this work, these results establish a solid foundation for further advances in the field.

2 Bayesian multiparameter quantum estimation theory

2.1 Preliminaries

Let $\boldsymbol{\Theta}$ denote a vector of unknown parameters and $\boldsymbol{\theta} \in \Omega$ a vector hypotheses—values the unknown parameter could take—both understood as $d\times 1$ column vectors in $\mathbb{R}^d$. Here, Ω represents a generic parameter space. The information about $\boldsymbol{\Theta}$ available prior to any measurement is encoded in a probability density $p(\boldsymbol{\theta})$, satisfying the non-negativity and normalisation conditions:

$$p(\boldsymbol{\theta}) \geq 0, \quad \int_{\Omega} \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta}) = 1. \tag{1}$$

This $p(\boldsymbol{\theta})$ is referred to as the *prior probability*.

To gather information about $\boldsymbol{\Theta}$ via measurements, the vector hypothesis is encoded in the quantum state of a suitable physical system, such as an atomic ensemble or a Mach–Zehnder interferometer, denoted by $\rho(\boldsymbol{\theta})$. This is a positive semidefinite operator with unit trace:

$$\rho(\boldsymbol{\theta}) \succeq 0, \quad \mathrm{Tr}[\rho(\boldsymbol{\theta})] = 1. \tag{2}$$

We focus on finite-dimensional quantum systems, and whilst we expect most results in Sec. 3 to hold more generally, we leave a more rigorous analysis for future work.

Once a state $\rho(\boldsymbol{\theta})$ is prepared, a measurement with outcome $x \in \chi$ is performed, where χ denotes the sampling space. Each outcome is associated with an element $M(x)$ of a positive operator-valued measure (POVM), which is positive semidefinite and resolves the identity:

$$M(x) \succeq 0, \quad \int_{\chi} \mathrm{d}x\, M(x) = \mathbb{I}. \tag{3}$$

For simplicity, in Sec. 4, we only consider examples with discrete outcomes, which is natural when focusing on finite-dimensional systems. In such cases, continuous outcomes can be realised by performing finite measurements depending on a classical parameter [114]. Conversely, discrete outcomes can be embedded in a continuous-outcome formulation by expressing $M(x)$ as a sum of Dirac delta functions centered on the discrete outcome n,

$$M(x) = \sum_n M_n\, \delta(x-n), \tag{4}$$

associated with a POVM element M_n satisfying $M_n \succeq 0$ and $\sum_n M_n = \mathbb{I}$. The continuous-outcome formalism is particularly convenient for variational calculus, which is frequently employed in Bayesian optimisation [25, 47].

At this point, Bayesian inference offers a straightforward solution to the problem of inferring the value of $\boldsymbol{\Theta}$ by combining Bayes' theorem with the Born rule. Specifically, the prior probability $p(\boldsymbol{\theta})$ is *updated* according to

$$p(\boldsymbol{\theta}) \mapsto p(\boldsymbol{\theta}|x) \propto p(\boldsymbol{\theta})\ \mathrm{Tr}[M(x)\rho(\boldsymbol{\theta})]. \tag{5}$$

Here, the probability density $p(\boldsymbol{\theta}|x)$ denotes the *posterior probability* while $\mathrm{Tr}[M(x)\rho(\boldsymbol{\theta})] = p(x|\boldsymbol{\theta})$ is identified as the *likelihood function*.

The posterior probability represents the full state of knowledge about the values of the parameters $\boldsymbol{\theta}$ available after observing the measurement outcome x. In that sense, it could be already considered a complete solution of the problem.

Calculation of the posterior probabilities underpins all Bayesian measurement protocols. However, more ergonomical theoretical constructions are generally required. For instance, it is customary to identify a single vector value representing our best guess for the unknown $\boldsymbol{\Theta}$. Intuitively, one possibility is taking the vector of parameters that maximises the posterior probability, which is particularly suitable in situations with a unimodal Gaussian profile. That is,

$$\tilde{\boldsymbol{\theta}}_{\mathrm{MAP}}(x) = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|x). \tag{6}$$

This is the so-called maximum *a posteriori* (MAP) estimator [115], which plays a role similar to that of the maximum likelihood estimator in local QET [116].

Furthermore, to select a state $\rho(\boldsymbol{\theta})$ and a POVM element $M(x)$ and guide optimal experimental design, together with an estimator $\tilde{\boldsymbol{\theta}}(x)$ for post-processing the data, one requires a figure of merit that quantifies estimation performance and serves as the optimisation criterion. To that end, one typically introduces a loss function[1] $l(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}})$ quantifying the inaccuracy that may arise by choosing an estimate $\tilde{\boldsymbol{\theta}}$ should the hypothesis $\boldsymbol{\theta}$ correspond to the true parameter values [47]. An estimator, state, and POVM are then said to be optimal when they minimise the mean loss

$$\mathcal{L} \coloneqq \int \mathrm{d}\boldsymbol{\theta}\, \mathrm{d}x\, p(\boldsymbol{\theta})\ \mathrm{Tr}[M(x)\rho(\boldsymbol{\theta})]\, l(\tilde{\boldsymbol{\theta}}(x), \boldsymbol{\theta}), \quad (7)$$

where, for simplicity, we omit explicit reference to the domains Ω and χ. This convention will be followed henceforth.

This work evaluates measurement incompatibility within the Bayesian paradigm by establishing lower and upper bounds on a particular type of mean loss $\mathcal{L}$. The remainder of this section addresses formal aspects of $\mathcal{L}$, as well as the notion of metrological optimality based on this figure of merit, that are necessary to derive and contextualise our main results in Sec. 3.

2.2 Estimation performance

We first observe that Eq. (7) quantifies estimation performance by integrating over all hypothesis values in Ω, whose contribution can vary from small to large, weighted by the prior probability $p(\boldsymbol{\theta})$. This is the key feature that makes the framework not only global but also Bayesian. There exists a more conservative, non-Bayesian approach to global estimation: the minimax approach, which is based on the *frequentist* loss [6]

$$\mathcal{L}(\boldsymbol{\theta}) \coloneqq \int \mathrm{d}x\ \mathrm{Tr}[M(x)\rho(\boldsymbol{\theta})]\, l(\tilde{\boldsymbol{\theta}}(x), \boldsymbol{\theta}), \quad (8)$$

and seeks to minimise the worst-case value of $\mathcal{L}(\boldsymbol{\theta})$ over relevant regions of the parameter space [117]. However, since the minimax criterion can be written as the supremum of Eq. (7) over all priors, our bounds in Sec. 3 are also relevant for minimax estimation. Note that, at the other extreme, Eq. (8) is employed in local QET by evaluating it in an infinitesimal neighbourhood of specific values of $\boldsymbol{\theta}$ [6, 116].

We further comment that, in Bayesian estimation, the additional integral over x is required only when searching for optimal estimation protocols, as in our case, since this task precedes the recording of specific outcomes. For data post-processing, estimation performance must instead be assessed using an outcome-dependent quantifier [118], e.g.,

$$\mathcal{L}(x) \coloneqq \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta}|x)\, l(\tilde{\boldsymbol{\theta}}(x), \boldsymbol{\theta}). \quad (9)$$

Alternatively, one may employ credibility intervals, also based on the posterior [118, 119]. Both experimental error assessment and the minimax and local estimation approaches discussed above are beyond the scope of this work, and we therefore focus exclusively on Eq. (7).

So far we have not specified any loss function l. The most prevalent choice is

$$l(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}) = (\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta})^T L\, (\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}), \quad (10)$$

where L is a positive semidefinite weight matrix, i.e., $L \succeq 0$, and we assume it to be parameter-independent. The resulting $\mathcal{L}$ is the standard MSL, and this loss quantifier leads to the posterior mean (PM) as the optimal Bayes estimator [47]:

$$\tilde{\theta}_{\mathrm{opt},i}(x) = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta}|x)\, \theta_i = \frac{\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, p(x|\boldsymbol{\theta})\, \theta_i}{\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, p(x|\boldsymbol{\theta})}. \quad (11)$$

There is a widespread belief that the MSL applies universally, regardless of the specific system, measurement, or type of parameters of interest. Strictly speaking, however, the MSL is only appropriate for location or shift parameters. This is because the topology for such parameters corresponds to the real line, so that the corresponding loss function must be translation invariant [47], i.e.,

$$l(\tilde{\boldsymbol{\theta}} + \boldsymbol{\gamma}, \boldsymbol{\theta} + \boldsymbol{\gamma}) = l(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}), \quad (12)$$

as Eq. (10) is[2]. Here, $\boldsymbol{\gamma}$ is an arbitrary constant vector.

In contrast, parameter types with different properties require alternative loss functions. Cycle parameters, for instance, are estimated using trigonometric loss functions—phases in interferometry are the archetypical example [6, 9, 120]. Moreover, recent results show that scale parameters such as temperature or decay rate require logarithmic loss functions for consistent estimation [121], while weight parameters such as probability of success or coherence parameters demand hyperbolic loss functions [122]. Naive use of the MSL, while it may occasionally yield correct results—particularly when $\boldsymbol{\Theta}$ is known to lie close to a reference point $\boldsymbol{\Theta}_0$—is physically unjustified and therefore discouraged unless there is no better choice.

[1] This function is also often referred to as a cost, deviation, or error function. The term *loss* is preferred because it conveys a general notion of estimation inaccuracy, encompassing statistical, systematic, and context-dependent sources. In contrast, *deviation* and *error* typically emphasise purely statistical discrepancies, whereas *cost* may suggest application-specific considerations more typical of engineering contexts.

[2] While one can generally consider parameter-dependent weight matrices, these break the translation invariance of Eq. (10).

Despite this, one can retain the mathematical simplicity of the MSL while respecting parameters' more general topologies in certain scenarios. Specifically, Refs. [101, 123] demonstrated that many parameter types admit mean losses equivalent to the MSL, up to a mapping that, in the context of multiparameter estimation, can be expressed as

$$l(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}) = (\boldsymbol{f}(\tilde{\boldsymbol{\theta}}) - \boldsymbol{f}(\boldsymbol{\theta}))^T L\, (\boldsymbol{f}(\tilde{\boldsymbol{\theta}}) - \boldsymbol{f}(\boldsymbol{\theta})), \tag{13}$$

where $\boldsymbol{f}$ is chosen according to symmetries of the problem [122]; see Ref. [37] for an experimental demonstration with cold Cs atoms. Note that phases fall outside this framework [120], though they can be approximated by the MSL itself when confined to intervals of order $\pi/2$ [8, 13].

Our results in Sec. 3 are formulated in terms of the standard MSL for pedagogical reasons, but the extension to Eq. (13) is immediate. While a general multiparameter symmetry-informed treatment is deferred, Sec. 4.2 provides an illustration of how to use Eq. (13).

2.3 Optimal quantum strategies

We proceed by providing the conditions that minimise the MSL and tools to establish an operational notion of measurement incompatibility in Bayesian multiparameter quantum estimation. For a fixed measurement, the optimal classical estimator is the posterior mean (PM), already introduced in Eq. (11). In the quantum setting, however, one must also minimise the MSL over all generalised measurements, thereby making the optimisation substantially more challenging.

To ease the manipulation of multiparameter expressions, we first rewrite the MSL as

$$\mathcal{L} = \lambda + L^{ij}\ \mathrm{Tr}[\rho_0 M_{ji} - 2\bar{\rho}_j M_i]. \tag{14}$$

Here, λ is the weighted second moment of the prior probability,

$$\lambda := \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \boldsymbol{\theta}^T L\, \boldsymbol{\theta}, \tag{15}$$

which differs from the *a priori* MSL for nonvanishing prior mean, i.e., when $\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \theta_i \neq 0$ for at least some i; see Eq. (37) and App. A. The element L^{ij} denotes the ij-th component of L, and we employ the Einstein summation convention over indices $i, j = 1, \ldots, d$.

We further define, respectively, the following zeroth and first state moments:

$$\rho_0 := \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \rho(\boldsymbol{\theta}), \tag{16a}$$

$$\bar{\rho}_i := \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \rho(\boldsymbol{\theta})\, \theta_i, \tag{16b}$$

as well as the following first and second measurement moments:

$$M_i := \int \mathrm{d}x\, M(x)\, \tilde{\theta}_i(x) = \int \mathrm{d}\tilde{\boldsymbol{\theta}}\, M(\tilde{\boldsymbol{\theta}})\, \tilde{\theta}_i, \tag{17a}$$

$$M_{ij} := \int \mathrm{d}x\, M(x)\, \tilde{\theta}_i(x)\tilde{\theta}_j(x) = \int \mathrm{d}\tilde{\boldsymbol{\theta}}\, M(\tilde{\boldsymbol{\theta}})\, \tilde{\theta}_i\tilde{\theta}_j, \tag{17b}$$

where to obtain the rightmost expressions we have combined the estimators and the measurement by using the relabelled POVM elements [10]

$$M(\tilde{\boldsymbol{\theta}}) = \int \mathrm{d}x\, M(x)\, \delta(\tilde{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}(x)); \tag{18}$$

this will be useful in following calculations.

Given that $p(\boldsymbol{\theta})$ must be fixed *a priori*, and L is fixed according to the priority assigned to certain linear combinations or specific parameters, λ remains constant throughout. The degrees of freedom for optimising the MSL are thus $\rho(\boldsymbol{\theta})$ and either the pair $(\tilde{\boldsymbol{\theta}}(x), M(x))$ or, as we will consider from now on, $M(\tilde{\boldsymbol{\theta}})$. Since our focus is measurement incompatibility, and not the trade-offs that arise from the choice of probing strategy before the final measurement [124–126], we shall also assume that $\rho(\boldsymbol{\theta})$—and therefore ρ_0 and $\bar{\rho}_i$—is fixed.

The general conditions constraining the form of the optimal $M(\tilde{\boldsymbol{\theta}})$, denoted $M_{\mathrm{opt}}(\tilde{\boldsymbol{\theta}})$, have been established since the seminal works of Personick, Helstrom, Holevo, and others [9, 25–28, 127] in the 1970s. In App. B, we revisit these conditions and show that, for the MSL, they read

$$\begin{aligned} L^{ij}[\, &(\tilde{\theta}_j\tilde{\theta}_i\mathbb{I} - M_{\mathrm{opt},ji})\rho_0 + \rho_0(\tilde{\theta}_j\tilde{\theta}_i\mathbb{I} - M_{\mathrm{opt},ji}) \\ &- 2(\tilde{\theta}_j\mathbb{I} - M_{\mathrm{opt},j})\bar{\rho}_i - 2\bar{\rho}_j(\tilde{\theta}_i\mathbb{I} - M_{\mathrm{opt},i})] \succeq 0, \end{aligned} \tag{19}$$

for all $\tilde{\boldsymbol{\theta}}$. Here, $M_{\mathrm{opt},ij}$ and $M_{\mathrm{opt},i}$ are obtained by inserting $M_{\mathrm{opt}}(\tilde{\boldsymbol{\theta}})$ into Eqs. (17). This leads to the minimum mean loss

$$\mathcal{L}_{\min} = \mathcal{L}[M_{\mathrm{opt}}(\tilde{\boldsymbol{\theta}})], \tag{20}$$

i.e., $\mathcal{L}$ in Eq. (14) evaluated at the optimal POVM.

Eq. (19) constitutes our first result. In general, determining the explicit form of $M_{\mathrm{opt}}(\tilde{\boldsymbol{\theta}})$ from Eq. (19) is a nontrivial task. Nevertheless, these conditions provide a practical criterion for verifying whether a candidate POVM is truly optimal. Note that employing a symmetry-informed yet still quadratic loss function [Eq. (13)] yields optimality conditions of the same structure, with the parameter space remapped according to the vector symmetry function $\boldsymbol{f}$.

2.4 Lower bounds from optimal single-parameter strategies

While exact computation of $\mathcal{L}_{\min}$ is generally difficult, it admits a lower bound [10, 43, 128]

$$\mathcal{L}_{\min} \geq \lambda - L^{ij}\, \mathrm{Tr}[\rho_0 S_j S_i] =: \mathcal{L}_{\mathrm{SPM}}, \tag{21}$$

where $\{S_i\}$ are the SPM operators, defined as the solutions to the Lyapunov equations

$$\bar{\rho}_i = \frac{S_i\rho_0 + \rho_0 S_i}{2} = E_{\rho_0,\mathrm{S}}\left[S_i\right]. \tag{22}$$

For later convenience, we have introduced the anti-commutator map (superoperator)

$$E_{\rho,\mathrm{S}}[X] \coloneqq \frac{1}{2}\left(X\rho + \rho X\right), \tag{23}$$

which admits a compact expression $S_i = E^{-1}_{\rho_0,\mathrm{S}}[\bar{\rho}_i]$ for the SPMs, and thus

$$\mathcal{L}_{\mathrm{SPM}} = \lambda - L^{ij}\,\mathrm{Tr}[\bar{\rho}_j E^{-1}_{\rho_0,\mathrm{S}}[\bar{\rho}_i]]. \tag{24}$$

Here and in the rest of this work, we assume that this inverse exists, which is guaranteed for a full-rank ρ_0. When this is not the case, one simply needs to work in the relevant subspace spanned by the parametric family $\rho(\boldsymbol{\theta})$, which corresponds to the support of ρ_0[3].

We note that several names have been adopted in the recent literature for the operators $\{S_i\}$, mostly in analogy with the SLD operators that appear in local QET. We propose instead the term SPM operators, by analogy with the optimal classical Bayes estimator for each parameter in Eq. (11).

Formally, the PM estimator is the component-wise ratio of the first-moment probability vector $\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, p(x|\boldsymbol{\theta})\, \theta_i$ to the average probability vector $\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, p(x|\boldsymbol{\theta})$. Intuitively, if $p(x|\boldsymbol{\theta})$ is formally replaced by a density operator $\rho(\boldsymbol{\theta})$, there are infinitely many inequivalent ways to "divide" $\bar{\rho}_i$ by ρ_0. The SPMs correspond to the symmetric choice, where "multiplication by ρ_0" is implemented by the map $E_{\rho_0,\mathrm{S}}$. Infinitely many other choices are possible and become inequivalent whenever $\bar{\rho}_i$ and ρ_0 do not commute, leading to different lower bounds; see App. D. This intuitive reasoning was originally identified by Personick [25] and is semantically consistent with local QET, in which the logarithmic derivative of a probability distribution admits inequivalent operator generalisations in the quantum setting [112].

Crucially, the SPM bound has the additional property of being attainable in single-parameter estimation by projecting onto the eigenbasis of the single SPM operator available in that case [10, 25, 43, 133]. Consequently, upon performing the eigendecomposition

$$S_i = \sum_n |s_{i,n}\rangle\langle s_{i,n}|s_{i,n} \tag{25}$$

for the i-th operator, where we assume a discrete eigenbasis for simplicity but without loss of generality, the spectrum $\{s_{i,n}\}$ defines the estimator (see Refs. [25, 133] for details, including a continuous-variable formulation)

$$s_{i,n} = \int \mathrm{d}\theta\, p(\theta|s_{i,n})\, \theta = \tilde{\theta}_{\mathrm{opt},i}(n), \tag{26}$$

which naturally corresponds to the posterior mean (PM) estimator for the i-th parameter.

While the PM estimators remain optimal in the multiparameter case [43], the SPM bound is generally not attainable due to the non-commutativity of the SPM operators. Nevertheless, it provides a valid lower bound on the minimum MSL and can be interpreted as the optimum obtained when incompatibility among the individually optimal measurements is neglected. Accordingly, the next section introduces a notion of measurement incompatibility based on $\mathcal{L}_{\mathrm{SPM}}$, establishing it as a result of Bayesian estimation theory analogous in fundamental significance to the Helstrom (SLD) CRB in local QET [10].

3 Measurement incompatibility: definition and bounds

3.1 Incompatibility figure of merit

As we have just argued, the benchmark to evaluate the impact of measurement incompatibility is the lower bound on the MSL, $\mathcal{L}_{\mathrm{SPM}}$, which is achievable when all SPM operators share a common eigenbasis and thence are compatible observables. Therefore, measurement incompatibility in Bayesian multiparameter estimation can be quantified by comparing $\mathcal{L}_{\mathrm{SPM}}$ with the minimum MSL achievable by measurements, $\mathcal{L}_{\min}$.

We introduce a figure of merit for measurement incompatibility of a given Bayesian multiparameter estimation problem as follows:

$$\mathcal{I} \coloneqq \frac{\mathcal{L}_{\min}}{\mathcal{L}_{\mathrm{SPM}}} - 1 \geq 0\,, \tag{27}$$

where the inequality follows from the SPM lower bound in Eq. (21). A similar figure of merit has been studied also in the framework of local QET [64]. In the following sections, we shall derive upper and lower bounds on $\mathcal{I}$ by obtaining corresponding bounds on the unknown $\mathcal{L}_{\min}$.

3.2 Upper bounds

The incompatibility quantifier $\mathcal{I}$ can be upper bounded by considering specific POVMs and estimators, since the corresponding MSL is, by definition, never smaller than the achievable minimum.

[3] This is possible because the support of the operators $\{\bar{\rho}_j\}$ is contained in the support of ρ_0, since $\langle\psi|\rho_0|\psi\rangle = 0$ implies $\langle\psi|\rho(\boldsymbol{\theta})|\psi\rangle = 0$ almost everywhere (in the measure-theoretic sense), which in turn implies $\langle\psi|\bar{\rho}_j|\psi\rangle = 0$. This is markedly different from local QET, where the partial derivatives operators may have components also in the kernel of the state, and properties of rank-deficient models are non-trivial. For example, necessary and sufficient conditions for attainability of the SLD CRB are not fully characterised [91, 129–131] (except for pure states [80, 132])

3.2.1 Pretty good measurement

Taking inspiration from hypothesis testing, we introduce the square root measurement, also known as the pretty good measurement (PGM) [108–111]:

$$M_{\mathrm{PGM}}(\tilde{\boldsymbol{\theta}}) \coloneqq p(\tilde{\boldsymbol{\theta}})\, \rho_0^{-\frac{1}{2}}\, \rho(\tilde{\boldsymbol{\theta}})\, \rho_0^{-\frac{1}{2}}. \tag{28}$$

This measurement is usually introduced for discriminating among discrete hypotheses, where the Barnum–Knill theorem guarantees the "pretty good" performance [134]. Recently, this theorem has been extended to continuous hypotheses [135] and applied to prove that the MSL of the PGM is within a factor two of the optimal $\mathcal{L}_{\min}$. While this remarkable result establishes the usefulness of the PGM for parameter estimation, it does not shed light on the issue of measurement incompatibility, i.e., understanding the discrepancy between $\mathcal{L}_{\min}$ and $\mathcal{L}_{\mathrm{SPM}}$. This will be the precise topic of the next subsection. In the remainder of this subsection, we set the stage by deriving the expression for the MSL achieved by the PGM.

Before proceeding, it is important to emphasise that the PGM is a POVM with outcomes in the parameter space and therefore already incorporates a classical estimator, which may not be optimal. Nonetheless, it yields simple expressions for the MSL that are easy to interpret and compare with the baseline provided by the SPM bound. In Sec. 3.2.3 we will show how the PM estimator—the optimal choice—can achieve a better MSL, at the expense of having a more involved expression.

The MSL of the PGM can be computed from Eq. (14), by first evaluating the operators in Eqs. (17) as

$$M_{\mathrm{PGM},i} = \rho_0^{-\frac{1}{2}} \bar{\rho}_i \rho_0^{-\frac{1}{2}}, \tag{29a}$$

$$M_{\mathrm{PGM},ij} = \rho_0^{-\frac{1}{2}} \left[\int \mathrm{d}\tilde{\boldsymbol{\theta}}\, p(\tilde{\boldsymbol{\theta}})\, \rho(\tilde{\boldsymbol{\theta}})\, \tilde{\theta}_i \tilde{\theta}_j\right] \rho_0^{-\frac{1}{2}}. \tag{29b}$$

Noticing that, from cyclicity of the trace,

$$L^{ij}\, \mathrm{Tr}[\rho_0\, M_{\mathrm{PGM},ji}] = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \boldsymbol{\theta}^T L\, \boldsymbol{\theta} = \lambda, \tag{30}$$

since ρ_0 is cancelled by the two factors $\rho_0^{-\frac{1}{2}}$ and $\mathrm{Tr}[\rho(\tilde{\boldsymbol{\theta}})] = 1$, we obtain

$$\begin{aligned} \mathcal{L}_{\mathrm{PGM}} &= 2\left(\lambda - L^{ij}\, \mathrm{Tr}\left[\bar{\rho}_j \rho_0^{-\frac{1}{2}} \bar{\rho}_i \rho_0^{-\frac{1}{2}}\right]\right) \\ &= 2\left(\lambda - L^{ij}\, \mathrm{Tr}\left[\bar{\rho}_j E^{-1}_{\rho_0,\mathrm{sq}}[\bar{\rho}_i]\right]\right). \end{aligned} \tag{31}$$

In the second line, we have expressed the MSL in terms of the inverse of the map

$$E_{\rho,\mathrm{sq}}[X] \coloneqq \sqrt{\rho} X \sqrt{\rho}. \tag{32}$$

This map, like $E_{\rho,S}$ in Eq. (23), is one of the infinitely many inequivalent ways of introducing a superoperator representing "multiplication by ρ". This plurality is due to noncommutativity in quantum theory, and indeed for diagonal ρ and X, i.e., embedding classical probability into quantum objects, they all correspond to an element-wise multiplication of probability vectors. Such maps are fundamental mathematical objects in the theory of monotone metrics [112, 136, 137].

3.2.2 Fundamental incompatibility upper bound

A crucial result for us in the theory of monotone metrics is that the inner product obtained from the inverse of E^S_ρ is the minimal one, implying the inequality

$$\mathrm{Tr}(X E^{-1}_{\rho,\mathrm{S}}[X]) \leq \mathrm{Tr}(X E^{-1}_{\rho,\mathrm{sq}}[X]). \tag{33}$$

Notice that the two sides of the inequality are norms induced by different inner products between operators. Since the weight matrix is symmetric and positive semidefinite, it admits a (non-unique) decomposition as a sum of rank-one outer products: $L = \sum_k \boldsymbol{v}_k \boldsymbol{v}_k^T$ for some real vectors $\boldsymbol{v}_k$. This, together with linearity of $E^{-1}_{\rho_0,\mathrm{sq}}$, allows to rewrite the second term of Eq. (31) as the sum of norms of the Hermitian operators $v^i_k \bar{\rho}_i$:

$$L^{ij}\, \mathrm{Tr}\left[\bar{\rho}_j E^{-1}_{\rho_0,\mathrm{sq}}[\bar{\rho}_i]\right] = \sum_k \mathrm{Tr}\left[(v^i_k \bar{\rho}_i) E^{-1}_{\rho_0,\mathrm{sq}}[(v^j_k \bar{\rho}_j)]\right]. \tag{34}$$

Plugging this identity into the second term of Eq. (31) and applying the inequality (33) to each summand, we prove the first inequality in the following chain

$$2\mathcal{L}_{\mathrm{SPM}} \geq \mathcal{L}_{\mathrm{PGM}} \geq \mathcal{L}_{\min} \geq \mathcal{L}_{\mathrm{SPM}}, \tag{35}$$

where the second inequality holds by definition and the third is the SPM lower bound in Eq. (21), which neglects potential incompatibility effects. Thus, we have succeeded in bounding the maximal increase in the MSL due to measurement incompatibility, which can equivalently be written as an upper bound on the incompatibility figure of merit:

$$0 \leq \mathcal{I} \leq 1. \tag{36}$$

We remark that the upper bound in Eq. (35) is our second and one of our key results. Interestingly, this mirrors a well-known result in local QET: the HCRB is at most twice the Helstrom (SLD) CRB [78, 79]. Notice, however, that the HCRB is in general attainable only asymptotically with collective measurements on multiple copies. Hence, only the fundamental asymptotic impact of measurement incompatibility is limited by the upper bound on the HCRB [64, 93]. On the other hand, our result is obtained in a Bayesian estimation setting, and holds for a single shot.

3.2.3 Pretty good measurement with optimal estimator

While generally valid, the bounds derived above can be trivial, i.e., less informative than the *a priori* MSL (cf. App. A): $\mathcal{L}_{\mathrm{prior}} \leq \mathcal{L}_{\mathrm{PGM}}$, where

$$\mathcal{L}_{\mathrm{prior}} = \lambda - \boldsymbol{\mu}^T L\, \boldsymbol{\mu}, \quad \mu_i = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \theta_i. \tag{37}$$

Such a behaviour may appear counter-intuitive, since measuring a system is supposed to increase the information about the parameters' values. The intuition is re-established if one realises that this happens only because of a suboptimal choice of the classical estimator in the construction of the PGM bound. Similar situations may appear also in quantum state discrimination, i.e., for discrete hypotheses. Indeed, it has been shown that, for specific ensembles, guessing the hypothesis that maximises the prior probability is optimal and the probability of success cannot be improved upon by performing measurements on the system [138, 139].

Instead of viewing the PGM as a POVM already equipped with a built-in estimator—one not specifically tailored to our quadratic loss function—we now drop the interpretation of its outcomes as estimates. For clarity, the POVM elements in Eq. (28) can be rewritten with a distinct label for the measurement outcomes as

$$M_{\mathrm{PGM}}(\boldsymbol{x}) = p(\boldsymbol{x})\, \rho_0^{-\frac{1}{2}}\, \rho(\boldsymbol{x})\, \rho_0^{-\frac{1}{2}}, \tag{38}$$

where each $\boldsymbol{x}$ is still a vector in the hypothesis space Ω, though it is not directly used as an estimator. Instead, the measurement outcomes are processed through the optimal PM estimator in Eq. (26), which is constructed from the likelihood function $p(\boldsymbol{x}|\boldsymbol{\theta}) = \mathrm{Tr}[\rho(\boldsymbol{\theta})\, M_{\mathrm{PGM}}(\boldsymbol{x})]$. From a straightforward calculation, we obtain the MSL

$$\begin{aligned}\mathcal{L}_{\mathrm{PGM}^*} \coloneqq \lambda - L^{ij} \int \mathrm{d}\boldsymbol{x}\, p(\boldsymbol{x})\ \mathrm{Tr}\left[\rho(\boldsymbol{x})\rho_0^{-\frac{1}{2}}\bar{\rho}_j\rho_0^{-\frac{1}{2}}\right] \\ \cdot\, \mathrm{Tr}\left[\rho(\boldsymbol{x})\rho_0^{-\frac{1}{2}}\bar{\rho}_i\rho_0^{-\frac{1}{2}}\right].\end{aligned} \tag{39}$$

Contrary to the standard PGM, this quantity is guaranteed to be no greater than $\mathcal{L}_{\mathrm{prior}}$, which can be seen explicitly by applying Jensen's inequality to the second term of Eq. (39), since the function of $\rho(\boldsymbol{\theta})$ is quadratic and convex. Whilst for an arbitrary POVM with the same estimator—which is guaranteed to be smaller than $\mathcal{L}_{\mathrm{prior}}$—the MSL may be larger than $\mathcal{L}_{\mathrm{PGM}}$. Note that, as discussed, for our choice of quadratic loss we have a formal guarantee that the PM is the optimal classical estimator, but there is numerical evidence that it also leads to a better fidelity than the standard PGM approach in state estimation [140].

Tightening the previous inequalities, by taking into account that they may be trivial, and adding this new tighter upper bound, we obtain

$$\begin{aligned}\min[2\mathcal{L}_{\mathrm{SPM}},\, \mathcal{L}_{\mathrm{prior}}] &\geq \min[\mathcal{L}_{\mathrm{PGM}},\, \mathcal{L}_{\mathrm{prior}}] \\ &\geq \mathcal{L}_{\mathrm{PGM}^*} \\ &\geq \mathcal{L}_{\min} \\ &\geq \mathcal{L}_{\mathrm{SPM}}\,.\end{aligned} \tag{40}$$

The corresponding bounds on $\mathcal{I}$ are

$$\begin{aligned}0 &\leq \mathcal{I} \\ &\leq \mathcal{I}_{\mathrm{PGM}^*} \\ &\leq \min[\mathcal{I}_{\mathrm{PGM}},\, \mathcal{I}_{\mathrm{prior}}] \\ &\leq \min[1,\, \mathcal{I}_{\mathrm{prior}}]\,,\end{aligned} \tag{41}$$

where $\mathcal{I}_{\mathrm{PGM}}$, $\mathcal{I}_{\mathrm{PGM}^*}$, and $\mathcal{I}_{\mathrm{prior}}$ are defined as in Eq. (27). The first three inequalities in Eq. (40) (the last three inequalities in Eq. (41)) constitute our third result.

3.3 Lower bounds

Having established an upper bound on the MSL to characterise the worst-case impact of measurement incompatibility, as well as a tighter version, our aim is now to tighten the lower bound in Eq. (36) for a given model. To this end, we seek lower bounds on the MSL that are tighter than $\mathcal{L}_{\mathrm{SPM}}$ in the presence of non-commuting SPM operators. In this section, we focus primarily on the tightest known lower bound on the MSL available in the literature, the Nagaoka–Hayashi (NH) bound, $\mathcal{L}_{\mathrm{NH}}$, recently introduced in Bayesian QET by Suzuki [104].

3.3.1 Nagaoka–Hayashi bound

To derive this bound, one starts by noticing that any choice of POVM and estimators satisfies the following inequality [104, Lemma 1],

$$\sum_{ij} |i\rangle\langle j| \otimes M_{ij} \succeq \sum_{ij} |i\rangle\langle j| \otimes M_i M_j\,, \tag{42}$$

in terms of the first and second moment operators introduced in Eq. (17). Note that an additional vector space of dimension d (number of parameters) has been introduced, but this is not a physical Hilbert space. While we use the bra–ket notation, $|k\rangle$ simply stands for the standard column basis vector e_k of $\mathbb{R}^d$ (we have also temporarily reintroduced an explicit summation for extra clarity). The two sides of Eq. (42) are block matrices: the left-hand side can be interpreted as an operator-valued second moment matrix, while the right-hand side as an outer product of operator-valued first moment vectors.

Since the MSL in Eq. (14) is expressed only in terms of these operators, a lower bound is obtained by treating the Hermitian operators $M_{ij} = M_{ji}$ and M_i as *independent* variables satisfying Eq. (42):

$$\begin{aligned}\mathcal{L}_{\mathrm{NH}} \coloneqq \lambda + \min_{\{M_{ij}, M_i\}} L^{ij}\ \mathrm{Tr}[\rho_0 M_{ji} - 2\bar{\rho}_j M_i] \\ \text{subject to Eq. (42)}\,.\end{aligned} \tag{43}$$

This means that, for a given set of M_{ij} and M_i attaining the minimum, there is no guarantee that an underlying POVM $M(x)$ exists from which they can be derived, according to Eqs. (17). Consequently, the

bound may be loose. However, despite not *fully* capturing measurement incompatibility, the NH bound can be shown to be tighter than the SPM one. This follows by noticing that Eq. (42) implies

$$L^{ij}\,\mathrm{Tr}[\rho_0 M_{ji}] \geq L^{ij}\,\mathrm{Tr}[\rho_0 M_j M_i], \tag{44}$$

which is a weaker constraint. Applying this inequality to Eq. (43) one obtains a relaxation of the original problem that can be solved analytically to obtain $\mathcal{L}_{\mathrm{SPM}}$.

The bound $\mathcal{L}_{\mathrm{NH}}$ formally mirrors the NH Cramér-Rao bound (NHCRB) in local QET [90], and analogously it can be evaluated by solving a semidefinite program (SDP) [104]. Notice, however, that the upper bound in Eq. (35) that holds in the Bayesian case has no direct counterpart in the local case, since the NHCRB can be much larger than twice the Helstrom (SLD) CRB, and this gap can grow with the dimension of the Hilbert space and number of parameters [93].

For two parameters, the optimisation over the second moment operators M_{11}, M_{12} and M_{22} can be solved exactly. For the sake of clarity, but without loss of generality[4], we present this expression for an identity weight matrix:

$$\begin{aligned}\mathcal{L}_{\mathrm{NH}} \overset{L=\mathbb{I}_2}{=} \lambda + \min_{M_1,M_2}\{&\mathrm{Tr}\left[\rho_0(M_1^2+M_2^2)\right] \\ &+\|\sqrt{\rho_0}[M_1,M_2]\sqrt{\rho_0}\|_1 - 2\,\mathrm{Tr}\left[\bar{\rho}_1 M_1 + \bar{\rho}_2 M_2\right]\}\,,\end{aligned} \tag{45}$$

where $\|A\|_1 = \mathrm{Tr}[\sqrt{A^\dagger A}]$ denotes the trace norm, i.e., the sum of the singular values. This expression can be derived from Eq. (43), following, e.g., Ref. [90, Supplementary Note 2]. This form is the one originally obtained by Nagaoka for local two-parameter estimation [87].

Further considering a single qubit and full-rank ρ_0 (the only nontrivial qubit case), the NH bound in Eq. (45) is in fact attainable, i.e., $\mathcal{L}_{\mathrm{NH}} = \mathcal{L}_{\mathrm{min}}$, because Nagaoka's argument carries over to the Bayesian setting; we show this in App. C. The proof is constructive and yields an explicit optimal POVM, reported in App. C.1. Moreover, this case can be solved in closed form: the remaining optimisation over M_1, M_2 in Eq. (45) is evaluated in App. C.2.

Overall, we have the following inequalities:

$$\mathcal{L}_{\mathrm{min}} \geq \mathcal{L}_{\mathrm{NH}} \geq \mathcal{L}_{\mathrm{SPM}}\,. \tag{46}$$

In turn, this restricts the allowed range of the incompatibility figure of merit for a given model:

$$\mathcal{I}_{\mathrm{NH}} \leq \mathcal{I} \leq 1\,, \tag{47}$$

where $\mathcal{I}_{\mathrm{NH}}$ is defined analogously to Eq. (27). Moreover, by combining this NH bound on incompatibility with the tightest upper bound in Eq. (40), we obtain a tight range for the possible values of incompatibility in Bayesian multiparameter quantum estimation:

$$\mathcal{I}_{\mathrm{NH}} \leq \mathcal{I} \leq \mathcal{I}_{\mathrm{PGM}^*}\,. \tag{48}$$

This constitutes our fourth result.

3.3.2 Other lower bounds

We conclude this section by highlighting other lower bounds on the MSL which, although less tight—and mostly derived using techniques from local QET—can be easier to compute and may therefore remain useful in practice.

We first mention the Holevo bound, here denoted as $\mathcal{L}_{\mathrm{H}}$, which is a Bayesian version of the HCRB [104, 105]. It can be obtained by loosening the constraints in Eq. (42) taking the expectation value with respect to ρ_0, thereby obtaining a minimisation problem with an inequality constraint involving $d\times d$ matrices. The HCRB sits between the SPM and NH bounds: $\mathcal{L}_{\mathrm{NH}} \geq \mathcal{L}_{\mathrm{H}} \geq \mathcal{L}_{\mathrm{SPM}}$. While in local QET the HCRB has a fundamental meaning, being asymptotically attainable, the Bayesian version does not seem to share the same significance[5]. The Bayesian Holevo bound may be computationally more efficient because, while it can be computed by solving an SDP as the NH bound, it involves positive semidefinite constraints in terms of matrices of smaller dimension [104].

Pursuing further the goal of identifying easily computable bounds, one may instead seek lower bounds that capture measurement incompatibility to some extent and can always be expressed in closed form, rather than indirectly as the outcome of an optimisation. A whole class of such bounds exists, based on the idea of non-commutative PM operators, of which SPMs are just one example. This is in full analogy with logarithmic derivative operators in local QET [112, Ch. 6]. In App. D, we derive this family of lower bounds in a unified manner, employing the theory of monotone metrics, and generalising the approach of Ref. [43]. Here, we only mention the right posterior mean (RPM) bound

$$\mathcal{L}_{\mathrm{RPM}} := \lambda - L^{ij}\mathcal{K}_{\mathrm{R},ji} + \|\sqrt{L}\,\mathrm{Im}[\mathcal{K}_{\mathrm{R}}]\sqrt{L}\|_1, \tag{49a}$$

$$\mathcal{K}_{\mathrm{R},ij} := \mathrm{Tr}[\bar{\rho}_i \rho_0^{-1} \bar{\rho}_j], \tag{49b}$$

written in terms of the complex matrix $\mathcal{K}_{\mathrm{R}}$. Such a matrix is analogous to the right logarithmic derivative quantum Fisher information matrix in local QET. This bound is less informative than the NH or Holevo ones, since $\mathcal{L}_{\mathrm{H}} \geq \mathcal{L}_{\mathrm{RPM}}$ [105, Thm. 8]. However, it can be tighter or looser than $\mathcal{L}_{\mathrm{SPM}}$ depending on the specific model. For example, it is always less tight

[4] Similarly to local QET, the scalar NH bound can be recast to an equivalent problem with $L = \mathbb{I}$ and suitably redefined first moment operators $\bar{\rho}'_i = \sum_j (\sqrt{L})^{ij}\bar{\rho}_j$. The explicit expression for a general weight matrix is also reported in Ref. [104, Eq. (20)].

[5] Note that such a Bayesian Holevo-type bound is different from the so-called Bayesian CRBs obtained by averaging the HCRB over the prior [141], conjectured to be attainable in the large sample limit of Bayesian estimation [10, Sec. 5.5]; see the discussion in Sec. 5.

than the SPM bound for $d = 1$, where the latter is attainable. Conversely, Holevo has shown the RPM bound to be attainable for a multiparameter problem with non-commuting SPM operators: estimation of the average position and momentum of a Gaussian state with a Gaussian prior [127].

As a means of summary, we combine the upper bounds in Eq. (40) with all the lower bounds discussed in this section to obtain the following hierarchy of bounds on the minimum MSL:

$$\begin{aligned} \min[2\mathcal{L}_{\mathrm{SPM}},\, \mathcal{L}_{\mathrm{prior}}] &\geq \min[\mathcal{L}_{\mathrm{PGM}},\, \mathcal{L}_{\mathrm{prior}}] \\ &\geq \mathcal{L}_{\mathrm{PGM}^*} \\ &\geq \mathcal{L}_{\min} \\ &\geq \mathcal{L}_{\mathrm{NH}} \\ &\geq \mathcal{L}_{\mathrm{H}} \\ &\geq \max[\mathcal{L}_{\mathrm{SPM}},\, \mathcal{L}_{\mathrm{RPM}}]\,. \end{aligned} \tag{50}$$

Correspondingly, the hierarchy of bounds on measurement incompatibility is

$$\begin{aligned} \max[0,\, \mathcal{I}_{\mathrm{RPM}}] &\leq \mathcal{I}_{\mathrm{H}} \\ &\leq \mathcal{I}_{\mathrm{NH}} \\ &\leq \mathcal{I} \\ &\leq \mathcal{I}_{\mathrm{PGM}^*} \\ &\leq \min[\mathcal{I}_{\mathrm{PGM}},\, \mathcal{I}_{\mathrm{prior}}] \\ &\leq \min[1,\, \mathcal{I}_{\mathrm{prior}}]\,, \end{aligned} \tag{51}$$

where $\mathcal{I}_{\mathrm{RPM}}$ and $\mathcal{I}_{\mathrm{H}}$ are defined as in Eq. (27).

3.4 Approaching the minimum MSL numerically

While the bounds derived above enable efficient estimation of the incompatibility range, the exact value of the minimum MSL $\mathcal{L}_{\min}$ typically requires numerical optimisation over POVMs. A practical approach, introduced in Ref. [107] in the more general setting of channel estimation, is an iterative *see-saw* (alternating optimisation) method. Here, we briefly summarise the variant relevant to our setting of parameter-dependent families of states.

We restrict, without loss of generality, to POVMs with a finite number of outcomes, denoted as $\{M_n\}_{n=1}^N$; see the definition in Eq. (4), with associated estimator values $\{\tilde{\boldsymbol{\theta}}_n\}_{n=1}^N$. For the quadratic loss in Eq. (10)—or its symmetry-remapped version in Eq. (13)—the MSL can be written as a constant term depending only on the prior moments plus a term that is linear in the POVM elements for fixed estimators. Concretely, defining $\tilde{\theta}_{n,i}$ as the i-th component of the estimator associated with outcome n, one can rewrite

$$\mathcal{L} = \lambda - \sum_{n=1}^{N} \mathrm{Tr}\big[M_n\, X_n(\tilde{\boldsymbol{\theta}}_n)\big], \tag{52}$$

where each $X_n(\tilde{\boldsymbol{\theta}}_n)$ is a Hermitian operator that depends only on the zeroth and first state moments ρ_0 and $\{\bar{\rho}_i\}$ in Eqs. (16a) and (16b), and on the weight matrix L. The explicit expression is obtained by inserting the definition of the measurement moments for a discrete-outcome POVM $M_i = \sum_n \tilde{\theta}_{n,i} M_n$ and $M_{ij} = \sum_n \tilde{\theta}_{n,i}\tilde{\theta}_{n,j} M_n$ into Eq. (14), which yields

$$X_n(\tilde{\boldsymbol{\theta}}_n) = 2\, L^{ij}\, \tilde{\theta}_{n,i}\, \bar{\rho}_j - L^{ij}\, \tilde{\theta}_{n,i}\tilde{\theta}_{n,j}\, \rho_0. \tag{53}$$

Therefore, for fixed $\{\tilde{\boldsymbol{\theta}}_n^{(t)}\}$, where t labels each iteration, the optimisation over the measurement reduces to a SDP:

$$\max_{\{M_n\}} \sum_{n=1}^{N} \mathrm{Tr}\big[M_n\, X_n(\tilde{\boldsymbol{\theta}}_n^{(t)})\big] \quad \text{s.t.} \quad M_n \succeq 0,\ \sum_{n=1}^{N} M_n = \mathbb{I}. \tag{54}$$

This step is convex and can be solved to global optimality for the chosen number of outcomes N. For a D-dimensional system, at most $N = D^2$ outcomes are needed [107].

For fixed POVM $\{M_n^{(t)}\}$, obtained by solving the SDP above, the optimal classical post-processing for quadratic loss is the PM estimator in Eq. (11) and no optimisation is needed; for different loss functions, a second optimisation is generally needed [107]. In our notation, this update can be expressed directly in terms of the state moments:

$$\tilde{\theta}_{n,i}^{(t+1)} = \frac{\mathrm{Tr}\big[M_n^{(t)}\, \bar{\rho}_i\big]}{\mathrm{Tr}\big[M_n^{(t)}\, \rho_0\big]}, \tag{55}$$

and analogously after the symmetry remapping $\boldsymbol{\theta} \mapsto \boldsymbol{f}(\boldsymbol{\theta})$.

Iterating Eqs. (54) and (55) yields a sequence of feasible strategies whose MSL is non-increasing at every step. The procedure is not guaranteed to find the global optimum in general—the joint optimisation over POVMs and estimators is non-convex—but in practice it provides a constructive upper bound on $\mathcal{L}_{\min}$. To mitigate the risk of converging to suboptimal stationary points, one may perform multiple random initialisations of the estimators and keep the best solution. Finally, candidate solutions produced by the see-saw could also be checked against the optimality conditions in Eq. (19), providing an *a posteriori* consistency test.

This gives (as any POVM would) an upper bound on incompatibility ($\mathcal{I}_{\mathrm{SSW}}$). However, iteratively solving the SDPs is an additional layer of computational complexity, potentially requiring a scaling—in the number of POVM elements to solve for—with the square of the state-space dimension.

4 Applications

In the following sections we evaluate the measurement incompatibility range in Eq. (48), as well as some of the less tight bounds derived, for a range of examples.

4.1 Discrete quantum phase imaging

Consider a $(d{+}1)$-mode interferometer, where d parameters $\boldsymbol{\theta}$ for a set of unknown phases $\boldsymbol{\Theta}$ are encoded via the unitary operator $\mathrm{e}^{-iN^j\theta_j}$. Here, N^j is the number operator of the jth mode, and the zeroth mode serves as a reference beam [142]. This configuration models discrete quantum imaging and has been extensively explored in both local and Bayesian frameworks [41, 43, 67, 143, 144]. In this context, a common probe state is the n-photon generalised N00N state

$$|\psi(\theta)\rangle = \frac{1}{\sqrt{d+\alpha^2}}\left(\alpha e^{-in\theta_0}|n_0\rangle + \sum_{i=1}^{d} e^{-in\theta_i}|n_i\rangle\right), \tag{56}$$

with $\alpha \in \mathbb{R}$ and

$$|n_i\rangle := |0\rangle_0 \otimes \cdots \otimes |0\rangle_{i-1} \otimes |n\rangle_i \otimes |0\rangle_{i+1} \otimes \cdots \otimes |0\rangle_d \tag{57}$$

representing a state with n photons in the i-th mode, such that $\sum_j N^j |\psi(\theta)\rangle = n|\psi(\theta)\rangle$.

To address this problem using the MSL, we restrict the hypothesis space Ω so that the phases are sufficiently localised to behave as location parameters, but remaining within a global regime. Specifically, we employ a flat prior $p(\boldsymbol{\theta}) = \left(\frac{n}{2\pi}\right)^d$, with $\theta_i \in [-\pi/n, \pi/n]$ and $n \geq 4$, so that the largest prior width for the i-th parameter is $\pi/2$. For this configuration—and choosing a normalised L with equal weights, i.e., $L = \mathbb{I}/d$—the average state, SPM operators, and SPM bound are, respectively [13],

$$\rho_0 = \frac{1}{d+\alpha^2}\left(\alpha^2|n_0\rangle\langle n_0| + \sum_{i=1}^{d}|n_i\rangle\langle n_i|\right), \tag{58}$$

$$\bar{\rho}_j = \frac{-i\alpha}{n(d+\alpha^2)}\left(|n_j\rangle\langle n_0| - |n_0\rangle\langle n_j|\right), \tag{59}$$

$$S_j = \frac{-2i\alpha}{n(1+\alpha^2)}\left(|n_j\rangle\langle n_0| - |n_0\rangle\langle n_j|\right), \tag{60}$$

and

$$\mathcal{L}_{\mathrm{SPM}} = \frac{1}{n^2}\left[\frac{\pi^2}{3} - \frac{4\alpha^2}{(1+\alpha^2)(d+\alpha^2)}\right]. \tag{61}$$

Two observations are in order. First, the Heisenberg scaling in Eq. (61) is achievable in a global setting only because the chosen prior width scales as $\sim 1/n$. Indeed, since all first moments of the prior vanish for this problem, the *a priori* MSL is given by (cf. Eq. (37))

$$\mathcal{L}_{\mathrm{prior}} = \lambda = \frac{\pi^2}{3n^2}, \tag{62}$$

which already exhibits Heisenberg scaling; this is consistent with the observations in Refs. [12, 40, 145–148].

Secondly, and more importantly for our purposes, the SPM operators do not commute:

$$[S_j, S_k] = \frac{4\alpha^2}{n^2(1+\alpha^2)^2}(|n_j\rangle\langle n_k| - |n_k\rangle\langle n_j|). \tag{63}$$

As such, the SPM bound is not tight, and we can therefore expect incompatibility to arise in this system.

To quantify the amount of incompatibility, we first need to calculate the bounds on the minimum MSL. We start with the PGM bound, given in Eq. (31). Since

$$\bar{\rho}_j\,\rho_0^{-\frac{1}{2}} = \frac{-i}{n\sqrt{d+\alpha^2}}\left(|n_j\rangle\langle n_0| - \alpha|n_0\rangle\langle n_j|\right), \tag{64}$$

and so

$$\mathrm{Tr}[\bar{\rho}_j\,\rho_0^{-\frac{1}{2}}\,\bar{\rho}_i\,\rho_0^{-\frac{1}{2}}] = \frac{2\alpha}{n^2(d+\alpha^2)}\delta_{ji}, \tag{65}$$

Eq. (31) yields

$$\mathcal{L}_{\mathrm{PGM}} = \frac{2}{n^2}\left(\frac{\pi^2}{3} - \frac{2\alpha}{d+\alpha^2}\right). \tag{66}$$

However, this upper bound is trivial, i.e., $\mathcal{L}_{\mathrm{PGM}} > \mathcal{L}_{\mathrm{prior}}$, and consequently so is $2\mathcal{L}_{\mathrm{SPM}}$. This follows from showing that $\mathcal{L}_{\mathrm{PGM}} - \mathcal{L}_{\mathrm{prior}}$ is strictly positive by virtue of the inequality

$$\frac{\pi^2}{3} > \max_{\alpha}\frac{4\alpha}{d+\alpha^2} = \frac{2}{\sqrt{d}}, \tag{67}$$

valid for all d.

In view of this, we proceed to calculate the PGM bound enhanced by the PM estimator, given in Eq. (39), which is guaranteed to be smaller than or equal to the *a priori* MSL. First, calculate

$$\rho_0^{-\frac{1}{2}}\,\bar{\rho}_j\,\rho_0^{-\frac{1}{2}} = \frac{i}{n}(|n_0\rangle\langle n_j| - |n_j\rangle\langle n_0|), \tag{68}$$

so that

$$\mathrm{Tr}[\rho(\boldsymbol{\theta})\,\rho_0^{-\frac{1}{2}}\,\bar{\rho}_j\,\rho_0^{-\frac{1}{2}}] = \frac{2\alpha\sin(n\theta_j)}{n(d+\alpha^2)}. \tag{69}$$

Since the following integrals hold:

$$\int_{-\pi/n}^{\pi/n} \mathrm{d}\theta_j\,\sin(n\theta_j) = 0, \quad \int_{-\pi/n}^{\pi/n} \mathrm{d}\theta_j\,\sin^2(n\theta_j) = \frac{\pi}{n}, \tag{70}$$

Eq. (39) leads to

$$\mathcal{L}_{\mathrm{PGM}^*} = \frac{1}{n^2}\left(\frac{\pi^2}{3} - \frac{2\alpha^2}{(d+\alpha^2)^2}\right). \tag{71}$$

This is our tightest upper bound on the minimum MSL.

We next turn to the lower bounds. We begin with the simplest tractable bound beyond the SPM bound, namely the RPM bound in Eq. (49a). Using

$$\bar{\rho}_j\,\rho_0^{-1}\,\bar{\rho}_k = \frac{1}{n^2(d+\alpha^2)}\left(|n_j\rangle\langle n_k| + \alpha^2\delta_{jk}|n_0\rangle\langle n_0|\right), \tag{72}$$

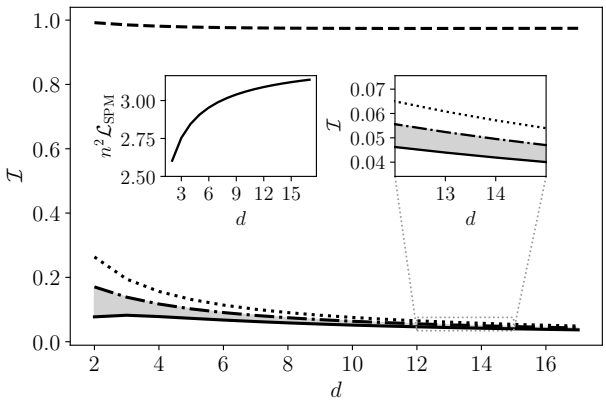

Figure 1: Measurement incompatibility bounds from the NH bound (solid), the PGM bound enhanced by the PM estimator (dash–dot), the standard PGM bound (dashed), and the *a priori* MSL (dotted) for a discrete quantum phase imaging protocol. The shaded region represents the range of incompatibility, as defined in Eq. (48), as a function of the number of parameters d, using a generalised N00N state with $\alpha = d^{1/4}$ as the initial state. The first inset displays the reference MSL, $\mathcal{L}_{\rm SPM}$. The incompatibility associated with the *a priori* MSL decreases with the number of parameters, thereby fundamentally limiting the extent to which measurement incompatibility can influence the overall precision; a zoomed view for larger parameter values is shown in the second inset. Furthermore, the incompatibility for the standard PGM bound is considerably larger than its *a priori* MSL counterpart and thus trivial for this system.

we obtain

$$\mathcal{K}_{{\rm R},jk} = \frac{1+\alpha^2}{n^2(d+\alpha^2)}\,\delta_{jk}, \tag{73}$$

and therefore

$$\mathcal{L}_{\rm RPM} = \frac{1}{n^2}\left(\frac{\pi^2}{3} - \frac{1+\alpha^2}{d+\alpha^2}\right). \tag{74}$$

However, it is straightforward to verify that $\mathcal{L}_{\rm SPM} \geq \mathcal{L}_{\rm RPM}$, consistent with the fact that $\mathcal{K}_{\rm R}$ in Eq. (73) is real-valued. Consequently, we shall instead evaluate the NH bound in Eq. (43) using the numerical algorithm of Ref. [113] to obtain a tighter lower bound on the minimum MSL.

We are now in a position to quantify measurement incompatibility. Building on the preceding results, Fig. 1 illustrates as a shaded region the range defined in Eq. (48), showing its dependence on the number of parameters d and using the state with $\alpha = d^{1/4}$ that minimises Eq. (61). The upper boundary of this region is given by $\mathcal{I}_{\rm PGM^*}$ (dashed-dot line), while the lower boundary corresponds to $\mathcal{I}_{\rm NH}$ (solid line). For comparison, we also include the incompatibility quantifier associated with the *a priori* MSL (dotted line) and the PGM bound (dashed line). As can be seen, the latter exceeds the former and is therefore trivial, as discussed above. Finally, the reference value $\mathcal{L}_{\rm SPM}$ is shown in the first inset, likewise plotted as a function of d.

Fig. 1 confirms the presence of measurement incompatibility within this protocol, with values not exceeding 0.2. This relatively low magnitude can initially be attributed to the proximity between $\mathcal{L}_{\rm prior}$ and $\mathcal{L}_{\rm SPM}$, evidenced by the dotted line. That is, within the Bayesian framework, the prior information imposes a finite and natural upper bound on the MSL, thereby limiting the maximum attainable relative precision gain. This effect becomes especially evident as the number of parameters increases, leading to a progressive reduction in relative precision gain and leaving increasingly limited scope for incompatibility to exert a noticeable influence on the precision. Nevertheless, the second inset demonstrates that the protocol consistently exhibits a nonzero degree of measurement incompatibility.

Here we see a deviation from the results for this model in local QET [41, 132], where the SLD bound is attainable—and no incompatibility is seen—due to the commuting generators which give rise to commuting SLD operators for pure state probes. With pure states leading to a mixed-state quantum statistical model in the Bayesian framework, the corresponding SPM operators do not directly commute and incompatibility is found as the SPM bound is unobtainable. Nevertheless that incompatibility is small, and numerical evidence using the method described in Sec. 3.4 suggests it is at the lower end of the incompatibility window shown in Fig. 1 as the NH bound appears attainable with $|\mathcal{I}_{\rm NH} - \mathcal{I}_{\rm SSW}| \leq 3\times10^{-4}$ for $2 \leq d \leq 10$ [149].

Symmetries, including the commuting generators, of the physical configuration may yet be responsible for both the low incompatibility, and apparent numeric saturability of the NH bound. It is perhaps informative to note that, while lacking the operational meaning it has in local QET [124], a Bayesian analogue of the so-called weak commutativity,

$$\mathrm{Tr}\,[\rho_0[S_j, S_k]] = 0, \tag{75}$$

does still hold in this case, since

$$\rho_0[S_j, S_k] = \frac{4\alpha^2}{n^2(1+\alpha^2)^2(d+\alpha^2)}\big(|n_j\rangle\langle n_k| - |n_k\rangle\langle n_j|\big). \tag{76}$$

4.2 Phase and dephasing estimation

While phase estimation serves as the archetypal problem in quantum estimation, the simultaneous estimation of phase and phase diffusion plays an analogous role in multiparameter quantum estimation. The simplest single-qubit incarnation of this model exhibits measurement incompatibility in local QET [150], which however asymptotically disappears as many identical copies are measured collectively [151]. This corresponds to preparing identical and factorised probe states, whereas asymptotically large entangled

probes generally allow to optimally estimate both parameters [152]. Nevertheless, we consider here independent identical copies collectively measured, since this is a paradigmatic setting for investigating measurement incompatibility [150, 151, 153].

For an initial state ρ_{in}, the single-qubit phase and phase-diffusion channels[6], which commute, are defined as

$$\Lambda_1(\rho_{\text{in}}) = \mathrm{e}^{-i\frac{\theta_1}{2}\sigma_z}\,\rho_{\text{in}}\,\mathrm{e}^{i\frac{\theta_1}{2}\sigma_z}, \tag{77a}$$

$$\Lambda_2(\rho_{\text{in}}) = \int_{-\infty}^{\infty} \mathrm{d}x\,\frac{\mathrm{e}^{-\frac{x^2}{4\theta_2^2}}}{2\sqrt{\pi}\,\theta_2}\,\mathrm{e}^{-i\frac{x}{2}\sigma_z}\,\rho_{\text{in}}\,\mathrm{e}^{i\frac{x}{2}\sigma_z}, \tag{77b}$$

respectively; σ_x, σ_y and σ_z denote Pauli matrices. The corresponding parameter-dependent state is

$$\rho(\boldsymbol{\theta}) = \frac{1}{2}(\mathbb{I}_2 + \cos\theta_1 \mathrm{e}^{-\theta_2^2}\sigma_x - \sin\theta_1 \mathrm{e}^{-\theta_2^2}\sigma_y)\,. \tag{78}$$

The MSL is not suitable for this problem. Specifically, while $\theta_1 \in [-\frac{W_1}{2}, \frac{W_1}{2}]$ is a cyclic variable that can be treated as a location parameter when $W_1 \lesssim \pi/2$—see Sec. 2.2—$\theta_2 > 0$ in Eq. (77b) is a *scale* parameter [47, 115], for which the appropriate loss function is logarithmic [133]. Consequently, one must employ a symmetry-informed loss function of the kind in Eq. (13).

A particular realisation, independently enforcing translation invariance for θ_1 and scale invariance for θ_2, is given by

$$l(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}) = (\tilde{\theta}_1 - \theta_1)^2 + \log^2\left(\frac{\tilde{\theta}_2}{\theta_2}\right), \tag{79}$$

which mixes two uncertainty types [37]. Here, we assume $L = \mathbb{I}_2$. This corresponds to using the multi-parameter symmetry function

$$\boldsymbol{f}(\boldsymbol{\theta}) = \begin{pmatrix} f_1(\boldsymbol{\theta}) \\ f_2(\boldsymbol{\theta}) \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \log(\theta_2) \end{pmatrix}. \tag{80}$$

Crucially, given that Eq. (79) preserves the quadratic structure of the square loss (10), the bounds derived in this work remain unchanged, except for the definitions of the prior moments,

$$\boldsymbol{\mu} = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \boldsymbol{f}(\boldsymbol{\theta}), \tag{81a}$$

$$\lambda = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \boldsymbol{f}(\boldsymbol{\theta})^T L\, \boldsymbol{f}(\boldsymbol{\theta}); \tag{81b}$$

the first state moments,

$$\bar{\rho}_i = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \rho(\boldsymbol{\theta})\, f_i(\boldsymbol{\theta}); \tag{82}$$

and the measurement moments,

$$M_i = \int \mathrm{d}\tilde{\boldsymbol{\theta}}\, M(\tilde{\boldsymbol{\theta}})\, f_i(\tilde{\boldsymbol{\theta}}), \tag{83a}$$

$$M_{ij} = \int \mathrm{d}\tilde{\boldsymbol{\theta}}\, M(\tilde{\boldsymbol{\theta}})\, f_i(\tilde{\boldsymbol{\theta}}) f_j(\tilde{\boldsymbol{\theta}}). \tag{83b}$$

Following Ref. [37], Eq. (80) further leads to the ignorance prior [47]

$$p(\theta_1, \theta_2) = \frac{1}{\theta_2 W_1 \log(W_2)} \propto \frac{1}{\theta_2}, \tag{84}$$

with $\theta_1 \in [-W_1/2, W_1/2]$ and $\theta_2 \in [1/\sqrt{W_2}, \sqrt{W_2}]$, for which the corresponding *a priori* MSL reads

$$\mathcal{L}_{\text{prior}} = \lambda = \frac{1}{12}\left[W_1^2 + \log^2(W_2)\right]. \tag{85}$$

Eq. (84) is the product of two independent priors—the first flat, representing maximum ignorance for a location, and the second Jeffreys's prior, representing maximum ignorance for a scale [47, 115]. The range for the dephasing parameter takes such a form when a symmetric, flat range of width W_2 is imposed in the log-transformed space.

Given this formulation, the state moments for a single copy can be calculated to have form

$$\rho_0 = \frac{1}{2}[\mathbb{I}_2 + g(W_1, W_2)\,\sigma_x], \tag{86a}$$

$$\bar{\rho}_1 \propto \sigma_y, \tag{86b}$$

$$\bar{\rho}_2 \propto \sigma_x, \tag{86c}$$

with the exact expressions given in App. E. Although this is a qubit system amenable to further analytic treatment, the resulting expressions are not very informative. For our purposes, it suffices to note that the SPM operators arising from Eqs. (183a) have the following structure:

$$S_1 \propto \sigma_y, \tag{87a}$$

$$S_2 \propto \mathbb{I}_2 + h(W_1, W_2)\,\sigma_x. \tag{87b}$$

As can be observed, the SPM operators do not generally commute, thereby guaranteeing the presence of measurement incompatibility at the single-copy level. However, the Bayesian weak commutativity condition proposed in Eq. (75) does hold (as the weak commutativity condition does for the local setting [150]). This also holds for the many-copy version [149], which we consider next.

Let $\rho^{\otimes n}(\boldsymbol{\theta})$ be the n-copy state. Due to the prior, the corresponding n-copy average state and first moments, i.e., $\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \rho^{\otimes n}(\boldsymbol{\theta})$ and $\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \rho^{\otimes n}(\boldsymbol{\theta})\, \theta_j$, are not directly tensor powers of the single-copy ρ_0 and $\bar{\rho}_j$ in Eqs. (183a).

In lieu of the PGM enhanced by PM estimator bound (due to the high dimension of the n-copy state), we consider a local qubit tomography—where each copy is measured at random in the basis of one of the

[6] The qubit can represent, e.g., a single-photon in optical interferometry, whereas the same setup for probe states with more photons is not equivalent to considering multiple copies of the single-photon state, but give rise to correlated dephasing [154, 155].

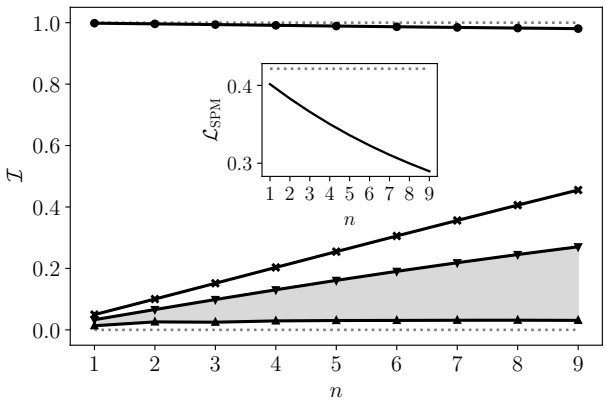

Figure 2: Measurement incompatibility bounds from the NH bound (upward triangles), the MSL obtained from a tomographic measurement (downward triangles), the standard PGM bound (circles), and the *a priori* MSL (crosses) for n copies of the qubit state (78), with prior widths $W_1 = \pi/2$ and $W_2 = 5$. The shaded region represents the range of incompatibility defined in Eq. (48) as a function of n. The upper and lower bounds on measurement incompatibility Eq. (36) (dotted) are represented for reference. The inset shows the reference MSL, $\mathcal{L}_{\rm SPM}$, which decreases as the number of copies increases, as well as the *a priori* MSL, which remains constant, leading the corresponding incompatibility to increase. The incompatibility associated with the standard PGM bound (circles), is, as in Sec. 4.1, is substantially larger than its *a priori* MSL counterpart and therefore again trivial.

three Pauli operators. The outcomes are then processed with the optimal PM estimator, such that the resulting MSL lower bounds $\mathcal{L}_{\rm prior}$, which is again the tighter bound than $\mathcal{L}_{\rm PGM}$.

In Fig. 2 we present numerical results on measurement incompatibility, by evaluating upper and lower bounds as a function of the number of copies, for prior widths $W_1 = \pi/2$ and $W_2 = 5$. The main plot shows that the incompatibility window grows with the number of copies, with the incompatibility NH lower bound increasing slightly, and the tomography upper bound increasing more rapidly. The latter result may be expected, since tomography corresponds to measuring each copy separately and thus cannot take advantage of collective effects.

Interestingly, the overall behaviour appears to be qualitatively different from local QET, where the initially large incompatibility decreases by measuring multiple copies collectively [150, 151]. In the asymptotic limit we would expect that decrease to manifest, with some eventual reduction in incompatibility. This is reinforced by the true incompatibility appearing to be at the lower end, with the NH bound being tight for $n = 1$ following from App. C, and appearing attainable numerically by the methods of Sec. 3.4 with $|\mathcal{I}_{\rm NH} - \mathcal{I}_{\rm SSW}| \leq 3{\times}10^{-3}$ for $n \leq 4$ [149].

Finally, we note that this growth in incompatibility, which may not represent the asymptotic behaviour, does come as the overall estimation precision increases, i.e., $\mathcal{L}_{\rm min}$ and its upper and lower bounds themselves decrease in the number of copies; see the plot of $\mathcal{L}_{\rm SPM}$ versus number of copies in the inset.

4.3 Qubit planar tomography

We conclude by analysing a configuration that, unlike the previous two examples, displays an *a priori* MSL larger than twice the SPM bound—the quantity leading to the upper bound on measurement incompatibility in Eq. (36). Before proceeding, we note that the purpose of this application is to illustrate specific mathematical properties of the theoretical bounds introduced, rather than to address a problem of physical relevance. For simplicity, we shall therefore use the MSL as a loss quantifier, notwithstanding the symmetry arguments discussed in Secs. 2.2 and 4.2.

Consider the following family of single-qubit states:

$$\rho(\boldsymbol{\theta}) = \frac{1}{2}\left(\mathbb{I}_2 + \theta_1\sigma_x + \theta_2\sigma_y\right), \tag{88}$$

where the unknown parameters satisfy the constraint $\theta_1^2 + \theta_2^2 \leq 1$ to ensure physicality. This setting corresponds to states lying on the plane $z = 0$ of the Bloch sphere, hence the term *planar tomography*. We assume independent, but not identical, prior distributions for the two parameters, i.e., $p(\boldsymbol{\theta}) = p_1(\theta_1)\, p_2(\theta_2)$, where $p_1(\theta_1)$ and $p_2(\theta_2)$ are supported on $[-W_1, W_1]$ and $[-W_2, W_2]$, respectively. Overall, we have a bivariate distribution over a rectangle contained inside the Bloch circle in the $x - y$ plane, ensured by the condition $W_1^2 + W_2^2 \leq 1$. Finally, we assume that the prior for each parameter has zero mean and a finite second moment:

$$\int \mathrm{d}\theta_i\, p_i(\theta_i)\, \theta_i = 0, \quad v_i := \int \mathrm{d}\theta_i\, p_i(\theta_i)\, \theta_i^2 \leq \infty. \tag{89}$$

Under these assumptions, and choosing $L = \mathbb{I}_2$, the *a priori* MSL is

$$\mathcal{L}_{\rm prior} = \lambda = v_1 + v_2\,, \tag{90}$$

while the zeroth and first state moments are

$$\rho_0 = \frac{\mathbb{I}}{2}, \quad \bar{\rho}_i = \frac{v_i}{2}\sigma_i, \tag{91}$$

where for simplicity we label Pauli matrices as $\sigma_1 \equiv \sigma_x$ and $\sigma_2 \equiv \sigma_y$. Given that ρ_0 is proportional to the identity, the SPM operators are trivially obtained:

$$S_i = 2\bar{\rho}_i = v_i\sigma_i. \tag{92}$$

Since they do not commute; this estimation problem is impacted by measurement incompatibility. However, the Bayesian weak commutativity condition proposed in Eq. (75) does again hold.

The SPM lower bound on the MSL can next be calculated as

$$\begin{aligned}\mathcal{L}_{\mathrm{SPM}} &= \lambda - \sum_i \mathrm{Tr}\left[\rho_0 S_i^2\right] \\ &= \lambda - \frac{1}{2}\sum_i v_i^2\, \mathrm{Tr}\left[\sigma_i^2\right] \\ &= v_1 + v_2 - \left(v_1^2 + v_2^2\right).\end{aligned} \tag{93}$$

The NH bound will be calculated numerically, as before.

Moving now to the upper bounds, in this scenario we have $\mathcal{L}_{\mathrm{PGM}} = 2\mathcal{L}_{\mathrm{SPM}}$, since ρ_0 commutes with the first state moments, and the symmetric and square-root inner products are therefore identical. Unlike in the two previous applications, here we find that the upper bound $2\mathcal{L}_{\mathrm{SPM}}$ is *not* trivial when

$$\mathcal{L}_{\mathrm{prior}} > 2\mathcal{L}_{\mathrm{SPM}} \iff 2\left(v_1^2 + v_2^2\right) > v_1 + v_2. \tag{94}$$

Nevertheless, in order to have a tighter upper bound, we also calculate the PGM with the optimal estimator in Eq. (39). Since here such a bound depends on the following quantity:

$$\mathrm{Tr}\left[\rho(\boldsymbol{\theta})\,\rho_0^{-\frac{1}{2}}\bar{\rho}_i\rho_0^{-\frac{1}{2}}\right]^2 = \left(\frac{v_i}{2}\,\theta_i\ \mathrm{Tr}[\sigma_i^2]\right)^2 = v_i^2\theta_i^2, \tag{95}$$

the integrals over θ_i with respect to the prior in Eq. (39) yield a further multiplication by the second moments v_i, giving the final expression

$$\mathcal{L}_{\mathrm{PGM}^*} = v_1 + v_2 - \left(v_1^3 + v_2^3\right). \tag{96}$$

We now show a particular choice that satisfies the inequality (94). Let each independent prior be based on a symmetric beta distribution rescaled and translated such that it is supported on the interval $[-W_i, W_i]$, and having zero mean as required:

$$p_i(\theta_i) = \frac{1}{2W_i} h_\beta\left(\frac{\theta_i}{2W_i} + \frac{1}{2}\right), \tag{97}$$

where

$$h_\beta(z) = \frac{(1-z)^{\beta-1} z^{\beta-1}}{B(\beta,\beta)}, \tag{98}$$

$\beta > 0$ controls the shape,

$$\begin{aligned}B(z_1, z_2) &= \int_0^1 \mathrm{d}t\, t^{z_1-1}(1-t)^{z_2-1} \\ &= \frac{\Gamma(z_1)\,\Gamma(z_2)}{\Gamma(z_1+z_2)}\end{aligned} \tag{99}$$

is the beta function, and $\Gamma(z)$ is the gamma function. For such an *a priori* assignment, the variances are

$$v_i = \frac{W_i^2}{1+2\beta}. \tag{100}$$

Therefore, by appropriately adjusting the parameters $\{W_i\}$ and β, one can find regions where the inequality (94) is satisfied.

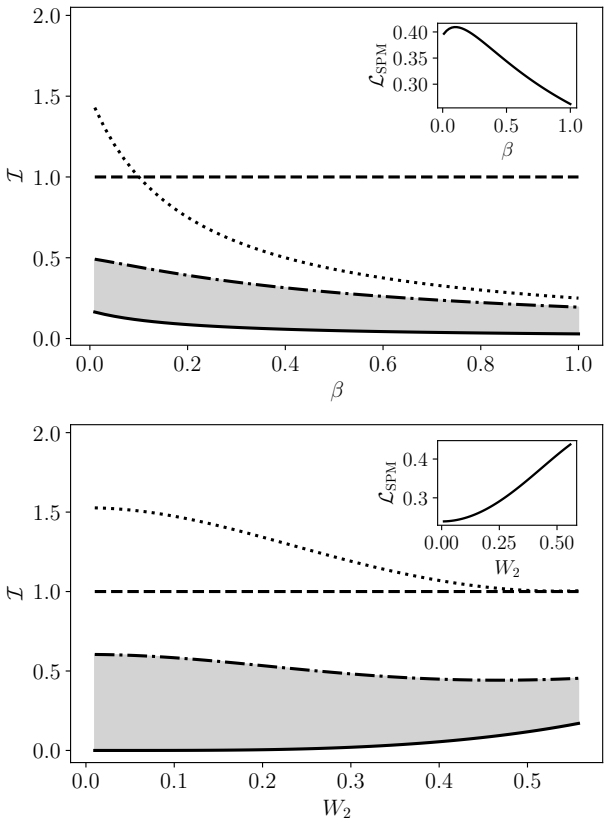

Figure 3: Measurement incompatibility bounds from the NH bound (solid), the PGM bound enhanced by the PM estimator (dash–dot), the standard PGM bound (dashed), and the *a priori* MSL (dotted). Top panel: Measurement incompatibility in qubit planar tomography as a function of the shape parameter β of the beta distribution (97), for fixed prior widths $W_1 = 0.85$ and $W_2 = 0.51$.The shaded region represents the range of incompatibility, as defined in Eq. (48). The inset shows the reference MSL, $\mathcal{L}_{\mathrm{SPM}}$. Bottom panel: Analogous configuration but represented as a function of the prior width W_2 for the parameter θ_2, for fixed shape parameter $\beta = 0.07$ and prior width $W_1 = 0.83$ for the parameter θ_1. Notably, the standard PGM bound is nontrivial in some regimes of these examples, with the prior incompatibility exceeding 1.

The incompatibility bounds corresponding to the MSL upper and lower bounds discussed above are plotted in Fig. 3, for various choices of the prior parameters W_1, W_2 and β. The graphical conventions distinguishing curves follow those established in Sec. 4.1. The region of incompatibility values permitted by the bounds is shaded in Fig. 3, however, for this class of models the NH bound is exactly attainable (see App. C), in analogy with local QET [88].

From these plots we observe that, when β is sufficiently small—roughly speaking, the prior distribution becomes more concentrated towards the edges—the upper bound $\mathcal{I} \leq \mathcal{I}_{\mathrm{PGM}} = 1$ becomes nontrivial as $\mathcal{I}_{\mathrm{prior}}$ can exceed the value of 1. In particular, in the bottom panel of Fig. 3 we show that for $\beta = 0.07$ and $W_1 = 0.83$ the inequality $\mathcal{I} \leq 1$ is not trivial for the whole range of valid values for the prior width

$0 < W_2 < \sqrt{1 - W_1^2}$, such that $W_2^2 + W_1^2 < 1$.

5 Discussion

One important aspect that we have not addressed is the relationship between Bayesian QET and the bounds derived in local QET for the frequentist loss in Eq. (8). In this regard, we recall a conjecture of Ref. [10]: under appropriate regularity conditions, the Bayesian MSL and the frequentist loss would become asymptotically equivalent, that is,

$$\lim_{n\to\infty} n\mathcal{L}_{\min}^{(n)} = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \mathcal{L}_{\mathrm{H}}(\boldsymbol{\theta}), \tag{101}$$

where $\mathcal{L}_{\min}^{(n)}$ is the MSL associated with n copies of a given state, $\rho^{\otimes n}(\boldsymbol{\theta})$, and prior $p(\boldsymbol{\theta})$, while $\mathcal{L}_{\mathrm{H}}(\boldsymbol{\theta})$ is the HCRB of local QET[7]. This expression would represent the minimum achievable frequentist loss in the asymptotic limit of large n.

Since in local QET it holds that [78, 79]

$$2\mathcal{L}_{\mathrm{S}}(\boldsymbol{\theta}) \geq \mathcal{L}_{\mathrm{H}}(\boldsymbol{\theta}) \geq \mathcal{L}_{\mathrm{S}}(\boldsymbol{\theta}), \tag{102}$$

where $\mathcal{L}_{\mathrm{S}}(\boldsymbol{\theta})$ is the Helstrom (SLD) CRB, our inequality in Eq. (35) is consistent with this asymptotic conjecture. Notice that, by definition, the Bayesian MSL $\mathcal{L}_{\min}^{(n)}$ is obtained by minimising over POVMs acting collectively on n copies, dubbed "correlated measurements" in Ref. [95].

Despite increasing efforts in the experimental implementation of few-copy collective measurements [151, 153, 156–160], scaling them to large n remains challenging in practice. Following this observation, recent work in local QET has addressed the minimum multiparameter frequentist loss achievable with single-copy measurements and classical feedback, referred to as "uncorrelated measurements" in Ref. [95]. This quantity corresponds to a minimisation of the weighted trace of the inverse Fisher information matrix (FIM) over all single-copy POVMs,

$$\mathcal{L}_{\min}(\boldsymbol{\theta}) \coloneqq \min_{M(\boldsymbol{\theta})} \mathrm{Tr}[LF(\boldsymbol{\theta})^{-1}], \tag{103}$$

where $F(\boldsymbol{\theta})$ is the FIM obtained by applying a POVM to a single copy $\rho(\boldsymbol{\theta})$. This is often called the "most informative" [87] or "tight" [95] bound.

Unlike the HCRB, $\mathcal{L}_{\min}(\boldsymbol{\theta})$ *can* be larger than $2\mathcal{L}_{\mathrm{S}}(\boldsymbol{\theta})$ [93, 95]. This appears to be in contrast with the Bayesian setting, where the inequality $2\mathcal{L}_{\mathrm{SPM}} \geq \mathcal{L}_{\min}$ holds at the single-copy level as well. The resolution of this apparent tension comes from noticing that the $\mathcal{L}_{\min}(\boldsymbol{\theta})$ is still attainable only asymptotically, but restricting each copy to be measured independently, albeit adaptively with classical feedback from previous outcomes. Therefore, we expect that imposing an analogous restriction in the Bayesian case should lead to a similar behaviour as the number of shots increases. More concretely, we conjecture that, under suitable regularity conditions

$$\lim_{n\to\infty} n\mathcal{L}_{\min}^{(n),\mathrm{uncorr}} = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\mathcal{L}_{\min}(\boldsymbol{\theta})\,, \tag{104}$$

where $\mathcal{L}_{\min}^{(n),\mathrm{uncorr}}$ would be the MSL for the prior $p(\boldsymbol{\theta})$ and family of states $\rho^{\otimes n}(\boldsymbol{\theta})$, but minimised over the uncorrelated measurement strategies defined above, instead of arbitrary collective (correlated) POVMs over all the copies.

We conclude by noting that hybrid figures of merit, incorporating methods from both local and global QET, have been introduced [21, 161]. However, these approaches are mostly heuristic, and understanding the connections outlined in this discussion could clarify their exact range of applicability.

6 Conclusions and outlook

QET was initiated in the 1960s and 1970s, but only in recent years have several lines of research converged, leading to a deeper understanding of multiparameter aspects [10, 67, 69]. Building on this progress, in this work we have laid the foundations for a theory of Bayesian multiparameter quantum estimation. We have achieved this by reviewing, streamlining, and contextualising existing results, and by deriving a number of new ones, including: the explicit set of conditions (19) that the optimal POVM must satisfy under the MSL criterion; the upper bound $2\mathcal{L}_{\mathrm{SPM}}$ on the minimum MSL in the first inequality of Eq. (35), showing that such a minimum can be at most twice the value of the bound when neglecting measurement incompatibility; the tighter upper bound through the PGM, and a tight bound on both $\mathcal{L}_{\mathrm{PGM}}$ and $\mathcal{L}_{\mathrm{prior}}$ using its POVM with the optimal PM estimator; the introduction of general PM operators and their corresponding matrix lower bounds via the monotonicity of inner products, presented in App. D; and finally, the demonstration that measurement incompatibility in a single-shot Bayesian setup is substantially affected by prior information, as illustrated through examples in qubit and optical configurations.

This work opens several avenues for future research that we now highlight, both theoretical and experimental. For example, in multiparameter quantum metrology—the simultaneous estimation of several parameters encoded by a quantum channel—additional precision trade-offs may arise from the choice of probe state and control strategy [95, 124, 126, 162], a phenomenon referred to as *probe incompatibility* [125]. While this aspect has not, to the best

[7] Bounds in local QET are most commonly denoted by the letter C, which stands for CRB—for example, $C_{\mathrm{H}}(\boldsymbol{\theta})$ for the HCRB. However, for consistency with the rest of this work we retain the terminology "loss" and the symbol $\mathcal{L}$, since these quantities bound the frequentist loss $\mathcal{L}(\boldsymbol{\theta})$ defined in Eq. (8).

of our knowledge, been examined in detail and in full generality within a Bayesian framework, numerical approaches for designing optimal multiparameter Bayesian estimation protocols [106, 107, 163] inherently account for probe incompatibility effects.

Still outstanding are conditions that lead to collapse in the hierarchy of lower or upper bounds. Perhaps the most likely candidate would be the Bayesian weak commutativity condition proposed in Eq. (75), whose local QET analogue implies equivalence of the Helstrom (SLD) and Holevo CRBs. Entirely consistent with the Bayesian weak-commutativity condition implying saturability of the NH bound, all our examples satisfy Eq. (75) and the NH bound appears attainable. Incidentally, the Bayesian Holevo bound also matches the SPM bound numerically in all examples[8], and—though not saturable and generally less significant in the Bayesian framework—this may be expected to hold more generally when the Bayesian weak-commutativity condition is satisfied, given the similarity of the underlying algebra.

Another important problem is addressing incompatibility—measurement and other forms thereof—for parameters obeying symmetries other than translation invariance. The symmetry remapping in Eq. (80) provides a starting point to generalise the results of this study to any location-isomorphic estimation problem, concrete illustrations of which have been discussed both in Sec. 4.2 and in Ref. [37]. A rigorous formulation should include the possibility of correlated priors and loss functions coupling different parameters, as well as the application of information geometry [119]. Moreover, a further extension to symmetries beyond location-isomorphic metrology is necessary to encompass phase parameters [120, 164].

In addition, while our analysis has focused on finite-dimensional systems, continuous-variable (CV) platforms—particularly quantum optical ones [2, 6, 165, 166]—play a central role in sensing applications. It will therefore be important to rigorously extend the theoretical framework developed here to CV systems. Given the rapid progress in learning theory for CV states [167–170], this direction appears especially promising. For example, the case of Gaussian states and priors [103] may allow for a tractable analysis while remaining relevant for practical implementations. In single-parameter Bayesian estimation, several specific results for Gaussian states are already known [171], and a general framework for multiparameter local QET with Gaussian states—capable of giving incompatibility with collective measurements—has recently been proposed [172].

[8]Supplemental data [149] shows that the Bayesian Holevo bound matches the SPM bound with $|\mathcal{I}_{\mathrm{H}}| \leq 2\times10^{-7}$ for $2 \leq d \leq 17$ for the example considered in Sec. 4.1, with $|\mathcal{I}_{\mathrm{H}}| \leq 2\times10^{-5}$ for $1 \leq n \leq 6$ (the Bayesian Holevo bound calculation becomes highly computationally demanding for larger n) for the example considered in Sec. 4.2, and with $|\mathcal{I}_{\mathrm{H}}| \leq 3\times10^{-5}$ for all points shown in Fig. 3.

At a practical level, identifying optimal strategies for even an individual shot in an adaptive, experimental protocol can be computationally difficult. Even methods reducible to convex problems [107] are currently far less performant than solving the corresponding NH bound, in turn more demanding than the SPM and PGM lower and upper bounds. The upper bound $\mathcal{L}_{\min} \leq 2\mathcal{L}_{\mathrm{SPM}}$ suggests that optimising the SPM affords a good compromise between optimality and tractability, compared to the exact but difficult minimum loss (20), with a smaller SPM bound also meaning a smaller upper bound on $\mathcal{L}_{\min}$ and so the worst-case impact of measurement-incompatibility is smaller in absolute terms. In practice this sees the pseudo-gain $L^{ij}\,\mathrm{Tr}[\rho_0 S_i S_j]$—i.e., the second term in the SPM lower bound (21); see App. A—become the optimiser of experimental control parameters for individual shots, fulfilling the role of a global alternative to the weighted trace of the inverse quantum FIM. Although such an approach will *not* generally lead to full optimality, we anticipate a substantial boost in quantum technological applications, in light of recent results from cold-atom experiments showing two-fold [36] and five-fold [37] precision enhancements obtained by employing the "classical" version of this quantifier, that is, the pseudo-gain computed using the PM estimator and a fixed POVM.

Altogether, we expect the Bayesian multiparameter framework for QET introduced here, together with the insight into the impact of measurement incompatibility, to become a fundamental building block for benchmarking both theory and experiment in quantum metrology, guiding the design of practically relevant, information-optimal quantum sensors.

Note added. During the preparation of the final version of this manuscript for publication, we became aware of the independent preprint by Zhang, Yamagata, and Suzuki [173], which develops a systematic theory of Bayesian monotone metrics and contains, among other results, a derivation of the monotone-metric matrix and scalar lower bounds presented in App. D. Revisiting our derivation, we found that an earlier version of App. D should have used complex, rather than only real, linear combinations of the parameter moments in order to justify the complex matrix inequality. We have amended the proof accordingly; the stated bounds are unchanged. Previous versions also specified an unnecessary property that linear maps E_ρ must satisfy; we thank Jun Suzuki for highlighting this. This property was not relied on in the subsequent proof.

Acknowledgments

F. A. acknowledges fruitful discussions with Stefano Chessa, Rafał Demkowicz-Dobrzański, Vittorio Giovannetti and Ludovico Lami. J. R. gratefully

thanks Luis Correa and Edward Gandar for insightful discussions. F. A. acknowledges partial financial support from Marie Skłodowska-Curie Action EUHORIZON-MSCA-2021PF-01 (project QECANM, Grant No. 101068347). D. B. acknowledges financial support from the European Union's Horizon Europe research and innovation program under the EPIQUE Project (Grant Agreement No. 101135288). J. R. acknowledges financial support from the Surrey Future Fellowship (University of Surrey), during which this project was developed and first released, and partial support from the UKRI-JST project "Quantum Control & Sensing: Enhancing high-precision quantum sensing in noisy environments via optimised control", during which this work was brought to publication.

Code and data generated for this paper are openly available [149]. An open-source python package for evaluating the bounds discussed is also available on GitHub [113].

Author contributions

F. A. and J. R. conceived the project, developed the theoretical framework and performed analytical calculations. D. B. carried out numerical simulations and also contributed to the theoretical framework and analytical calculations. All authors discussed the results and wrote the manuscript. The derivations in Appendix C and the example in Sec. 4.3 were obtained in a conversational style with a large language model (ChatGPT 5). This tool was also employed to assist with text editing and code optimisation. All scientific content and final text were reviewed and approved by the authors, who take full responsibility for the work.

References

[1] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nat. Photonics **5**, 222 (2011), arXiv:1102.2318.

[2] E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino, Photonic quantum metrology, AVS Quantum Sci. **2**, 024703 (2020), arXiv:2003.05821.

[3] J. Liu, M. Zhang, H. Chen, L. Wang, and H. Yuan, Optimal Scheme for Quantum Metrology, Adv. Quantum Technol. **5**, 2100080 (2022), arXiv:2111.12279.

[4] L. Jiao, W. Wu, S.-Y. Bai, and J.-H. An, Quantum Metrology in the Noisy Intermediate-Scale Quantum Era, Adv. Quantum Techol. **8**, 2300218 (2025), arXiv:2307.07701.

[5] V. Montenegro, C. Mukhopadhyay, R. Yousefjani, S. Sarkar, U. Mishra, M. G. A. Paris, and A. Bayat, Review: Quantum metrology and sensing with many-body systems, Phys. Rep. **1134**, 1 (2025), arXiv:2408.15323.

[6] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, Quantum Limits in Optical Interferometry, in *Progress in Optics, Volume 60*, edited by E. Wolf (Elsevier, Amsterdam, 2015) pp. 345–435, arXiv:1405.7703.

[7] J. Sahota and N. Quesada, Quantum correlations in optical metrology: Heisenberg-limited phase estimation without mode entanglement, Phys. Rev. A **91**, 013808 (2015), arXiv:1404.7110.

[8] N. Friis, D. Orsucci, M. Skotiniotis, P. Sekatski, V. Dunjko, H. J. Briegel, and W. Dür, Flexible resources for quantum metrology, New J. Phys. **19**, 063044 (2017), arXiv:1610.09999.

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

[10] R. Demkowicz-Dobrzański, W. Górecki, and M. Guţă, Multi-parameter estimation beyond quantum Fisher information, J. Phys. A **53**, 363001 (2020), arXiv:2001.11742.

[11] W. Górecki and R. Demkowicz-Dobrzański, Multiparameter quantum metrology in the Heisenberg limit regime: Many-repetition scenario versus full optimization, Phys. Rev. A **106**, 012424 (2022), arXiv:2203.09541.

[12] M. Tsang, Ziv-Zakai Error Bounds for Quantum Parameter Estimation, Phys. Rev. Lett. **108**, 230401 (2012), arXiv:1111.3568.

[13] J. Rubio and J. Dunningham, Quantum metrology in the presence of limited data, New J. Phys. **21**, 043037 (2019), arXiv:1810.12857.

[14] M. Valeri, E. Polino, D. Poderini, I. Gianani, G. Corrielli, A. Crespi, T. Osellame, N. Spagnolo, and F. Sciarrino, Experimental adaptive Bayesian estimation of multiple phases with limited data, npj Quantum Inf. **6**, 92 (2020), arXiv:2002.01232.

[15] M. Gessner and A. Smerzi, Hierarchies of Frequentist Bounds for Quantum Metrology: From Cramér-Rao to Barankin, Phys. Rev. Lett. **130**, 260801 (2023), arXiv:2303.06108.

[16] M. Hassani, C. Macchiavello, and L. Maccone, Digital Quantum Estimation, Phys. Rev. Lett. **119**, 200502 (2017), arXiv:1705.06666.

[17] V. Gebhart, M. Gessner, and A. Smerzi, Fundamental bounds for parameter estimation with few measurements, Phys. Rev. Res. **6**, 043261 (2024), arXiv:2402.14495.

[18] J. R. Hervas, A. Z. Goldberg, A. S. Sanz, Z. Hradil, J. Řeháček, and L. L. Sánchez-Soto, Beyond the Quantum Cramér-Rao Bound, Phys. Rev. Lett. **134**, 010804 (2025).

[19] J. R. Hervas, A. Z. Goldberg, A. S. Sanz, Z. Hradil, J. Řeháček, and L. L. Sánchez-Soto, Higher-order corrections to the quantum

Cramér-Rao bound, Phys. Rev. A **112**, 022426 (2025).

[20] J. J. Meyer, S. Khatri, D. Stilck França, J. Eisert, and P. Faist, Quantum Metrology in the Finite-Sample Regime, PRX Quantum **6**, 030336 (2025), arXiv:2307.06370.

[21] C. Mukhopadhyay, V. Montenegro, and A. Bayat, Current trends in global quantum metrology, J. Phys. A **58**, 063001 (2025), arXiv:2411.03850.

[22] A. M. Polloreno, J. L. Beckey, J. Levin, A. Shlosberg, J. K. Thompson, M. Foss-Feig, D. Hayes, and G. Smith, Opportunities and Limitations in Broadband Sensing, Phys. Rev. Appl. **19**, 014029 (2023), arXiv:2203.05520.

[23] R. R. Allen, F. Machado, I. L. Chuang, H.-Y. Huang, and S. Choi, Quantum Computing Enhanced Sensing (2025), arXiv:2501.07625.

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

[25] S. Personick, Application of quantum estimation theory to analog communication over quantum channels, IEEE Trans. Inf. Theory **17**, 240 (1971).

[26] A. S. Holevo, Statistical decision theory for quantum systems, J. Multivar. Anal. **3**, 337 (1973).

[27] A. S. Holevo, Statistical problems in quantum physics, in *Proc. Second Jpn.-USSR Symp. Probab. Theory*, edited by G. Maruyama and Yu. V. Prokhorov (Springer, Berlin, Heidelberg, 1973) pp. 104–119.

[28] H. Yuen and M. Lax, Multiple-parameter quantum estimation and measurement of nonselfadjoint observables, IEEE Trans. Inf. Theory **19**, 740 (1973).

[29] C. Helstrom and R. Kennedy, Noncommuting observables in quantum detection and estimation theory, IEEE Trans. Inf. Theory **20**, 16 (1974).

[30] S. E. D'Aurelio, M. Valeri, E. Polino, V. Cimini, I. Gianani, M. Barbieri, G. Corrielli, A. Crespi, R. Osellame, F. Sciarrino, and N. Spagnolo, Experimental investigation of Bayesian bounds in multiparameter estimation, Quantum Sci. Technol. **7**, 025011 (2022), arXiv:2304.06405.

[31] M. Valeri, V. Cimini, S. Piacentini, F. Ceccarelli, E. Polino, F. Hoch, G. Bizzarri, G. Corrielli, N. Spagnolo, R. Osellame, and F. Sciarrino, Experimental multiparameter quantum metrology in adaptive regime, Phys. Rev. Res. **5**, 013138 (2023), arXiv:2208.14473.

[32] V. Cimini, E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino, Benchmarking Bayesian quantum estimation, Quantum Sci. Technol. **9**, 035035 (2024), arXiv:2401.14900.

[33] V. Cimini, M. Valeri, S. Piacentini, F. Ceccarelli, G. Corrielli, R. Osellame, N. Spagnolo, and F. Sciarrino, Variational quantum algorithm for experimental photonic multiparameter estimation, npj Quantum Inf. **10**, 26 (2024), arXiv:2308.02643.

[34] F. Belliardo, V. Cimini, E. Polino, F. Hoch, B. Piccirillo, N. Spagnolo, V. Giovannetti, and F. Sciarrino, Optimizing quantum-enhanced Bayesian multiparameter estimation of phase and noise in practical sensors, Phys. Rev. Res. **6**, 023201 (2024), arXiv:2211.04747.

[35] G. Minati, E. Urbani, N. Spagnolo, V. Cimini, and F. Sciarrino, Multiparameter quantum-enhanced adaptive metrology with squeezed light, Nat. Commun. **17**, 6418 (2026), arXiv:2510.14739.

[36] J. Glatthard, J. Rubio, R. Sawant, T. Hewitt, G. Barontini, and L. A. Correa, Optimal cold atom thermometry using adaptive bayesian strategies, PRX Quantum **3**, 040330 (2022), arXiv:2204.11816.

[37] M. Overton, J. Rubio, N. Cooper, D. Baldolini, D. Johnson, J. Anders, and L. Hackermüller, Adaptive, symmetry-informed bayesian metrology for precise quantum technology measurements, Phys. Rev. Lett. **136**, 140801 (2026), arXiv:2410.10615.

[38] H. Espinós, O. Dhungel, A. Wickenbrock, D. Budker, R. Puebla, and E. Torrontegui, Microwave-free vector magnetometry and crystal orientation determination with Nitrogen-Vacancy centers using Bayesian inference (2025), arXiv:2512.13835.

[39] A. Rivas and A. Luis, Sub-Heisenberg estimation of non-random phase shifts, New J. Phys. **14**, 093052 (2012), arXiv:1105.6310.

[40] L. Pezzè, Sub-Heisenberg phase uncertainties, Phys. Rev. A **88**, 060101 (2013).

[41] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, Quantum Enhanced Multiple Phase Estimation, Phys. Rev. Lett. **111**, 070403 (2013), arXiv:1307.7653.

[42] P. A. Knott, T. J. Proctor, A. J. Hayes, J. F. Ralph, P. Kok, and J. A. Dunningham, Local versus global strategies in multiparameter estimation, Phys. Rev. A **94**, 062312 (2016), arXiv:1601.05912.

[43] J. Rubio and J. Dunningham, Bayesian multiparameter quantum metrology with limited data, Phys. Rev. A **101**, 032114 (2020), arXiv:1906.04123.

[44] 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), arXiv:2107.10863.

[45] W. Górecki, R. Demkowicz-Dobrzański, H. M.

Wiseman, and D. W. Berry, π-Corrected Heisenberg Limit, Phys. Rev. Lett. **124**, 030501 (2020), arXiv:1907.05428.

[46] R. Ravell Rodríguez and S. Morelli, Knowledge-dependent optimal Gaussian strategies for phase estimation, SciPost Phys. **18**, 167 (2025), arXiv:2412.16023.

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

[48] V. Gebhart, R. Santagati, A. A. Gentile, E. M. Gauger, D. Craig, N. Ares, L. Banchi, F. Marquardt, L. Pezzè, and C. Bonato, Learning quantum systems, Nat. Rev. Phys. **5**, 141 (2023), arXiv:2207.00298.

[49] Y. Berquin, Bayesian statistical learning using density operators (2023), arXiv:2212.14715.

[50] L. Banchi, J. L. Pereira, S. T. Jose, and O. Simeone, Statistical Complexity of Quantum Learning, Adv. Quantum Technol. **8**, 2300311 (2024), arXiv:2309.11617.

[51] S. Chen, C. Oh, S. Zhou, H.-Y. Huang, and L. Jiang, Tight Bounds on Pauli Channel Learning without Entanglement, Phys. Rev. Lett. **132**, 180805 (2024), arXiv:2309.13461.

[52] C. Oh, S. Chen, Y. Wong, S. Zhou, H.-Y. Huang, J. A. H. Nielsen, Z.-H. Liu, J. S. Neergaard-Nielsen, U. L. Andersen, L. Jiang, and J. Preskill, Entanglement-Enabled Advantage for Learning a Bosonic Random Displacement Channel, Phys. Rev. Lett. **133**, 230604 (2024), arXiv:2402.18809.

[53] A. Dutt, *Accelerating Learning of Quantum Systems Using Prior Information*, Ph.D thesis, Massachusetts Institute of Technology (2024).

[54] Y. Wang, C. Oh, J. Liu, L. Jiang, and S. Zhou, Advancing quantum imaging through learning theory, Nat Commun **17**, 1124 (2025), arXiv:2501.15685.

[55] H.-Y. Hu, M. Ma, W. Gong, Q. Ye, Y. Tong, S. T. Flammia, and S. F. Yelin, Ansatz-Free Hamiltonian Learning with Heisenberg-Limited Scaling, PRX Quantum **6**, 040315 (2025), arXiv:2502.11900.

[56] F. Belliardo, E. M. Gauger, M. H. Abobeih, T. H. Taminiau, Y. Altmann, and C. Bonato, Multidimensional Quantum Estimation and Model Learning Framework Based on Variational Bayesian Inference, PRX Quantum **7**, 020360 (2026), arXiv:2507.23130.

[57] T. Heinosaari, T. Miyadera, and M. Ziman, An invitation to quantum incompatibility, J. Phys. A **49**, 123001 (2016), arXiv:1511.07548.

[58] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, Colloquium: Incompatible measurements in quantum information science, Rev. Mod. Phys. **95**, 011003 (2023), arXiv:2112.06784.

[59] M. Szczykulska, T. Baumgratz, and A. Datta, Multi-parameter quantum metrology, Adv. Phys. X **1**, 621 (2016), arXiv:1604.02615.

[60] G. Li and X.-M. Lu, Ultimate tradeoff relation of quantum precision limits in multiparameter linear measurement, Quantum Sci. Technol. **11**, 035058 (2026), arXiv:2412.15031.

[61] B.-S. Hu and X.-M. Lu, Quantum Parameter Estimation Uncertainty Relation (2025), arXiv:2506.15352.

[62] I. Kull, P. A. Guérin, and F. Verstraete, Uncertainty and trade-offs in quantum multiparameter estimation, J. Phys. Math. Theor. **53**, 244001 (2020), arXiv:2002.05961.

[63] S. Razavian, M. G. A. Paris, and M. G. Genoni, On the Quantumness of Multiparameter Estimation Problems for Qubit Systems, Entropy **22**, 1197 (2020), arXiv:2010.12630.

[64] F. Belliardo and V. Giovannetti, Incompatibility in quantum parameter estimation, New J. Phys. **23**, 063055 (2021), arXiv:2102.13417.

[65] A. Candeloro, M. G. A. Paris, and M. G. Genoni, On the properties of the asymptotic incompatibility measure in multiparameter quantum estimation, J. Phys. A **54**, 485301 (2021), arXiv:2107.13426.

[66] A. Candeloro, Z. Pazhotan, and M. G. A. Paris, Dimension matters: Precision and incompatibility in multi-parameter quantum estimation models, Quantum Sci. Technol. **9**, 045045 (2024), arXiv:2403.07106.

[67] F. Albarelli, M. Barbieri, M. G. Genoni, and I. Gianani, A perspective on multiparameter quantum metrology: From theoretical tools to applications in quantum imaging, Phys. Lett. A **384**, 126311 (2020), arXiv:1911.12067.

[68] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, Quantum Fisher information matrix and multiparameter estimation, J. Phys. A **53**, 023001 (2020), arXiv:1907.08037.

[69] L. Pezzè and A. Smerzi, Advances in multiparameter quantum sensing and metrology (2025), arXiv:2502.17396.

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

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

[72] O. E. Barndorff-Nielsen and R. D. Gill, Fisher information in quantum statistics, J. Phys. A **33**, 4481 (2000), arXiv:quant-ph/9808009.

[73] M. Hayashi and K. Matsumoto, Asymptotic performance of optimal state estimation in qubit system, J. Math. Phys. **49**, 102101 (2008), arXiv:quant-ph/0411073.

[74] J. Kahn and M. Guţă, Local Asymptotic Normality for Finite Dimensional Quantum Sys-

tems, Commun. Math. Phys. **289**, 597 (2009), arXiv:0804.3876.

[75] K. Yamagata, A. Fujiwara, and R. D. Gill, Quantum local asymptotic normality based on a new quantum likelihood ratio, Ann. Stat. **41**, 2197 (2013), arXiv:1210.3749.

[76] Y. Yang, G. Chiribella, and M. Hayashi, Attaining the Ultimate Precision Limit in Quantum State Estimation, Commun. Math. Phys. **368**, 223 (2019), arXiv:1802.07587.

[77] A. Fujiwara and K. Yamagata, Efficiency of estimators for locally asymptotically normal quantum statistical models, Ann. Stat. **51**, 1159 (2023), arXiv:2209.00832.

[78] M. Tsang, F. Albarelli, and A. Datta, Quantum Semiparametric Estimation, Phys. Rev. X **10**, 031023 (2020), arXiv:1906.09871.

[79] A. Carollo, B. Spagnolo, A. A. Dubkov, and D. Valenti, On quantumness in multi-parameter quantum estimation, J. Stat. Mech. Theory Exp. **2019**, 094010 (2019), arXiv:1911.11621.

[80] K. Matsumoto, A new approach to the Cramér-Rao-type bound of the pure-state model, J. Phys. A **35**, 3111 (2002), arXiv:quant-ph/9711008.

[81] R. D. Gill and S. Massar, State estimation for large ensembles, Phys. Rev. A **61**, 042312 (2000), arXiv:quant-ph/9902063.

[82] X.-M. Lu and X. Wang, Incorporating Heisenberg's Uncertainty Principle into Quantum Multiparameter Estimation, Phys. Rev. Lett. **126**, 120503 (2021), arXiv:2008.08888.

[83] H. Chen, Y. Chen, and H. Yuan, Incompatibility measures in multiparameter quantum estimation under hierarchical quantum measurements, Phys. Rev. A **105**, 062442 (2022), arXiv:2109.05807.

[84] H. Chen, Y. Chen, and H. Yuan, Information Geometry under Hierarchical Quantum Measurement, Phys. Rev. Lett. **128**, 250502 (2022), arXiv:2206.13095.

[85] H. Chen, L. Wang, and H. Yuan, Simultaneous measurement of multiple incompatible observables and tradeoff in multiparameter quantum estimation, npj Quantum Inf. **10**, 98 (2024), arXiv:2310.11925.

[86] L. Wang, H. Chen, and H. Yuan, Tight tradeoff relation and optimal measurement for multiparameter quantum estimation, Phys. Rev. A **114**, 012609 (2026), arXiv:2504.09490.

[87] H. Nagaoka, A New Approach to Cramér-Rao Bounds for Quantum State Estimation, IEICE Tech. Rep. **IT 89-42**, 9 (1989).

[88] H. Nagaoka, A Generalization of the Simultaneous Diagonalization of Hermitian Matrices and its Relation to Quantum Estimation Theory, Trans. Jpn. Soc. Ind. Appl. Math. **1**, 43 (1991).

[89] M. Hayashi, On simultaneous measurement of noncommutative observables, Dev. Infin.-Dimens. Non-Commut. Anaysis , 96 (1999).

[90] L. O. Conlon, J. Suzuki, P. K. Lam, and S. M. Assad, Efficient computation of the Nagaoka–Hayashi bound for multiparameter estimation with separable measurements, npj Quantum Inf. **7**, 110 (2021), arXiv:2008.02612.

[91] L. O. Conlon, J. Suzuki, P. K. Lam, and S. M. Assad, The gap persistence theorem for quantum multiparameter estimation (2024), arXiv:2208.07386.

[92] S. K. Yung, L. O. Conlon, J. Zhao, P. K. Lam, and S. M. Assad, Comparison of estimation limits for quantum two-parameter estimation, Phys. Rev. Res. **6**, 033315 (2024), arXiv:2407.12466.

[93] A. Das, L. O. Conlon, J. Suzuki, S. K. Yung, P. K. Lam, and S. M. Assad, Holevo Cramér-Rao bound: How close can we get without entangling measurements?, Quantum **9**, 1867 (2025), arXiv:2405.09622.

[94] L. O. Conlon, J. Suzuki, P. K. Lam, and S. M. Assad, Role of the extended Hilbert space in the attainability of the quantum Cramér–Rao bound for multiparameter estimation, Phys. Lett. A **542**, 130445 (2025), arXiv:2404.01520.

[95] M. Hayashi and Y. Ouyang, Tight Cramér-Rao type bounds for multiparameter quantum metrology through conic programming, Quantum **7**, 1094 (2023), arXiv:2209.05218.

[96] M. Hayashi and Y. Ouyang, Finding the optimal probe state for multiparameter quantum metrology using conic programming, npj Quantum Inf. **10**, 111 (2024), arXiv:2401.05886.

[97] H. Yuen, R. Kennedy, and M. Lax, On optimal quantum receivers for digital signal detection, Proc. IEEE **58**, 1770 (1970).

[98] H. Yuen, RS. Kennedy, and M. Lax, Optimum testing of multiple hypotheses in quantum detection theory, IEEE Trans. Inf. Theory **21**, 125 (1975).

[99] B. Coutts, M. Girard, and J. Watrous, Certifying optimality for convex quantum channel optimization problems, Quantum **5**, 448 (2021), arXiv:1810.13295.

[100] D. C. Brody and B. Meister, Bayesian inference in quantum systems, Physica A **223**, 348 (1996).

[101] J. Rubio, First-principles construction of symmetry-informed quantum metrologies, Phys. Rev. A **110**, L030401 (2024), arXiv:2402.16410.

[102] V. P. Belavkin and B. A. Grishanin, Optimum Estimation in Quantum Channels by the Generalized Heisenberg Inequality Method, Probl. Peredachi Inf. **9**, 44 (1973).

[103] X.-B. Wang, T. Hiroshima, A. Tomita, and M. Hayashi, Quantum information with

Gaussian states, Phys. Rep. **448**, 1 (2007), arXiv:0801.4604.

[104] J. Suzuki, Bayesian Nagaoka-Hayashi Bound for Multiparameter Quantum-State Estimation Problem, IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **E107.A**, 510 (2024), arXiv:2302.14223.

[105] J. Zhang and J. Suzuki, Bayesian Logarithmic Derivative Type Lower Bounds for Quantum Estimation (2024), arXiv:2405.10525.

[106] R. Kaubruegger, A. Shankar, D. V. Vasilyev, and P. Zoller, Optimal and Variational Multiparameter Quantum Metrology and Vector-Field Sensing, PRX Quantum **4**, 020333 (2023), arXiv:2302.07785.

[107] J. Bavaresco, P. Lipka-Bartosik, P. Sekatski, and M. Mehboudi, Designing optimal protocols in bayesian quantum parameter estimation with higher-order operations, Phys. Rev. Res. **6**, 023305 (2024), arXiv:2311.01513.

[108] V. P. Belavkin, Optimum distinction of non-orthogonal quantum signals, Radiotekhnika Elektron. **20**, 1177 (1975).

[109] V. P. Belavkin, Optimal multiple quantum statistical hypothesis testing, Stochastics **1**, 315 (1975).

[110] A. S. Holevo, On Asymptotically Optimal Hypothesis Testing in Quantum Statistics, Theory Probab. Its Appl. **23**, 411 (1979).

[111] P. Hausladen and W. K. Wootters, A 'Pretty Good' Measurement for Distinguishing Quantum States, J. Mod. Opt. **41**, 2385 (1994).

[112] M. Hayashi, *Quantum Information Theory* (Springer, Berlin, Heidelberg, 2017).

[113] `multipar_bayes` available on github: `dbranford/multipar_bayes` (2026).

[114] G. Chiribella, G. M. D'Ariano, and D. Schlingemann, How continuous quantum measurements in finite dimensions are actually discrete, Phys. Rev. Lett. **98**, 190403 (2007), arXiv:quant-ph/0702068.

[115] W. von derLinden, V. Dose, and U. von Toussaint, *Bayesian Probability Theory: Applications in the Physical Sciences* (Cambridge University Press, 2014).

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

[117] W. Górecki, *Heisenberg Limit beyond Quantum Fisher Information*, Ph.D. thesis, Faculty of Physics, Univ. of Warsaw (2023), arXiv:2304.14370.

[118] E. T. Jaynes and O. Kempthorne, Confidence intervals vs bayesian intervals, in *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science: Proceedings of an International Research Colloquium held at the University of Western Ontario, London, Canada, 10–13 May 1973 Volume II Foundations and Philosophy of Statistical Inference*, edited by W. L. Harper and C. A. Hooker (Springer Netherlands, Dordrecht, 1976) pp. 175–257.

[119] J. Boeyens, S. Seah, and S. Nimmrichter, Uninformed Bayesian quantum thermometry, Phys. Rev. A **104**, 052214 (2021), arXiv:2108.07025.

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

[121] J. Rubio, J. Anders, and L. A. Correa, Global Quantum Thermometry, Phys. Rev. Lett. **127**, 190402 (2021), arXiv:2011.13018.

[122] J. Boeyens, J. Glatthard, E. Gandar, S. Nimmrichter, L. A. Correa, and J. Rubio, On the role of symmetry and geometry in global quantum sensing, Quantum Sci. Technol. **10**, 045053 (2025), arXiv:2502.14817.

[123] M. Tsang, Generalized conditional expectations for quantum retrodiction and smoothing, Phys. Rev. A **105**, 042213 (2022), arXiv:1912.02711.

[124] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, Compatibility in multiparameter quantum metrology, Phys. Rev. A **94**, 052108 (2016), arXiv:1608.02634.

[125] F. Albarelli and R. Demkowicz-Dobrzański, Probe Incompatibility in Multiparameter Noisy Quantum Metrology, Phys. Rev. X **12**, 011039 (2022), arXiv:2104.11264.

[126] Z. Hu, S. Wang, L. Qiao, T. Isogawa, C. Li, Y. Yang, G. Wang, H. Yuan, and P. Cappellaro, Control incompatibility in multiparameter quantum metrology (2024), arXiv:2411.18896.

[127] A. S. Holevo, *Investigations in the General Theory of Statistical Decisions*, Proc. Steklov Institute of Mathematics No. 124 (American Mathematical Soc., 1978 [1976]).

[128] J. S. Sidhu and P. Kok, Geometric perspective on quantum parameter estimation, AVS Quantum Sci. **2**, 014701 (2020), arXiv:1907.06628.

[129] J. Yang, S. Pang, Y. Zhou, and A. N. Jordan, Optimal measurements for quantum multiparameter estimation with general states, Phys. Rev. A **100**, 032104 (2019), arXiv:1806.07337.

[130] P. Horodecki, Ł. Rudnicki, and K. Życzkowski, Five Open Problems in Quantum Information Theory, PRX Quantum **3**, 010101 (2022), arXiv:2002.03233.

[131] H. I. Nurdin, Corrections to "Saturability of the Quantum Cramér-Rao Bound in Multiparameter Quantum Estimation at the Single-Copy Level", IEEE Control Syst. Lett. **8**, 2111 (2024).

[132] L. Pezzè, M. A. Ciampini, N. Spagnolo, P. C. Humphreys, A. Datta, I. A. Walmsley, M. Barbieri, F. Sciarrino, and A. Smerzi, Optimal Measurements for Simultaneous Quantum Es-

timation of Multiple Phases, Phys. Rev. Lett. **119**, 130504 (2017), arXiv:1705.03687.

[133] J. Rubio, Quantum scale estimation, Quantum Sci. Technol. **8**, 015009 (2022), arXiv:2111.11921.

[134] H. Barnum and E. Knill, Reversing quantum dynamics with near-optimal quantum and classical fidelity, J. Math. Phys. **43**, 2097 (2002), arXiv:quant-ph/0004088.

[135] H. K. Mishra, L. Lami, and M. M. Wilde, Near-optimal performance of square-root measurement for general score functions and quantum ensembles (2025), arXiv:2505.20447.

[136] D. Petz, Monotone metrics on matrix spaces, Linear Algebra Its Appl. **244**, 81 (1996).

[137] M. Scandi, P. Abiuso, J. Surace, and D. De Santis, Quantum Fisher information and its dynamical nature, Rep. Prog. Phys. **88**, 076001 (2025), arXiv:2304.14984.

[138] K. Hunter, Measurement does not always aid state discrimination, Phys. Rev. A **68**, 012306 (2003), arXiv:quant-ph/0211148.

[139] S. M. Barnett and S. Croke, Quantum state discrimination, Adv. Opt. Photonics **1**, 238 (2009), arXiv:0810.1970.

[140] M. Quadeer, Haar-random and pretty good measurements for Bayesian state estimation (2024), arXiv:2310.20565.

[141] R. D. Gill, Conciliation of Bayes and Pointwise Quantum State Estimation: Asymptotic information bounds in quantum statistics, in *Quantum Stoch. Inf. Stat. Filter. Control*, edited by V. P. Belavkin and M. Guţă (World Scientific, 2008) pp. 239–261, arXiv:math/0512443.

[142] M. Jarzyna and R. Demkowicz-Dobrzański, Quantum interferometry with and without an external phase reference, Phys. Rev. A **85**, 011801 (2012), arXiv:1108.3844.

[143] C. Macchiavello, Optimal estimation of multiple phases, Phys. Rev. A **67**, 062302 (2003), arXiv:quant-ph/0304126.

[144] V. Gebhart, A. Smerzi, and L. Pezzè, Bayesian quantum multiphase estimation algorithm, Phys. Rev. Appl. **16**, 014035 (2021), arXiv:2010.09075.

[145] M. J. W. Hall and H. M. Wiseman, Does Nonlinear Metrology Offer Improved Resolution? Answers from Quantum Information Theory, Phys. Rev. X **2**, 041006 (2012), arXiv:1205.2405.

[146] V. Giovannetti and L. Maccone, Sub-Heisenberg Estimation Strategies Are Ineffective, Phys. Rev. Lett. **108**, 210404 (2012), arXiv:1201.1878.

[147] J. Rubio, P. A. Knott, and J. A. Dunningham, Non-asymptotic analysis of quantum metrology protocols beyond the Cramér-Rao bound, J. Phys. Commun. **2**, 015027 (2018), arXiv:1707.05022.

[148] D. Branford and J. Rubio, Average number is an insufficient metric for interferometry, New J. Phys. **23**, 123041 (2021), arXiv:2107.06698.

[149] F. Albarelli, D. Branford, and J. Rubio, Code and data for Measurement incompatibility in Bayesian multiparameter quantum estimation , 10.5281/zenodo.17814129 (2026).

[150] M. D. Vidrighin, G. Donati, M. G. Genoni, X.-M. Jin, W. S. Kolthammer, M. S. Kim, A. Datta, M. Barbieri, and I. A. Walmsley, Joint estimation of phase and phase diffusion for quantum metrology, Nat. Commun. **5**, 3532 (2014), arXiv:1410.5353.

[151] E. Roccia, I. Gianani, L. Mancino, M. Sbroscia, F. Somma, M. G. Genoni, and M. Barbieri, Entangling measurements for multiparameter estimation with two qubits, Quantum Sci. Technol. **3**, 01LT01 (2018), arXiv:1704.03327.

[152] A. Kobus and R. Demkowicz-Dobrzański, Asymptotically optimal joint phase and dephasing strength estimation using spin-squeezed states (2025), arXiv:2507.22997.

[153] S. K. Yung, A. Das, J. Suzuki, P. K. Lam, J. Zhao, L. O. Conlon, and S. M. Assad, Saturating the Quantum Cramér–Rao Bound in Prioritised Parameter Estimation (2025), arXiv:2511.06704.

[154] M. G. Genoni, S. Olivares, and M. G. A. Paris, Optical Phase Estimation in the Presence of Phase Diffusion, Phys. Rev. Lett. **106**, 153603 (2011), arXiv:1012.1123.

[155] M. Szczykulska, T. Baumgratz, and A. Datta, Reaching for the quantum limits in the simultaneous estimation of phase and phase diffusion, Quantum Sci. Technol. **2**, 044004 (2017), arXiv:1701.07520.

[156] Z. Hou, J.-F. Tang, J. Shang, H. Zhu, J. Li, Y. Yuan, K.-D. Wu, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, Deterministic realization of collective measurements via photonic quantum walks, Nat. Commun. **9**, 1414 (2018), arXiv:1710.10045.

[157] M. Parniak, S. Borówka, K. Boroszko, W. Wasilewski, K. Banaszek, and R. Demkowicz-Dobrzański, Beating the Rayleigh Limit Using Two-Photon Interference, Phys. Rev. Lett. **121**, 250503 (2018), arXiv:1803.07096.

[158] Y. Yuan, Z. Hou, J.-F. Tang, A. Streltsov, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, Direct estimation of quantum coherence by collective measurements, npj Quantum Inf. **6**, 1 (2020), arXiv:2001.01384.

[159] L. O. Conlon, T. Vogl, C. D. Marciniak, I. Pogorelov, S. K. Yung, F. Eilenberger, D. W. Berry, F. S. Santana, R. Blatt, T. Monz, P. K. Lam, and S. M. Assad, Approaching optimal entangling collective measurements on quantum

computing platforms, Nat. Phys. **19**, 351 (2023), arXiv:2205.15358.

[160] M. Mi, B. Wang, and L. Zhang, Approaching the Multiparameter Quantum Cramér-Rao Bound via Classical Correlation and Entangling Measurements, Phys. Rev. Lett. **135**, 110804 (2025), arXiv:2509.10196.

[161] D. V. Vasilyev, A. Shankar, R. Kaubruegger, and P. Zoller, Optimal Multiparameter Metrology: The Quantum Compass Solution (2024), arXiv:2404.14194.

[162] W. Górecki, S. Zhou, L. Jiang, and R. Demkowicz-Dobrzański, Optimal probes and error-correction schemes in multi-parameter quantum metrology, Quantum **4**, 288 (2020), arXiv:1901.00896.

[163] E. L. André, J. Bavaresco, and M. Mehboudi, Strategy optimization for Bayesian quantum parameter estimation with finite copies: Adaptive greedy, parallel, sequential, and general strategies, Quantum Sci. Technol. **11**, 035040 (2026), arXiv:arXiv:2602.09655.

[164] R. Demkowicz-Dobrzański, Optimal phase estimation with arbitrary a priori knowledge, Phys. Rev. A **83**, 061802 (2011), arXiv:1102.0786.

[165] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Advances in photonic quantum sensing, Nat. Photonics **12**, 724 (2018), arXiv:1811.01969.

[166] M. Barbieri, Optical Quantum Metrology, PRX Quantum **3**, 010202 (2022), arXiv:2507.22680.

[167] Y.-D. Wu, Y. Zhu, G. Chiribella, and N. Liu, Efficient learning of continuous-variable quantum states, Phys. Rev. Res. **6**, 033280 (2024), arXiv:2303.05097.

[168] F. A. Mele, A. A. Mele, L. Bittel, J. Eisert, V. Giovannetti, L. Lami, L. Leone, and S. F. E. Oliviero, Learning quantum states of continuous-variable systems, Nature Physics **21**, 2002 (2025), arXiv:2405.01431.

[169] L. Bittel, F. A. Mele, A. A. Mele, S. Tirone, and L. Lami, Optimal estimates of trace distance between bosonic Gaussian states and applications to learning, Quantum **9**, 1769 (2025), arXiv:2411.02368.

[170] M. Fanizza, C. Rouzé, and D. S. França, Efficient Hamiltonian, structure and trace distance learning of Gaussian states (2025), arXiv:2411.03163.

[171] S. Morelli, A. Usui, E. Agudelo, and N. Friis, Bayesian parameter estimation using Gaussian states and measurements, Quantum Sci. Technol. **6**, 025018 (2021), arXiv:2009.03709.

[172] S. Chang, M. G. Genoni, and F. Albarelli, Efficiently evaluating Holevo, RLD and SLD Cramér-Rao bounds for multiparameter quantum estimation with Gaussian states, Commun. Phys. **9**, 126 (2026), arXiv:2504.17873.

[173] J. Zhang, K. Yamagata, and J. Suzuki, Bayesian Monotone Metrics for Multiparameter Quantum Estimation (2026), arXiv:2607.01685.

A Prior loss and precision gain in Bayesian quantum estimation

Suppose we make a random guess of the unknown parameters based solely on the prior probability. Let us denote the corresponding vector of estimates by $\tilde{\boldsymbol{\theta}}$, noting the absence of any outcome dependence. The MSL in this case is given by

$$\mathcal{L}_p(\tilde{\boldsymbol{\theta}}) \coloneqq \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, (\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta})^T L\, (\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}), \tag{105}$$

which we may refer to as the *prior loss*. Alternatively, one may obtain Eq. (105) by inserting a trivial POVM consisting of a single element, the identity operator, into the MSL in Eq. (14).

It is straightforward to verify that the estimator minimising Eq. (105) is the prior mean,

$$\boldsymbol{\mu} \coloneqq \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \boldsymbol{\theta}, \tag{106}$$

which leads to the minimum prior loss

$$\begin{aligned} \mathcal{L}_{\text{prior}} &\coloneqq \mathcal{L}_p(\boldsymbol{\mu}) \\ &= \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \boldsymbol{\theta}^\top L\, \boldsymbol{\theta} - \boldsymbol{\mu}^T L\, \boldsymbol{\mu} \\ &= \lambda - \boldsymbol{\mu}^T L\, \boldsymbol{\mu}, \end{aligned} \tag{107}$$

where λ was already defined in Eq. (15). Such an optimal prior loss is nothing but $\operatorname{tr} L\Sigma_p$, where Σ_p is the covariance matrix of the multivariate prior probability $p(\boldsymbol{\theta})$ and we have introduced a lowercase trace of $d\times d$ matrices, tr, distinct from the uppercase trace of operators acting on the physical Hilbert space of the system.

One may further show that $\mu_i = \mathrm{Tr}[\rho_0 S_i]$, where ρ_0 and S_i are defined in Sec. 2.3. This, together with Eq. (106), allows us to rewrite the SPM bound in Eq. (21) as

$$\mathcal{L}_{\mathrm{SPM}} = \mathcal{L}_{\mathrm{prior}} - \mathcal{G}_{\mathrm{SPM}}, \tag{108}$$

where

$$\mathcal{G}_{\mathrm{SPM}} \coloneqq L^{ij}(\mathrm{Tr}[\rho_0 S_j S_i] - \mathrm{Tr}[\rho_0 S_j]\ \mathrm{Tr}[\rho_0 S_i]). \tag{109}$$

Written in this form, the SPM bound can be interpreted as the result of reducing the prior loss by the amount of precision gained through individually optimal measurements, $\mathcal{G}_{\mathrm{SPM}}$.

Recent literature has referred to the quantity $L^{ij}\,\mathrm{Tr}[\rho_0 S_j S_i]$ as an alternative definition of precision gain when the SPM bound is used as an optimiser for control parameters [37, 122]. From a computational standpoint, this definition is often more convenient, and it has likewise been adopted in our code [113]. Accordingly, we refer to this quantity as the *pseudo-gain*, following the discussion above. Naturally, the gain and pseudo-gain coincide whenever the vector prior mean $\boldsymbol{\mu}$ vanishes.

B Optimality conditions for the MSL

Let $M_{\mathrm{opt}}(\tilde{\boldsymbol{\theta}})$ denote the optimal POVM for a Bayesian quantum multiparameter estimation problem. Following Holevo [26, 27, 127] and Helstrom [9, 29], a necessary and sufficient condition for such a POVM to exist is the existence of a Hermitian operator

$$\Upsilon = \int \mathrm{d}\tilde{\boldsymbol{\theta}}\, M_{\mathrm{opt}}(\tilde{\boldsymbol{\theta}})\, \varrho(\tilde{\boldsymbol{\theta}}) = \int \mathrm{d}\tilde{\boldsymbol{\theta}}\, \varrho(\tilde{\boldsymbol{\theta}})\, M_{\mathrm{opt}}(\tilde{\boldsymbol{\theta}}), \tag{110}$$

with

$$\varrho(\tilde{\boldsymbol{\theta}}) \coloneqq \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \rho(\boldsymbol{\theta})\, l(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}), \tag{111}$$

that satisfies

$$\varrho(\tilde{\boldsymbol{\theta}}) - \Upsilon \succeq 0 \tag{112}$$

for all $\tilde{\boldsymbol{\theta}}$. The minimum loss is then given by $\mathcal{L}_{\min} = \mathrm{Tr}(\Upsilon)$. A discrete version of this optimality condition exists for hypothesis testing [97, 98], and a generalisation to convex functions has recently been proposed [99].

We now derive the explicit form of these conditions for the MSL. Using the notation introduced in Sec. 2.3, Eq. (111) becomes

$$\varrho(\tilde{\boldsymbol{\theta}}) = \Lambda + L^{ij}\big(\tilde{\theta}_j \tilde{\theta}_i \rho_0 - 2\tilde{\theta}_j \bar{\rho}_i\big), \tag{113}$$

where $\mathrm{Tr}(\Lambda) = \lambda$. Accordingly, the operator in Eq. (110) takes the form

$$\Upsilon = \Lambda + L^{ij}\big(M_{\mathrm{opt},ji}\rho_0 - 2M_{\mathrm{opt},j}\bar{\rho}_i\big) = \Lambda + L^{ij}\big(\rho_0 M_{\mathrm{opt},ji} - 2\bar{\rho}_j M_{\mathrm{opt},i}\big). \tag{114}$$

By adding both expressions for Υ and dividing by two, we obtain

$$\Upsilon = \Lambda + \tfrac{1}{2} L^{ij}(M_{\mathrm{opt},ji}\rho_0 + \rho_0 M_{\mathrm{opt},ji}) - M_{\mathrm{opt},j}\bar{\rho}_i - \bar{\rho}_j M_{\mathrm{opt},i}. \tag{115}$$

Finally, inserting Eqs. (113) and (115) into Eq. (112), we obtain the conditions that any POVM optimal under the MSL criterion must satisfy:

$$L^{ij}[(\tilde{\theta}_j\tilde{\theta}_i\mathbb{I} - M_{\mathrm{opt},ji})\rho_0 + \rho_0(\tilde{\theta}_j\tilde{\theta}_i\mathbb{I} - M_{\mathrm{opt},ji}) - 2(\tilde{\theta}_j\mathbb{I} - M_{\mathrm{opt},j})\bar{\rho}_i - 2\bar{\rho}_j(\tilde{\theta}_i\mathbb{I} - M_{\mathrm{opt},i})] \succeq 0, \tag{116}$$

for all $\tilde{\boldsymbol{\theta}}$, as stated in the main text.

C Tightness of the NH bound for two parameters on a single qubit

In the main text we have shown that the Bayesian NH bound is obtained by relaxing the requirement that the first- and second-moment operators $\{M_i\}$ and $\{M_{ij}\}$ originate from an underlying POVM $M(x)$, exploiting the operator-valued moment-matrix inequality (42). In general, this relaxation can be loose. However, in this appendix we show that, for two parameters on a single qubit, the relaxation is instead exact, and it coincides with an attainability result proved by Nagaoka for 2×2 matrices. Thus, this appendix shows that the reason for the tightness of Nagaoka's two-parameter bound in local estimation naturally carries over to the Bayesian domain.

For $d = 2$ and the identity weight matrix $L = \mathbb{I}_2$ (without loss of generality; see footnote 4 in the main text), the optimisation over the second-moment operators M_{11}, M_{12}, M_{22} can be carried out analytically, leading to Eq. (45), which we reproduce here for convenience:

$$\mathcal{L}_{\mathrm{NH}} = \lambda + \min_{M_1,M_2} \left\{ \mathrm{Tr}\big[\rho_0(M_1^2 + M_2^2)\big] + \big\|\sqrt{\rho_0}[M_1, M_2]\sqrt{\rho_0}\big\|_1 - 2\,\mathrm{Tr}\big[\bar{\rho}_1 M_1 + \bar{\rho}_2 M_2\big] \right\}. \tag{117}$$

Here, M_1, M_2 are the operator-valued first moments, defined in Eq. (17) in the main text, and $\|\cdot\|_1$ is the trace norm. The only source of looseness is that, in the derivation of $\mathcal{L}_{\mathrm{NH}}$, the pair (M_1, M_2) is treated as a free pair of Hermitian operators, constrained only indirectly through Eq. (42), without explicitly enforcing that it comes from a common POVM $M(x)$.

Nagaoka proved the following result, for which we adapt the notation to the one used in this paper:

Theorem (Ref. [88, Theorem 2])**.** *Fix a pair of Hermitian operators M_1, M_2 acting on a two-dimensional Hilbert space, and consider the set of all POVMs $M(x_1, x_2)$ on $\mathbb{R}^2$ whose first moments match this pair:*

$$\int \mathrm{d}x_1 \mathrm{d}x_2\, M(x_1, x_2)\, x_i = M_i, \qquad i = 1, 2. \tag{118}$$

Among these POVMs, define the minimal achievable "quadratic" second-moment cost $\sum_i \mathrm{Tr}[\rho_0 M_{ii}]$ to be

$$C^\star(M_1, M_2) := \inf_{M(x_1,x_2)\ s.t.\ (118)} \mathrm{Tr}\left[\rho_0 \int \mathrm{d}x_1 \mathrm{d}x_2\, (x_1^2 + x_2^2)\, M(x_1, x_2)\right]. \tag{119}$$

Then, for a qubit system, this infimum is attainable and admits the closed form

$$C^\star(M_1, M_2) = \mathrm{Tr}\big[\rho_0(M_1^2 + M_2^2)\big] + \big\|\sqrt{\rho_0}[M_1, M_2]\sqrt{\rho_0}\big\|_1. \tag{120}$$

This means that, for any M_1, M_2, an MSL

$$\mathcal{L} = \lambda + \mathrm{Tr}\big[\rho_0(M_1^2 + M_2^2)\big] + \big\|\sqrt{\rho_0}[M_1, M_2]\sqrt{\rho_0}\big\|_1 - 2\,\mathrm{Tr}\big[\bar{\rho}_1 M_1 + \bar{\rho}_2 M_2\big] \tag{121}$$

is attainable, which entails attainability of the Bayesian NH lower bound Eq. (117). Nagaoka also provided explicitly an optimal POVM attaining the equality above. Notice that this result is proven under the assumption $\rho_0 \succ 0$, which, for qubit systems, holds in essentially all nontrivial cases. The only exception arises when the prior is a Dirac delta, in which case the estimation problem becomes trivial.

The Bayesian NH bound is obtained by optimising over $\{M_i\}$ and $\{M_{ij}\}$ subject to the moment-matrix inequality (42), instead of optimising over POVMs directly. For two parameters, Eq. (45) (restated as Eq. (117)) already encodes the best possible contribution of second moments compatible with a given pair (M_1, M_2), namely Eq. (120). Therefore, in the qubit case for every candidate pair (M_1, M_2), there exists a POVM whose second moments attain the value $\mathrm{Tr}[\rho_0(M_1^2 + M_2^2)] + \|\sqrt{\rho_0}[M_1, M_2]\sqrt{\rho_0}\|_1$; hence the minimisation in Eq. (117) is not merely a relaxation: it coincides with the true minimum MSL over single-qubit POVMs. Operationally, one may view Eq. (117) as first optimising over the "effective first-moment operators" M_1, M_2, and then invoking Nagaoka's construction to realise the corresponding optimal second moments by a concrete measurement on the qubit.

C.1 Explicit optimal POVM attaining the two parameter single qubit NH bound

For completeness, we briefly give the details of an explicit POVM attaining Eq. (120) for any fixed Hermitian pair (M_1, M_2) on a qubit with a full-rank ρ_0. We assume that M_1 and M_2 are linearly independent and non-commuting; otherwise, the problem is trivial, as they would share a common eigenbasis and the SPM bound would therefore be tight. Let $x := \mathrm{Tr}[\rho_0 M_1]$ and $y := \mathrm{Tr}[\rho_0 M_2]$ in order to define the centred operators $X := M_1 - x\mathbb{I}_2$ and $Y := M_2 - y\mathbb{I}_2$, so that $\mathrm{Tr}[\rho_0 X] = \mathrm{Tr}[\rho_0 Y] = 0$. Introduce an inner product induced by the state ρ_0 (see the next section for more details on such inner products)

$$\langle A, B\rangle_{\rho_0}^{\mathrm{S},(e)} := \frac{1}{2}\,\mathrm{Tr}\left[\rho_0\,(AB + BA)\right], \tag{122}$$

and let $\mathcal{W} := \mathrm{span}_{\mathbb{R}}\{X, Y\}$, i.e., the two-dimensional subspace of Hermitian operators with zero first moments spanned by X and Y. Define a linear map $\mathsf{A} : \mathcal{W} \to \mathcal{W}$ as

$$\mathsf{A}(W) := \langle W, X\rangle_{\rho_0}^{\mathrm{S},(e)}\, X \;+\; \langle W, Y\rangle_{\rho_0}^{\mathrm{S},(e)}\, Y. \tag{123}$$

This map is symmetric and positive definite with respect to $\langle\cdot,\cdot\rangle_{\rho_0}^{\mathrm{S},(e)}$, hence it has an orthonormal eigenbasis of Hermitian operators $\{U_1, U_2\} \subset \mathcal{W}$ with eigenvalues $\nu_1, \nu_2 > 0$ satisfying $\mathsf{A}(U_j) = \nu_j U_j$ and $\langle U_j, U_k\rangle_{\rho_0}^{\mathrm{S},(e)} = \delta_{jk}$.

Now, the optimal POVM can be obtained as a random projective measurement on the eigenbasis of U_1 and U_2. Explicitly, diagonalise each U_j as $U_j = \sum_{k=1}^{2} u_{jk} E_{jk}$, where $u_{jk} \in \mathbb{R}$ are eigenvalues and E_{jk} are rank-one projectors with $\sum_{k=1}^{2} E_{jk} = \mathbb{I}_2$. Define

$$p_j := \frac{\sqrt{\nu_j}}{\sqrt{\nu_1} + \sqrt{\nu_2}}, \qquad M(j,k) := p_j\, E_{jk}, \tag{124}$$

so that $\{M(j,k)\}_{j,k\in\{1,2\}}$ is a 4-outcome POVM satisfying $\sum_{j,k} M(j,k) = \mathbb{I}_2$. Now, introducing the coefficients

$$x_{jk} := p_j^{-1} \langle X, U_j\rangle_{\rho_0}^{\mathrm{S},(e)} u_{jk} + x, \qquad y_{jk} := p_j^{-1} \langle Y, U_j\rangle_{\rho_0}^{\mathrm{S},(e)} u_{jk} + y, \tag{125}$$

the first-moment constraints hold as follows: $\sum_{j,k} x_{jk} M(j,k) = M_1$ and $\sum_{j,k} y_{jk} M(j,k) = M_2$. It can be checked that the second-moment cost for this POVM and coefficients $\{x_{jk}, y_{jk}\}$ attains the minimum in Eq. (119):

$$\mathrm{Tr}\left[\rho_0 \sum_{jk} (x_{jk}^2 + y_{jk}^2) M(j,k)\right] = x^2 + y^2 + (\sqrt{\nu_1} + \sqrt{\nu_2})^2 = x^2 + y^2 + \mathrm{Tr}[\rho_0(X^2 + Y^2)] + \|\sqrt{\rho_0}[X,Y]\sqrt{\rho_0}\|_1\,; \tag{126}$$

this matches Eq. (120), since $x^2 + y^2 + \mathrm{Tr}[\rho_0(X^2 + Y^2)] = \mathrm{Tr}[\rho_0(M_1^2 + M_2^2)]$, and $[X,Y] = [M_1, M_2]$.

C.2 Closed form of the qubit NH bound

We now show that, for two parameters and a single qubit, the remaining minimisation in Eq. (117) also admits a closed form. We only prove the result for $L = \mathbb{I}_2$, as the bounds for other loss matrices can be obtained simply by reparametrization of the first moment operators (see footnote 4 in the main text).

Let S_i be the SPM operators for the qubit model, i.e., $\bar{\rho}_i = E_{\rho_0,\mathrm{S}}[S_i] = \frac{S_i\rho_0 + \rho_0 S_i}{2}$ with $i = 1, 2$, and also introduce their zero-mean version $\Delta S_i := S_i - \mu_i \mathbb{I}_2$, where $\mu_i = \mathrm{Tr}[\rho_0 S_i]$ are the prior mean values in Eq. (106). Introducing the real 2×2 Gram matrix of the zero-mean SPM operators

$$\mathcal{H}ij := \frac{1}{2}\,\mathrm{Tr}[\rho_0\{\Delta S_i, \Delta S_j\}] = \mathrm{Re}\,\mathrm{Tr}[\rho_0 \Delta S_i \Delta S_j] = \langle \Delta S_i, \Delta S_j\rangle_{\rho_0}^{\mathrm{S},(e)}, \tag{127}$$

which corresponds to the matrix-valued version of the precision gain in Eq. (109), the two-parameter qubit Bayesian NH bound is equivalent to

$$\mathcal{L}_{\mathrm{NH}} = \lambda - \mu_1^2 - \mu_2^2 - \lambda_{\max}(\mathcal{H}) = \mathcal{L}_{\mathrm{prior}} - \lambda_{\max}(\mathcal{H}) = \mathcal{L}_{\mathrm{SPM}} + \lambda_{\min}(\mathcal{H}), \tag{128}$$

where $\lambda_{\max}(\mathcal{H})$ and $\lambda_{\min}(\mathcal{H})$ denote the largest and smallest eigenvalues of of $\mathcal{H}$, respectively. The last equality holds because for $L = \mathbb{I}_2$ the SPM bound can be written as $\mathcal{L}_{\mathrm{SPM}} = \lambda - \sum_{i=1}^{2} \mathrm{Tr}[\bar{\rho}_i S_i] = \lambda - \mu_1^2 - \mu_2^2 - \mathrm{tr}\,\mathcal{H}$ and $\mathrm{tr}\,\mathcal{H} = \lambda_{\max}(\mathcal{H}) + \lambda_{\min}(\mathcal{H})$. The quantity $\lambda_{\min}(\mathcal{H})$ can be written explicitly in terms of a commutator between the SPM operators, yielding

$$\mathcal{L}_{\mathrm{NH}} = \mathcal{L}_{\mathrm{SPM}} + \frac{1}{2}\left[\mathrm{tr}\,\mathcal{H} - \sqrt{(\mathrm{tr}\,\mathcal{H})^2 - \|\sqrt{\rho_0}[S_1, S_2]\sqrt{\rho_0}\|_1^2}\right]. \tag{129}$$

Proof. Starting from Eq. (117), one can write any Hermitian first-moment operator M_i in terms of its zero-mean version X_i as

$$M_i = m_i \mathbb{I}_2 + X_i, \qquad \mathrm{Tr}[\rho_0 X_i] = 0. \tag{130}$$

From the SPM definition we have

$$\mathrm{Tr}[\bar{\rho}_i M_i] = \frac{1}{2}\,\mathrm{Tr}[\rho_0(S_i M_i + M_i S_i)] = \mu_i m_i + \mathrm{Re}\,\mathrm{Tr}[\rho_0 \Delta S_i X_i], \tag{131}$$

where the mixed terms vanish because both X_i and ΔS_i have zero mean with respect to ρ_0. Since the commutator term is invariant under shifts by multiples of the identity, the optimisation in Eq. (117) separates into a scalar part and a part involving the zero-mean operators.

$$\min_{M_1,M_2} \left\{\mathrm{Tr}\big[\rho_0(M_1^2 + M_2^2)\big] + \big\|\sqrt{\rho_0}[M_1, M_2]\sqrt{\rho_0}\big\|_1 - 2\,\mathrm{Tr}\big[\bar{\rho}_1 M_1 + \bar{\rho}_2 M_2\big]\right\} \tag{132}$$

$$= \min_{m_1,m_2} \sum_{i=1}^{2} \left(m_i^2 - 2\mu_i m_i\right) + \min_{X_1,X_2} \left\{\sum_{i=1}^{2} \left(\mathrm{Tr}\left[\rho_0 X_i^2\right] - 2\,\mathrm{Re}\,\mathrm{Tr}[\rho_0 \Delta S_i X_i]\right) + \big\|\sqrt{\rho_0}[X_1, X_2]\sqrt{\rho_0}\big\|_1\right\} \tag{133}$$

$$= -\mu_1^2 - \mu_2^2 + \min_{X_1,X_2} \left\{\sum_{i=1}^{2} \left(\mathrm{Tr}\left[\rho_0 X_i^2\right] - 2\,\mathrm{Re}\,\mathrm{Tr}[\rho_0 \Delta S_i X_i]\right) + \big\|\sqrt{\rho_0}[X_1, X_2]\sqrt{\rho_0}\big\|_1\right\}, \tag{134}$$

where the scalar part has been explicitly minimised by $m_i = \mu_i$.

Thus, what remains is a minimisation over the zero-mean Hermitian operators X_1, X_2 which belong to the real vector space $\mathcal{V} := \{X = X^\dagger : \mathrm{Tr}[\rho_0 X] = 0\}$. We employ the symmetric inner product $\langle A, B\rangle_{\rho_0}^{\mathrm{S},(e)}$ defined in Eq. (122) and introduce the Gram matrix of the operators X_1, X_2 (denoted collectively as X):

$$\Phi[\mathsf{X}]_{ij} = \langle X_i, X_j\rangle_{\rho_0}^{\mathrm{S},(e)}. \tag{135}$$

This Gram matrix can be used to express part of the objective function in Eq. (134). First, we have

$$\mathrm{Tr}\left[\rho_0\left(X_1^2 + X_2^2\right)\right] = \langle X_1, X_1\rangle_{\rho_0}^{\mathrm{S},(e)} + \langle X_2, X_2\rangle_{\rho_0}^{\mathrm{S},(e)} = \mathrm{tr}\,\Phi[\mathsf{X}]. \tag{136}$$

Second, for a single qubit, one can check explicitly the following identity by Nagaoka [88, Lemma 8]

$$\|\sqrt{\rho_0}[X_1, X_2]\sqrt{\rho_0}\|_1 = 2\sqrt{\langle X_1, X_1\rangle_{\rho_0}^{\mathrm{S},(e)}\langle X_2, X_2\rangle_{\rho_0}^{\mathrm{S},(e)} - \left(\langle X_1, X_2\rangle_{\rho_0}^{\mathrm{S},(e)}\right)^2} = 2\sqrt{\det\Phi[\mathsf{X}]}. \tag{137}$$

Notice that this identity was already implicitly used to show the attainability result in Eq. (126).

Thus, the remaining optimisation is rewritten in terms of the Gram matrix and inner product as

$$\min_{X_1, X_2} \mathrm{tr}\,\Phi[\mathsf{X}] + 2\sqrt{\det\Phi[\mathsf{X}]} - 2\langle X_1, \Delta S_1\rangle_{\rho_0}^{\mathrm{S},(e)} - 2\langle X_2, \Delta S_2\rangle_{\rho_0}^{\mathrm{S},(e)}. \tag{138}$$

Since $\Phi[\mathsf{X}]$ is a positive semidefinite 2×2 matrix, if its eigenvalues are $\alpha_1, \alpha_2 \geq 0$, then

$$\mathrm{tr}\,\Phi[\mathsf{X}] + 2\sqrt{\det\Phi[\mathsf{X}]} = \alpha_1 + \alpha_2 + 2\sqrt{\alpha_1\alpha_2} = (\sqrt{\alpha_1} + \sqrt{\alpha_2})^2 = \left(\mathrm{tr}\sqrt{\Phi[\mathsf{X}]}\right)^2. \tag{139}$$

To perform the last minimisation, choose an arbitrary real basis of $\mathcal{V}$, i.e. three linearly independent zero-mean Hermitian operators $\{F_a\}_{a=1}^3$. Then every $X_i \in \mathcal{V}$ and every $\Delta S_i \in \mathcal{V}$ can be represented uniquely by three-dimensional vectors

$$X_i = \sum_{a=1}^{3} x_{ai} F_a, \qquad \Delta S_i = \sum_{a=1}^{3} d_{ai} F_a. \tag{140}$$

Then we can introduce the matrices $X, D \in \mathbb{R}^{3\times 2}$ with columns $x_i = (x_{1i}, x_{2i}, x_{3i})^T$ and $d_i = (d_{1i}, d_{2i}, d_{3i})^T$, respectively. In this basis, the inner product $\langle\cdot,\cdot\rangle_{\rho_0}^{\mathrm{S},(e)}$ is represented by the real positive-definite matrix

$$G_{\rho_0,ab} := \langle F_a, F_b\rangle_{\rho_0}^{\mathrm{S},(e)}. \tag{141}$$

Therefore we can express the Gram matrices of X_1, X_2 and $\Delta S_1, \Delta S_2$ as

$$\Phi[X] = X^T G_{\rho_0} X, \qquad \mathcal{H} = D^T G_{\rho_0} D, \tag{142}$$

and the mixed product term as

$$\sum_{i=1}^{2} \langle X_i, \Delta S_i\rangle_{\rho_0}^{\mathrm{S},(e)} = \mathrm{tr}\, X^T G_{\rho_0} D. \tag{143}$$

The remaining optimisation becomes

$$\min_{X\in\mathbb{R}^{3\times 2}} \left\{\left(\mathrm{tr}\sqrt{X^T G_{\rho_0} X}\right)^2 - 2\,\mathrm{tr}\, X^T G_{\rho_0} D\right\}. \tag{144}$$

Now we can introduce the matrices $Y := G_{\rho_0}^{1/2} X$ and $Z := G_{\rho_0}^{1/2} D$ so that

$$X^T G_{\rho_0} X = Y^T Y, \qquad \mathrm{tr}\, X^T G_{\rho_0} D = \mathrm{tr}\, Y^T Z. \tag{145}$$

Hence Eq. (144) becomes

$$\min_{Y} \left\{\|Y\|_1^2 - 2\,\mathrm{tr}\, Y^T Z\right\}, \tag{146}$$

where $\|Y\|_1 := \mathrm{tr}\sqrt{Y^T Y}$ is the trace (nuclear) norm of a real matrix (i.e. the sum of the singular values). By the standard duality between the trace norm and the operator norm, i.e. the maximum singular value of a matrix denoted as $\|\cdot\|_\infty$, we have:

$$\mathrm{tr}\, Y^T Z \leq \|Y\|_1 \|Z\|_\infty, \tag{147}$$

which implies

$$\|Y\|_1^2 - 2\,\mathrm{tr}\,Y^T Z \geq \|Y\|_1^2 - 2\|Y\|_1\,\|Z\|_\infty \geq -\|Z\|_\infty^2. \tag{148}$$

where the last inequality follows by minimising the scalar $t^2 - 2t\|Z\|_\infty$ over $t \geq 0$. Equality is attained by taking $Y = \|Z\|_\infty uv^T$, where u, v are left and right singular vectors associated with the largest singular value of Z. Consequently,

$$\min_Y \left\{\|Y\|_1^2 - 2\,\mathrm{tr}\,Y^T Z\right\} = -\|Z\|_\infty^2. \tag{149}$$

Finally, by noticing that $Z^T Z = D^T G_{\rho_0} D = \mathcal{H}$, we have $\|Z\|_\infty^2 = \lambda_{\max}(\mathcal{H})$. Thus the zero-mean part of the optimisation contributes $-\lambda_{\max}(\mathcal{H})$, and substituting into Eq. (134) gives

$$\mathcal{L}_{\mathrm{NH}} = \lambda - \mu_1^2 - \mu_2^2 - \lambda_{\max}(\mathcal{H}). \tag{150}$$

The equivalent expression in Eq. (129) comes from rewriting $\lambda_{\min}(\mathcal{H})$ in terms of $\mathrm{tr}\,\mathcal{H} = \lambda_{\min}(\mathcal{H}) + \lambda_{\max}(\mathcal{H})$ and $\det \mathcal{H} = \lambda_{\min}(\mathcal{H})\lambda_{\max}(\mathcal{H})$, and using the identity in Eq. (137) to express the determinant in terms of the commutator.

□

D Matrix lower bounds from monotone metrics

D.1 Brief overview of monotone metrics

In this subsection we follow the notation of Ref. [112, Ch. 6.1], but see Ref. [137] for a recent survey on the topic of monotone metrics. For a given quantum state ρ, we define an inner product as follows:

$$\langle Y,\, X\rangle_\rho^{(e)} \;:=\; \mathrm{Tr}\left[Y^\dagger\, E_\rho[X]\right], \tag{151}$$

where E_ρ is a linear map satisfying

$$\mathrm{Tr}\left[X^\dagger E_\rho[X]\right] \geq 0, \tag{152}$$

$$\mathrm{Tr}\left[Y^\dagger E_\rho[X]\right] = \mathrm{Tr}\left[E_\rho[Y]^\dagger\, X\right], \tag{153}$$

and

$$E_\rho\left[U^\dagger X U\right] = U^\dagger\, E_{U\rho U^\dagger}[X]\; U, \tag{154}$$

$$E_\rho\left[\mathbb{I}\right] = \rho, \tag{155}$$

Following Hayashi's formulation, these maps and inner products are defined on arbitrary finite-dimensional operators, so X and Y need not be Hermitian. The choice of the map E_ρ determines the inner product. We will primarily work with a generic one; however, three notable examples are:

$$E_{\rho,\mathrm{S}}[X] = \frac{1}{2}\left(X\rho + \rho X\right) \tag{156}$$

$$E_{\rho,\mathrm{R}}[X] = X\rho \tag{157}$$

$$E_{\rho,\mathrm{sq}}[X] = \sqrt{\rho}X\sqrt{\rho}, \tag{158}$$

which we will call "symmetric", "right" and "square-root", respectively. As in the main text, we assume $\rho \succ 0$ and for the standard monotone maps we will consider next it ensures that E_ρ is invertible, so that we can introduce the dual inner product in terms of the inverse linear map E_ρ^{-1} without further regularisations:

$$\langle A,\, B\rangle_\rho^{(m)} \;=\; \mathrm{Tr}\left[E_\rho^{-1}[A]^\dagger B\right]; \tag{159}$$

it is denoted by the superscript (m), following Hayashi's notation. More specifically, in what follows we consider inner products with the standard classical normalisation: when ρ and X, Y are diagonal in the same basis, the dual inner product reduces to the classical Fisher-type expression $\langle X, Y\rangle_\rho^{(m)} = \sum_k \frac{X_{kk}^* Y_{kk}}{\rho_{kk}}$.

One of the crucial properties of such dual inner products is the monotonicity under the action of quantum channels, which makes them a good measure of information. In the following, we restrict to choices of E_ρ belonging to the standard class of monotone quantum metrics characterised by Petz [136, 137], or more generally to maps whose dual inner products satisfy the data-processing inequality[9]:

$$\langle A, A\rangle_\rho^{(m)} \geq \langle \Lambda[A], \Lambda[A]\rangle_{\Lambda[\rho]}^{(m)}, \tag{160}$$

[9] Petz's work gives a complete characterisation of normalised monotone Riemannian metrics on the tangent space of full-rank quantum states, i.e., on Hermitian traceless operators. Here we also use the associated inner products on the full operator space, so we explicitly assume that the chosen extension continues to satisfy Eq. (160).

where Λ is a quantum channel, i.e., a completely positive trace-preserving map, such that also the output state is full rank $\Lambda[\rho] \succ 0$. In the derivation below we will only use classical normalisation and this data-processing inequality.

For normalised monotone metrics in the Petz class, another important property which holds for Hermitian operators $A = A^\dagger$ is that the norm induced by this class of inner products is minimal and maximal for the symmetric and right versions, respectively:

$$\langle A, A\rangle_\rho^{\mathrm{S},(m)} \leq \langle A, A\rangle_\rho^{(m)} \leq \langle A, A\rangle_\rho^{\mathrm{R},(m)} . \tag{161}$$

Eq. (33) in the main text follows from the first inequality. In the derivation of the matrix lower bounds below, where complex linear combinations of posterior moments are used, we only rely on the monotonicity property (160), which holds for arbitrary operators.

D.2 Posterior mean operators

Introducing the MSL matrix, with elements

$$\Sigma_{ij} = \int \mathrm{d}x \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta}, x)\, (\tilde{\theta}_i(x) - \theta_i)(\tilde{\theta}_j(x) - \theta_j)\,, \tag{162}$$

where $p(\boldsymbol{\theta}, x) = p(\boldsymbol{\theta})\, p(x|\boldsymbol{\theta})$, the scalar MSL can be written as

$$\mathcal{L} = L^{ij}\Sigma_{ji} = \operatorname{tr} L\Sigma\,. \tag{163}$$

Classically, and thus also for quantum problems where the measurement is fixed, the optimal Bayes estimator that minimises the multiparameter MSL matrix is the PM:

$$\tilde{\theta}_{\mathrm{opt},i}(x) = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta}|x)\, \theta_i = \frac{\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, p(x|\boldsymbol{\theta})\, \theta_i}{\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, p(x|\boldsymbol{\theta})}, \tag{164}$$

which was already defined in Eq. (26) of the main text. Applying this estimator, the optimal MSL matrix has elements

$$\Sigma_{ij} = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \theta_i\theta_j - \int \mathrm{d}x\, \frac{\left[\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, p(x|\boldsymbol{\theta})\, \theta_i\right]\left[\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, p(x|\boldsymbol{\theta})\, \theta_j\right]}{\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, p(x|\boldsymbol{\theta})}\,. \tag{165}$$

Following Ref. [43], we now proceed to consider the positive scalar

$$\boldsymbol{u}^\dagger\Sigma\boldsymbol{u} = u^{*i}u^j\Sigma_{ij} = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, |\theta_{\boldsymbol{u}}|^2 - \int \mathrm{d}x\, \frac{\left|\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, p(x|\boldsymbol{\theta})\, \theta_{\boldsymbol{u}}\right|^2}{\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, p(x|\boldsymbol{\theta})}, \tag{166}$$

where $\boldsymbol{u} \in \mathbb{C}^d$ is a complex vector and $\theta_{\boldsymbol{u}} = \boldsymbol{u}^\dagger\boldsymbol{\theta} = u^{*i}\theta_i$ is a linear combination of parameters. For quantum systems, the outcome probability for a given value of the parameters is obtained from Born's rule $p(x|\boldsymbol{\theta}) = \operatorname{Tr}[M(x)\rho(\boldsymbol{\theta})]$, which leads to a complexified version of Eq. (6) in Ref. [43]:

$$\boldsymbol{u}^\dagger\Sigma\boldsymbol{u} = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, |\theta_{\boldsymbol{u}}|^2 - \int \mathrm{d}x\, \frac{|\operatorname{Tr}[M(x)\bar{\rho}_{\boldsymbol{u}}]|^2}{\operatorname{Tr}[M(x)\rho_0]}\,, \tag{167}$$

where $\bar{\rho}_{\boldsymbol{u}} = u^{*i}\bar{\rho}_i = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \rho(\boldsymbol{\theta})\, \theta_{\boldsymbol{u}}$.

The same result can be arrived at from the formalism of Sec. 2.3 with the posterior mean estimator Eq. (11) being

$$\tilde{\theta}_j(x) = \frac{\operatorname{Tr}[\bar{\rho}_j M(x)]}{\operatorname{Tr}[\rho_0 M(x)]},$$

assuming a parameter-independent measurement. As such, the operators M_j and M_{jk} introduced in Eq. (17) take the form

$$M_i = \int \mathrm{d}x\, M(x)\, \frac{\operatorname{Tr}[\bar{\rho}_i M(x)]}{\operatorname{Tr}[\rho_0 M(x)]}, \quad M_{ij} = \int \mathrm{d}x\, M(x)\, \frac{\operatorname{Tr}[\bar{\rho}_i M(x)]\operatorname{Tr}[\bar{\rho}_j M(x)]}{\operatorname{Tr}[\rho_0 M(x)]^2}; \tag{168}$$

the application in Eq. (14) then leads to Eq. (167).

A crucial observation from Ref. [43] is that the second term resembles the classical Fisher information and so it is possible to apply a Braunstein–Caves–like inequality [71] and obtain a matrix SPM lower bound on Σ. Here, we take a different route and use a more general family of inequalities which stem from the monotonicity

of the inner products defined in the previous sections. This allows us to derive a whole family of matrix lower bounds, similarly to the case of local QET.

For simplicity, we focus on a discrete set of measurement outcomes x, which can always be encoded on a register composed by an orthonormal basis of vectors $\{|e_x\rangle\}$. We introduce the measurement channel

$$\kappa[\bullet] \coloneqq \sum_x \mathrm{Tr}[M(x)\bullet]|e_x\rangle\langle e_x|\,; \tag{169}$$

this is an entanglement-breaking quantum channel that formally represents the map between quantum states and probability distributions (diagonal quantum states) implemented by the POVM with elements $M(x)$. Outcomes with $\mathrm{Tr}[M(x)\rho_0] = 0$ can be discarded, or equivalently one works on the support of $\kappa[\rho_0]$. Applying this channel to ρ_0 and $\bar{\rho}_{\boldsymbol{u}}$ we obtain

$$\begin{aligned}\kappa[\rho_0] &= \sum_x \mathrm{Tr}\left[M(x)\rho_0\right]|e_x\rangle\langle e_x|,\\ \kappa[\bar{\rho}_{\boldsymbol{u}}] &= \sum_x \mathrm{Tr}\left[M(x)\bar{\rho}_{\boldsymbol{u}}\right]|e_x\rangle\langle e_x|.\end{aligned} \tag{170}$$

These operators are diagonal in the orthonormal basis $\{|e_x\rangle\}$ and they commute with each other; hence, by the classical normalisation stated above, the dual inner product gives exactly the second term of Eq. (167):

$$\langle \kappa[\bar{\rho}_{\boldsymbol{u}}], \kappa[\bar{\rho}_{\boldsymbol{u}}]\rangle^{(m)}_{\kappa[\rho_0]} = \mathrm{Tr}\left[E^{-1}_{\kappa[\rho_0]}(\kappa[\bar{\rho}_{\boldsymbol{u}}])^\dagger \kappa[\bar{\rho}_{\boldsymbol{u}}]\right] = \sum_x \frac{\left|\mathrm{Tr}\left[M(x)\bar{\rho}_{\boldsymbol{u}}\right]\right|^2}{\mathrm{Tr}\left[M(x)\rho_0\right]}. \tag{171}$$

Therefore, by the monotonicity property (160) we finally obtain that

$$\sum_x \frac{\left|\mathrm{Tr}\left[M(x)\bar{\rho}_{\boldsymbol{u}}\right]\right|^2}{\mathrm{Tr}\left[M(x)\rho_0\right]} \leq \langle \bar{\rho}_{\boldsymbol{u}}, \bar{\rho}_{\boldsymbol{u}}\rangle^{(m)}_{\rho_0} = \langle L_{\boldsymbol{u}}, L_{\boldsymbol{u}}\rangle^{(e)}_{\rho_0}, \tag{172}$$

where now $L_{\boldsymbol{u}}$ is a generic PM operator obtained as

$$L_{\boldsymbol{u}} \coloneqq E^{-1}_{\rho_0}[\bar{\rho}_{\boldsymbol{u}}] = u^{*i} E^{-1}_{\rho_0}[\bar{\rho}_i] = u^{*i} L_i, \tag{173}$$

where we have also introduced PM operators for each original parameter, satisfying the equations $E_{\rho_0}(L_i) = \bar{\rho}_i$. Such PM operators generalise the symmetric version introduced in the main text in terms of the map $E_{\rho,\mathrm{S}}$. For example, they also include the operators introduced in Ref. [105, Theorem 8], where they are called "λ logarithmic derivatives".

From the inequality in Eq. (172), the lower bound for the scalar introduced before immediately follows:

$$\boldsymbol{u}^\dagger \Sigma \boldsymbol{u} \geq \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, |\theta_{\boldsymbol{u}}|^2 - \boldsymbol{u}^\dagger \mathcal{K} \boldsymbol{u}, \tag{174}$$

where we have introduced the complex-valued Hermitian positive semidefinite matrix with elements

$$\mathcal{K}_{ij} \coloneqq \langle \bar{\rho}_j, \bar{\rho}_i\rangle^{(m)}_{\rho_0} = \langle L_j, L_i\rangle^{(e)}_{\rho_0} = \mathrm{Tr}[L_j^\dagger E_{\rho_0}[L_i]]\,; \tag{175}$$

notice that with this index convention, $\boldsymbol{u}^\dagger \mathcal{K} \boldsymbol{u} = \langle \bar{\rho}_{\boldsymbol{u}}, \bar{\rho}_{\boldsymbol{u}}\rangle^{(m)}_{\rho_0} \geq 0$. By choosing the map $E_{\rho,\mathrm{S}}$, this matrix becomes the real-valued SPM matrix introduced in Ref. [43]. However, for other choices, such as $E_{\rho,\mathrm{R}}$, this matrix is still Hermitian but complex-valued. Regardless of the particular choice of monotone metric, since the scalar inequality holds for an arbitrary choice of $\boldsymbol{u} \in \mathbb{C}^d$, we obtain the family of matrix lower bounds on the MSL matrix

$$\Sigma \succeq \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \boldsymbol{\theta}\boldsymbol{\theta}^T - \mathcal{K}\,. \tag{176}$$

This is again a generalisation of the SPM matrix bound obtained in Ref. [43].

D.3 Scalar bounds from matrix inequalities

We start from the matrix inequality in Eq. (176) and denote the right-hand side by the symbol $\mathcal{B} \coloneqq \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\, \boldsymbol{\theta}\boldsymbol{\theta}^\top - \mathcal{K}$, where $\mathcal{K} \geq 0$ is a *complex-valued* Hermitian positive semidefinite matrix. We now want to show that also the imaginary part of $\mathcal{K}$ sets constraints on the minimum MSL $\mathcal{L}$. This fact is well-known [120]; for completeness, we provide a proof here, roughly following Ref. [10].

First, let us rewrite the function to minimise, subject to the constraint (176), as the trace of a symmetric matrix $\operatorname{tr} L\Sigma = \operatorname{tr}\sqrt{L}\Sigma\sqrt{L}$. From (176) we can obtain the two matrix inequalities

$$\sqrt{L}\,\Sigma\sqrt{L} \geq \sqrt{L}\,(\operatorname{Re}\mathcal{B} \pm i\operatorname{Im}\mathcal{B})\,\sqrt{L}, \tag{177}$$

where the two signs correspond to taking the transpose of the two sides of the original equation. Here $\operatorname{Re}\mathcal{B}$ and $\operatorname{Im}\mathcal{B}$ denote the entrywise real and imaginary parts, so that $\operatorname{Re}\mathcal{B}$ is real symmetric and $\operatorname{Im}\mathcal{B}$ is real antisymmetric. Now, this matrix inequality implies

$$\sum_k \boldsymbol{z}_k^\dagger \sqrt{L}\,(\Sigma - \operatorname{Re}\mathcal{B})\,\sqrt{L}\boldsymbol{z}_k \geq \pm\sum_k \boldsymbol{z}_k^\dagger \left(i\sqrt{L}\operatorname{Im}\mathcal{B}\sqrt{L}\right)\boldsymbol{z}_k\,, \tag{178}$$

for any set of complex-valued vectors $\{\boldsymbol{z}_i\}$. Since $i\sqrt{L}\operatorname{Im}\mathcal{B}\sqrt{L}$ is Hermitian, we can choose its eigenbasis as the vectors $\boldsymbol{z}_i$; noting the corresponding eigenvalues as λ_i, we obtain

$$\operatorname{tr} L(\Sigma - \operatorname{Re}\mathcal{B}) \geq \sum_i |\lambda_i| = \operatorname{tr}|\sqrt{L}\operatorname{Im}\mathcal{B}\sqrt{L}| = \|\sqrt{L}\operatorname{Im}\mathcal{B}\sqrt{L}\|_1. \tag{179}$$

The absolute values are obtained by applying the positive inequality separately to eigenvectors with positive eigenvalues and the negative inequality separately to eigenvectors with negative eigenvalues, and then summing the resulting inequalities. The matrix $|\sqrt{L}\operatorname{Im}\mathcal{B}\sqrt{L}|$, where $|A| \coloneqq \sqrt{A^\dagger A}$ denotes the operator absolute value, is a positive semidefinite Hermitian matrix that can be diagonalised in the same eigenbasis as $i\sqrt{L}\operatorname{Im}\mathcal{B}\sqrt{L}$, with eigenvalues $|\lambda_i|$. Moreover, $\|A\|_1 = \operatorname{tr}|A| = \operatorname{tr}\sqrt{A^\dagger A}$ is the Schatten 1-norm, also known as the trace norm or nuclear norm. This inequality for the trace of the real part of a positive semidefinite complex matrix is known as the Belavkin–Grishanin inequality [78, 102].

Assuming $L \succ 0$ (otherwise one should restrict to the support of L), the minimum can be achieved by setting

$$\Sigma = \operatorname{Re}\mathcal{B} + \sqrt{L^{-1}}|\sqrt{L}\operatorname{Im}\mathcal{B}\sqrt{L}|\sqrt{L^{-1}}\,, \tag{180}$$

as seen from direct computation.

In conclusion, we have proven that

$$\min_{\Sigma\geq\mathcal{B}} \operatorname{tr} L\Sigma = \operatorname{tr} L\operatorname{Re}\mathcal{B} + \|\sqrt{L}\operatorname{Im}\mathcal{B}\sqrt{L}\|_1. \tag{181}$$

Substituting back the original right-hand side of the inequality (176), we get

$$\mathcal{L} \geq \lambda - \operatorname{Tr}[L\operatorname{Re}\mathcal{K}] + \|\sqrt{L}\operatorname{Im}\mathcal{K}\sqrt{L}\|_1, \tag{182}$$

since $\lambda = \operatorname{tr} L\int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\,\boldsymbol{\theta}\boldsymbol{\theta}^T = \int \mathrm{d}\boldsymbol{\theta}\, p(\boldsymbol{\theta})\,\boldsymbol{\theta}^T L\,\boldsymbol{\theta}$, and $\operatorname{Im}\mathcal{B} = -\operatorname{Im}\mathcal{K}$, given that the other matrix terms are real-valued by assumption. In general, this bound is not necessarily tight, as it is not always possible to find a measurement and estimator that produce a MSL matrix as in (180).

We finally note that the bound in Ref. [105, Theorem 8] in terms of the λ logarithmic derivatives is also in the class of scalar bounds derived in this Appendix.

E Phase-dephasing multi-copy expressions

For the phase-dephasing model introduced in Sec. 4.2, the full single-copy zeroth and first state moments are

$$\rho_0 = \frac{1}{2}\left\{\mathbb{I}_2 + \frac{\sin\left(\frac{W_1}{2}\right)\left[\operatorname{Ei}(-W_2) - \operatorname{Ei}\left(-\frac{1}{W_2}\right)\right]}{W_1\log(W_2)}\sigma_x\right\}, \tag{183a}$$

$$\bar{\rho}_1 = \frac{1}{2W_1\log(W_2)}\left[\frac{W_1}{2}\cos\left(\frac{W_1}{2}\right) - \sin\left(\frac{W_1}{2}\right)\right]\left[\operatorname{Ei}(-W_2) - \operatorname{Ei}\left(-\frac{1}{W_2}\right)\right]\sigma_y, \tag{183b}$$

$$\bar{\rho}_2 = \frac{\sin(W_1/2)}{2W_1}\left(\frac{W_2\Xi(-W_2) - \frac{1}{W_2}\Xi(-\frac{1}{W_2})}{2\log(W_2)} + \frac{\operatorname{Ei}(-W_2) + \operatorname{Ei}(-\frac{1}{W_2})}{2} - \gamma\right)\sigma_x, \tag{183c}$$

where $\operatorname{Ei}(z)$ denotes the exponential integral, $\Xi(z) = {}_3F_3((1,1,1),(2,2,2),z)$ is a generalised hypergeometric function, and γ is Euler's constant.

By virtue of the tensor product structure—although the moments themselves are not direct tensor products—the multi-copy state $\rho(\boldsymbol{\theta}) = \varrho(\boldsymbol{\theta})^{\otimes n}$ consists of terms $\prod_{j,k=1}^{d} [\varrho_{j,k}(\boldsymbol{\theta})]^{\kappa_{j,k}}$ for a d-dimensional system. For the multi-copy phase and dephasing example of Sec. 4.2, the pertinent integrals for the respective moments are given below. For ρ_0:

$$\int \mathrm{d}\boldsymbol{\theta}\, \frac{[\varrho_{1,1}(\boldsymbol{\theta})]^j [\varrho_{2,2}(\boldsymbol{\theta})]^n}{\theta_2 W_1 \log W_2} = \frac{1}{2^{j+n}}, \tag{184a}$$

$$\int \mathrm{d}\boldsymbol{\theta}\, \frac{[\varrho_{1,1}(\boldsymbol{\theta})]^j [\varrho_{1,2}(\boldsymbol{\theta})]^k [\varrho_{2,1}(\boldsymbol{\theta})]^k [\varrho_{2,2}(\boldsymbol{\theta})]^n}{\theta_2 W_1 \log W_2} = \frac{\mathrm{Ei}(-2kW_2) - \mathrm{Ei}\left(-\frac{2k}{W_2}\right)}{2^{j+2k+n+1} \log W_2}, \tag{184b}$$

$$\int \mathrm{d}\boldsymbol{\theta}\, \frac{[\varrho_{1,1}(\boldsymbol{\theta})]^j [\varrho_{1,2}(\boldsymbol{\theta})]^k [\varrho_{2,1}(\boldsymbol{\theta})]^m [\varrho_{2,2}(\boldsymbol{\theta})]^n}{\theta_2 W_1 \log W_2} = \frac{\left[\mathrm{Ei}(-(k+m)W_2) - \mathrm{Ei}\left(-\frac{(k+m)}{W_2}\right)\right] \sin\left(\frac{(k-m)W_1}{2}\right)}{2^{j+k+m+n}(k-m) W_1 \log W_2}, \tag{184c}$$

where Eq. (184b) holds for $k > 0$, and Eq. (184c) for $k \neq m$. For $\bar{\rho}_1$:

$$\int \mathrm{d}\boldsymbol{\theta}\, \frac{[\varrho_{1,1}(\boldsymbol{\theta})]^j [\varrho_{1,2}(\boldsymbol{\theta})]^k [\varrho_{2,1}(\boldsymbol{\theta})]^k [\varrho_{2,2}(\boldsymbol{\theta})]^n \theta_1}{\theta_2 W_1 \log W_2} = 0, \tag{185a}$$

$$\begin{aligned}\int \mathrm{d}\boldsymbol{\theta}\, \frac{[\varrho_{1,1}(\boldsymbol{\theta})]^j [\varrho_{1,2}(\boldsymbol{\theta})]^k [\varrho_{2,1}(\boldsymbol{\theta})]^m [\varrho_{2,2}(\boldsymbol{\theta})]^n \theta_1}{\theta_2 W_1 \log W_2} &= \frac{i\left[\mathrm{Ei}(-(k+m)W_2) - \mathrm{Ei}\left(-\frac{k+m}{W_2}\right)\right]}{2^{j+k+m+n}(k-m)^2 W_1 \log W_2} \\ &\quad \times \left[\sin\left(\frac{(k-m)W_1}{2}\right) - \frac{(k-m)W_1}{2}\cos\left(\frac{(k-m)W_1}{2}\right)\right],\end{aligned} \tag{185b}$$

where Eq. (185b) holds for $k \neq m$. For $\bar{\rho}_2$:

$$\int \mathrm{d}\boldsymbol{\theta}\, \frac{[\varrho_{1,1}(\boldsymbol{\theta})]^j [\varrho_{2,2}(\boldsymbol{\theta})]^n \log\theta_2}{\theta_2 W_1 \log W_2} = 0, \tag{186a}$$

$$\int \mathrm{d}\boldsymbol{\theta}\, \frac{[\varrho_{1,1}(\boldsymbol{\theta})]^j [\varrho_{1,2}(\boldsymbol{\theta})]^k [\varrho_{2,1}(\boldsymbol{\theta})]^k [\varrho_{2,2}(\boldsymbol{\theta})]^n \log\theta_2}{\theta_2 W_1 \log W_2} = \frac{\xi_{2k}(W_1, W_2)}{2^{j+2k+n+1}}, \tag{186b}$$

$$\int \mathrm{d}\boldsymbol{\theta}\, \frac{[\varrho_{1,1}(\boldsymbol{\theta})]^j [\varrho_{1,2}(\boldsymbol{\theta})]^k [\varrho_{2,1}(\boldsymbol{\theta})]^m [\varrho_{2,2}(\boldsymbol{\theta})]^n \log\theta_2}{\theta_2 W_1 \log W_2} = \frac{\sin\left(\frac{(k-m)W_1}{2}\right) \xi_{k+m}(W_1, W_2)}{2^{j+k+m+n} W_1 (k-m)} \tag{186c}$$

where

$$\xi_q(W_1, W_2) = \frac{q}{2\log W_2}\left(W_2 \Xi(-qW_2) - \frac{\Xi\left(-\frac{q}{W_2}\right)}{W_2}\right) + \frac{\mathrm{Ei}\left(-qW_2\right) + \mathrm{Ei}\left(-\frac{q}{W_2}\right)}{2} - \log(q) - \gamma,$$

and Eq. (186b) holds for $k > 0$, and Eq. (186c) for $k \neq m$.