PURPOSE-LED PUBLISHING™

PAPER • OPEN ACCESS

Benchmarking Bayesian quantum estimation

View the article online for updates and enhancements.

You may also like

Quantum Science and Technology

Benchmarking Bayesian quantum estimation

Valeria Cimini[1,*] , **Emanuele Polino**[1,2] , **Mauro Valeri**[1], **Nicolò Spagnolo**[1] and **Fabio Sciarrino**[1]

[1] Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 5, I-00185 Roma, Italy
[2] Centre for Quantum Dynamics and Centre for Quantum Computation and Communication Technology, Griffith University, Yuggera Country, Brisbane, Queensland 4111, Australia
* Author to whom any correspondence should be addressed.

E-mail: valeria.cimini@uniroma1.it

Abstract

The quest for precision in parameter estimation is a fundamental task in different scientific areas. The relevance of this problem thus provided the motivation to develop methods for the application of quantum resources to estimation protocols. Within this context, Bayesian estimation offers a complete framework for optimal quantum metrology techniques, such as adaptive protocols. However, the use of the Bayesian approach requires extensive computational resources, especially in the multiparameter estimations that represent the typical operational scenario for quantum sensors. Hence, the requirement to characterize protocols implementing Bayesian estimations can become a significant challenge. This work focuses on the crucial task of robustly benchmarking the performances of these protocols in both single and multiple-parameter scenarios. By comparing different figures of merits, evidence is provided in favor of using the median of the quadratic error in the estimations in order to mitigate spurious effects due to the numerical discretization of the parameter space, the presence of limited data, and numerical instabilities. These results, providing a robust and reliable characterization of Bayesian protocols, find natural applications to practical problems within the quantum estimation framework.

1. Introduction

Parameter estimation is a fundamental requirement of most scientific studies. To this end, a fundamental goal consists of developing method to extract accurate information about the unknown parameters of interest from observed data. Usually, in order to achieve sufficiently high precision, measurements are repeated several times, allowing the gathering of statistical information about the studied parameters which are then treated as random variables. Measurement outcomes are indeed used to reconstruct an estimator that correctly addresses the values of the investigated parameters. The two most widely adopted and efficient estimators belong to different perspectives: the so-called frequentist approach and the Bayesian one [1]. The first gives a description of the estimate as a function of the experimental outcomes and often uses the maximum likelihood estimator [2]. The latter relies on describing the process in view of the current knowledge on the parameters, i.e. probability distributions, using both the information gained with the measurement and the prior information on the parameter values to reconstruct, through Bayes's theorem, the posterior probability distribution [3]. Such a distribution is then used to give both the estimate of the parameters under investigation and of its precision [4].

Independently from the adopted approach, the interest of the metrology field is to identify which is the ultimate achievable sensitivity, intrinsically limited by the nature of the investigated process and the underlying statistical model, and which are the estimation strategies allowing the saturation of such bounds [5]. This brought to the investigation of different kinds of probes allowing the optimization of the estimation task. In particular, the use of quantum probes revealed to promise the capability of achieving improved estimation precision with respect to classical ones, thus disclosing applications in different tasks from biological sensing [6] to gravitational wave detection [7]. For this reason, quite recently the emergent field of

quantum metrology has redefined the boundaries of sensitivity, reachable when exploiting quantum resources [8–12].

However, often the amount of available quantum resources is limited. Therefore, to achieve optimal estimation performances is necessary to properly engineer either the probe preparation or the measurement settings, or both, often based on the information acquired during the measurement process through adaptive estimation strategies [13–18]. This optimization process can be done through different online and offline [19–23] protocols. One of the most implemented ones is developed within the Bayesian inference framework, and is based on updating the knowledge of the parameter value depending on the registered measurement outcome and on the setting of some control parameter [24–27] valid for both the qubit framework [28] and the continuous-variable one [29]. The advantage becomes particularly significant in the multiparameter framework where the interest relies on the simultaneous estimation of several parameters [30–33]. In this regime, adaptive strategies take into account the covariance structure of the parameters, optimizing the allocation of resources to estimate the parameters with the greatest precision, outperforming non-adaptive methods.

A central challenge in multiparameter estimation based on quantum strategies is the identification of an appropriate figure of merit that quantifies estimation precision and the exploration of the associated precision bounds. In this work, we analyze which are the correct figures of merit to address in the limited data regime [34–36] considering the Bayesian framework where both the estimation and the optimization parameters are retrieved with a discrete computation. Therefore, the appropriate figure of merit used to benchmark the quality of the estimate should also consider all the numerical caveats related to the estimation algorithm.

Here, we examine the different figures of merit for benchmarking estimation processes in single and multiparameter quantum scenarios. Given the Bayesian framework, we consider the aforementioned caveats in the numerical calculations and the issue of the dimension of the estimation problems, within different measurement strategies. We develop a general procedure to tackle the performance benchmarking of different protocols, focusing on the median as a metric able to mitigate the presence of outliers that are intrinsic to the Bayesian estimation and to compare general protocols. A relation is demonstrated between the achieved performances and both the number of parameters and the discretization necessary to retrieve the estimate. In this way, our results provide an effective working procedure that can be adapted to different systems and scenarios to select the best-performing estimation algorithms in quantum metrology problems.

2. Figures of merit in estimation protocols

In the quest for the identification of the most appropriate probe and measurement strategy, a crucial analysis is to compare the achieved results with the fundamental precision bounds of the investigated process. Considering an unknown parameter ϕ or a vector of parameters $\vec{\phi} = (\phi_1, \phi_2, \dots, \phi_p)$ in a multiparameter case, a central role is played by the Cramér–Rao bound (CRB) [37]. Such a bound sets a lower fundamental limit on the variance (the covariance matrix $\Sigma(\vec{\phi})$ in the multiparameter case) of unbiased estimators:

$$\Delta^2 \phi \geqslant \frac{1}{NF(\phi)}; \qquad \Sigma\left(\vec{\phi}\right) \geqslant \frac{\mathbb{F}^{-1}\left(\vec{\phi}\right)}{N}, \tag{1}$$

where N represents the number of employed independent probes, $F(\phi)$ and $\mathbb{F}(\vec{\phi})$ are the Fisher Information and the Fisher Information matrix respectively [38–40]. Although the CRB serves as a fundamental benchmark of the achievable ultimate precision, other relevant bounds proper of the Bayesian framework have been introduced, such as the Van Trees and the Ziv-Zakai bounds [41–43].

In the limit of a large number of probes, i.e. $N \gg 1$, both the maximum likelihood estimator and the Bayesian one converge to the CRB. Indeed they are efficient estimators [11], and therefore in this regime, the sensitivity reached by the frequentist and Bayesian methods asymptotically agrees. However, in real experimental conditions, it is usually desired to exploit the minimum number of probes to show such a convergence. This is possible by optimizing the measurement and the probe preparation in order to extract the highest possible amount of information from each probe. This allows a faster convergence to the bound and the possibility of saturating it already in the limited resource regime.

In the quantum scenario, implementing such optimization protocol, however, is not a trivial task in particular for multiparameter problems. Indeed, the solution of the optimization problem requires in general expansive computational efforts since it involves the computation of complex multidimensional integrals. To speed up such computation an efficient solution is represented by the sequential Monte Carlo (SMC) [44]

algorithm which has indeed been already employed to solve different problems from Hamiltonian learning [44, 45], to state tomography [46] and parameter estimation [24, 26, 47] often using an adaptive scheme.

The precision of an estimator in quantum metrology is usually assessed in terms of its variance. In the Bayesian scenario, this figure of merit corresponds to the variance of the posterior distribution $P(\phi|x)$ updated after the measurement outcome x:

$$\sigma^2 = \int \mathrm{d}\phi\, P(\phi|x) \left[\tilde{\phi}(x) - \phi\right]^2, \tag{2}$$

where $\tilde{\phi}(x)$ represents the estimate of the parameter, obtained as follows:

$$\tilde{\phi}(x) = \int \mathrm{d}\phi\, P(\phi|x)\, \phi. \tag{3}$$

The estimator generalizes to the multiparameter scenario as:

$$\tilde{\phi}_i(\vec{x}) = \int \phi_i P\left(\vec{\phi}|\vec{x}\right) \prod_j \mathrm{d}\phi_j, \tag{4}$$

thus considering the posterior mean for each of the parameter ϕ_i. Analogously, the variance of the posterior is replaced by its covariance matrix according to the standard definition.

However, the variance (covariance matrix) alone may not fully capture the performance of the estimator since it is related only to the spread of the estimator probability distribution, and it does not take into account potential systematic errors or biases. To face this issue, in practical applications, another figure of merit that is often taken into consideration, allowing a more comprehensive evaluation of the achieved estimation performances, is the quadratic loss, i.e. the quadratic error. The latter is defined as the quadratic difference between the estimated parameter $\tilde{\phi}$ and its true values ϕ^{true} (ϕ_i and its true values $\tilde{\phi}_i$ for the multiparameter case):

$$Q(\phi) = \left(\tilde{\phi} - \phi^{\text{true}}\right)^2 \tag{5}$$

$$Q\left(\vec{\phi}\right) = \sum_{i=1}^{p} \left(\tilde{\phi}_i - \phi_i^{\text{true}}\right)^2. \tag{6}$$
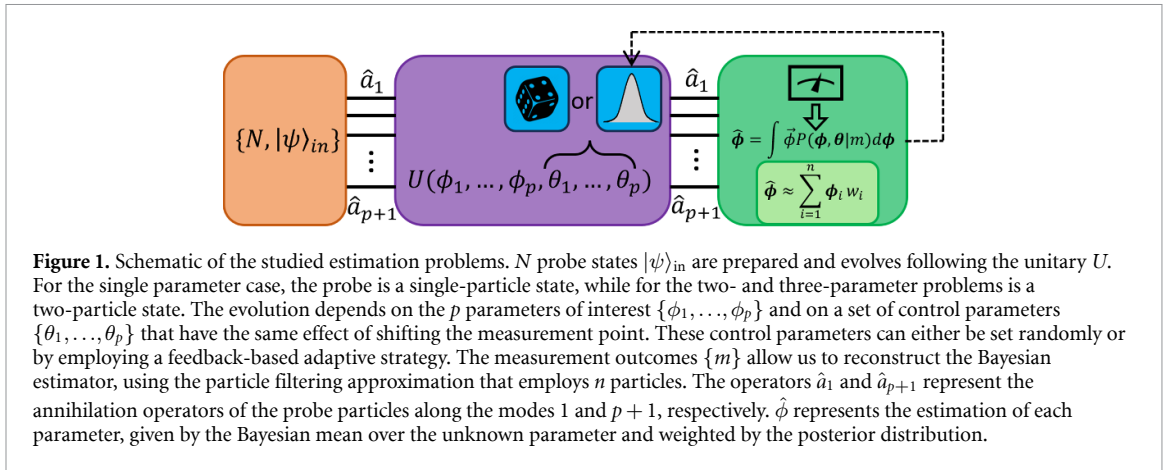
The quadratic loss naturally embeds both the variance and the bias in the estimator, often resulting in a more robust quantity to evaluate the estimation performances. While in actual experiments the true values of the parameters are unknown and thus the quadratic loss is inaccessible, this quantity is a precious tool to pre-calibrate the performances of estimation strategies, for instance through numerical simulations that take into account a detailed model of the apparatus. Furthermore, the quadratic loss is the preferred choice with respect to other possible figures of merits, e.g. the mean absolute error, given that for an unbiased estimator such a quantity is bounded by the Fisher information (or the Fisher information matrix in the multiparameter case) as per the Cramer–Rao bound.

3. Benchmarking quantum metrology strategies

We now focus on practical case studies of Bayesian quantum metrology estimations, within which we study the performances of protocols as a function of the relevant parameters of the processes by using different figures of merit. To this goal, we consider a test-bed example starting from the estimation of the rotation angle ϕ of a single-qubit state. Then, we enter in the multiparameter regime studying two- and three-parameter estimation problems by considering a testbench platform with two-particle probe states and multiarm interferometers. The general estimation process considered here is schematized in figure 1.

3.1. Single parameter case
We start considering the estimate of ϕ encoded in the single-qubit state $|\psi(\phi)\rangle = 1/\sqrt{2}\left(|0\rangle \pm \mathrm{e}^{\mathrm{i}\phi}|1\rangle\right)$, with $\phi \in [0, 2\pi)$. This coincides also, in the optical case, with the output single-photon state of a two-arm interferometer. We study the performances achieved with numerical simulations when adopting a Bayesian estimator updated after each measurement outcome. Quantum parameter estimation protocols featuring non-monotonic likelihood probabilities, such as most phase estimation problems, have an ambiguity issue. Indeed, a fixed probability can be generated by more than one phase even if the parameters are inside the

Figure 1. Schematic of the studied estimation problems. N probe states $|\psi\rangle_{\text{in}}$ are prepared and evolves following the unitary U. For the single parameter case, the probe is a single-particle state, while for the two- and three-parameter problems is a two-particle state. The evolution depends on the p parameters of interest $\{\phi_1, \ldots, \phi_p\}$ and on a set of control parameters $\{\theta_1, \ldots, \theta_p\}$ that have the same effect of shifting the measurement point. These control parameters can either be set randomly or by employing a feedback-based adaptive strategy. The measurement outcomes $\{m\}$ allow us to reconstruct the Bayesian estimator, using the particle filtering approximation that employs n particles. The operators $\hat{a}_1$ and $\hat{a}_{p+1}$ represent the annihilation operators of the probe particles along the modes 1 and $p+1$, respectively. $\hat{\phi}$ represents the estimation of each parameter, given by the Bayesian mean over the unknown parameter and weighted by the posterior distribution.

same periodicity interval. In these cases, it is necessary to implement strategies that can distinguish between these ambiguities, and this can be achieved by changing the measurement settings during the estimation process. For this purpose, and also for faster convergence of the estimation to the precision bound, adaptive protocols can be used [16, 48]. Different adaptive algorithms are characterized, in general, by different performances.

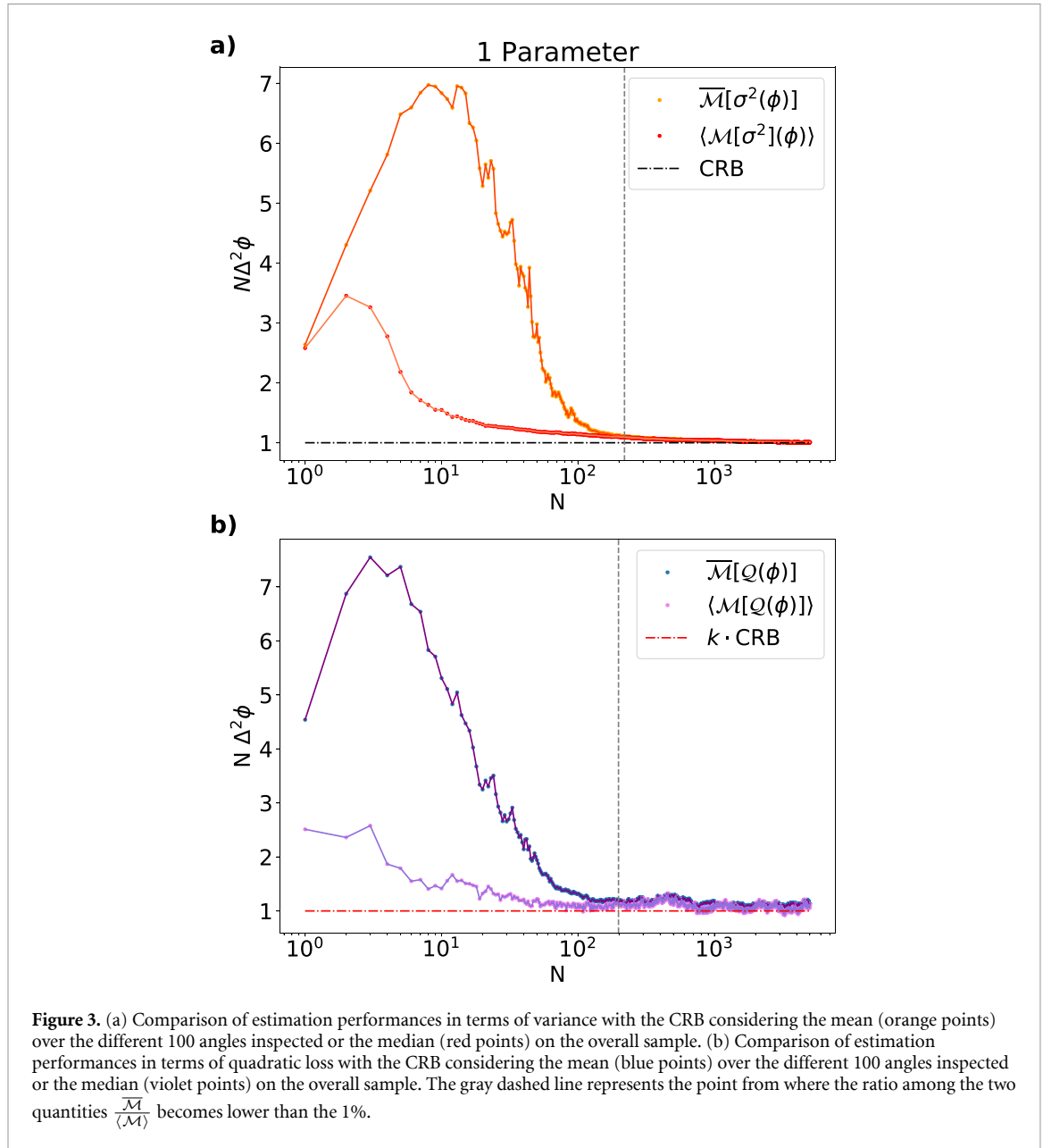In order to make the results of this work independent of the specific optimization protocol chosen, we focus only on the easiest and computationally efficient protocol that consists of shifting the measurement settings randomly after the measurement of each probe. More specifically, to disambiguate the phase values in the whole periodicity interval, it is necessary to adopt a procedure that changes the measurement settings during the estimation experiment. The easiest and computationally efficient one is to shift the measurement settings randomly after the measurement of each probe. This random control-setting strategy permits to saturate the CRB in the single parameter estimation. The reported performances are obtained considering $M = 100$ different rotation angles ϕ chosen randomly in $[0, 2\pi)$ and repeating the estimate of each phase $r = 30$ times to study its robustness. We study the performances as a function of the number of probes N, and we compare them with the CRB that provides a lower bound on the achievable estimation precision. The performances are computed both in terms of variance and quadratic loss and the distributions, obtained from the respective histogram over the 30 repetitions, for 10 of the 100 investigated phases are reported in figure 2. As it appears from the comparison of the distributions of different repetitions for each rotation angle, the influence of the outliers on the mean value of the quadratic loss is greater. This is observed from the tails of the distributions, while those of the variance are more shrunk around a central value.

The outliers above can impact the benchmarking procedure of a specific strategy. For this reason, it can be useful to replace the analysis of the mean with the median of the set of different estimation runs. The different behavior of the distribution is due to the fact that for N sufficiently large, the posterior becomes a Gaussian distribution centered around the true value of the parameter and with variance given by the inverse of the Fisher information. It follows that the variance is Gaussian distributed. In this case, the mean and the median coincide, and the bound of the two figures of merit remains the CRB. Hence, the common choice is to use as Bayesian estimator the mean of the posterior distribution (equation (3)), which provides the results of a given estimation experiment by sending N probes. Additionally, the variance of the posterior provides an estimation of the degree of confidence of the single run. Conversely, as described in section 2, if one needs to evaluate the overall performances of a given protocol, it is necessary to look at repeated estimation experiments via the quadratic loss. Considering it as a random variable, the quadratic loss is the square of a variable described by a Gaussian distribution, resulting in a positively skewed distribution. In this case, the mean is greater than the median, and the proportionality factor among them can be estimated numerically computing the ratio of the median of a generic squared Gaussian distribution centered in zero and its mean. This factor, in the single-parameter case, results equal to $k \simeq 0.4549$ [25] and it must be taken into consideration when comparing the performances in terms of medians with the ultimate precision bound. Moreover, we underline that when comparing the overall estimation performances to the precision bound, after having considered either the mean or the median over the different repetitions of the estimate (that we indicate respectively with $M[\cdot]$ and $\mathcal{M}[\cdot]$), it is necessary to consider the average performances among all the different parameter values. Also in this case, there are some differences in the considered figure of merit but only in the range of very low probes as reported in figure 3. Hereafter, we indicate the mean over different angles as $\overline{\phi}$ while the median as $\langle\phi\rangle$.

Figure 2. Normalized probability distribution of the performances of the estimation algorithm over different repetitions. The performances are reported in terms of quadratic loss values (purple) and variances (orange) obtained over the 30 different repetitions for each phase value ϕ, after sending $N = 300$ probes. The distributions are obtained with the kernel density estimation [49] which gives the estimated probability density function smoothed on the histograms of the data. The performances are reported for 10 selected values among the 100 tested ones in order to cover the whole periodicity interval.
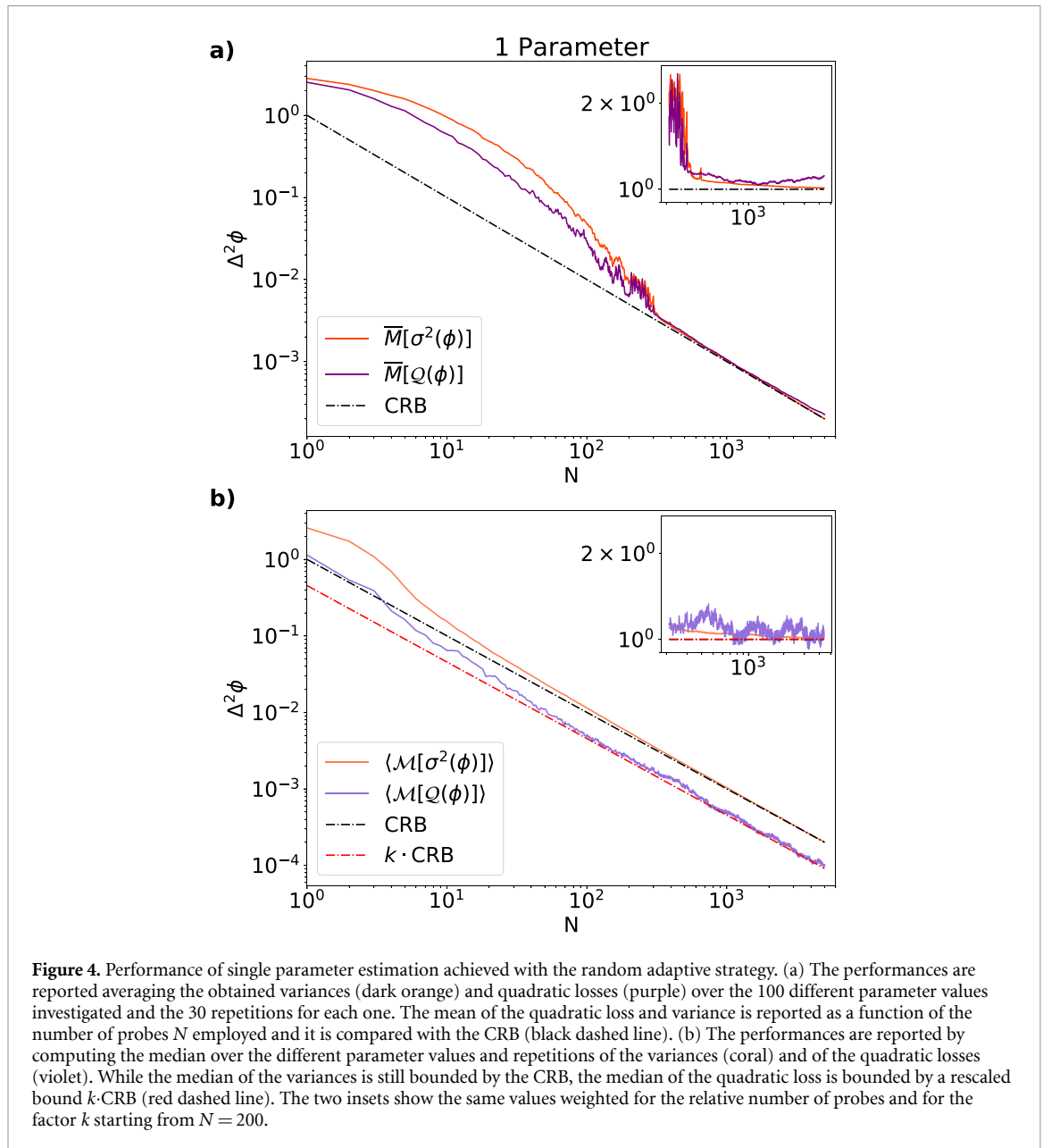
Due to the shape of the quadratic loss distribution, it becomes more sensitive to outliers. The conventional reliance on the mean estimate across different repetitions of the estimate and parameter values can be problematic as it may not converge to the CRB, largely because of the impact of these outliers. This may thus include a bias when benchmarking or comparing different techniques. This is shown in panel (a) of figure 4 where it is reported the mean of the achieved variance and quadratic loss computed over the same set of data. The median, on the other hand, offers distinct advantages, especially in scenarios where resources are limited. By adopting the median and considering the associated bound, as reported in panel (b) of figure 4 we observe a clear convergence to the bound, also for the quadratic loss, with fewer probes compared to using the mean estimate. In such a way, using the median provides an approach to identify the convergence to the ultimate bound of a given estimation algorithm by reducing the impact of outliers.

As previously discussed, our approach to computing the necessary integrals for Bayesian estimation involved discretizing the parameter space into n possible values covering the entire periodicity interval. This discretization was conducted in accordance with SMC, also known as the particle filtering method [44], used in recent multiparameter estimation experiments [24–26]. The objective was to reconstruct the posterior probability distribution of the parameter and then use it to estimate its value. Such an approach is based on choosing a set of n random particles $(\vec{\phi}_i, w_i)$, where $\vec{\phi}_i$ are vectors of specific phase values in the p-dimensional parameter space and w_i are the associated weights, that are representative of the posterior distribution. This is done initializing the particles by drawing n phase vectors $\vec{\phi}_i$ from the prior distribution, and by setting $w_i = 1/n$. At each step of the experiment, when an outcome m has been measured, the weights are updated as $w_i \rightarrow w_i p(m|\vec{\phi}_i, \vec{\theta})$, where $\vec{\theta}$ is the (eventual) vector of control parameters used in the experiment. As shown in references [24, 26, 44], to avoid concentration of the weight values only to a small subsets of the particles, that may occur during the update process, resampling techniques can be used to appropriately reset the particles values. At the end of the estimation process, the calculation of the required multi-dimensional integrals for Bayesian protocols is then replaced by a finite sum over the particle points $\vec{\phi}_i$, weighted by the parameters w_i.
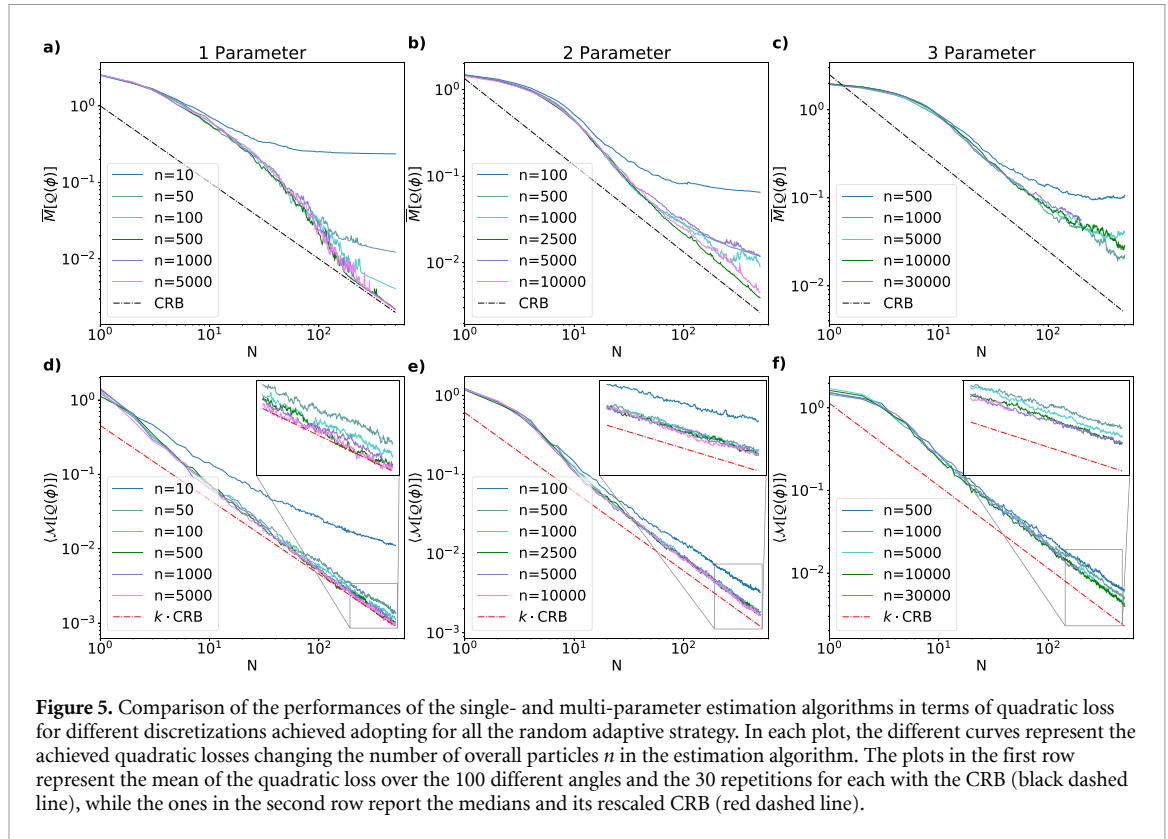
Figure 3. (a) Comparison of estimation performances in terms of variance with the CRB considering the mean (orange points) over the different 100 angles inspected or the median (red points) on the overall sample. (b) Comparison of estimation performances in terms of quadratic loss with the CRB considering the mean (blue points) over the different 100 angles inspected or the median (violet points) on the overall sample. The gray dashed line represents the point from where the ratio among the two quantities $\frac{\overline{\mathcal{M}}}{\langle\mathcal{M}\rangle}$ becomes lower than the 1%.

It is important to recognize that the convergence to the ultimate bound of precision is closely tied to the choice of the number of particles *n* used in this process. The number of particles employed in the estimation process impacts how well the parameter space is sampled and how effectively the estimation algorithm converges. A higher number of particles can improve precision, but it also increases computational requirements. Therefore, the selection of the appropriate number of particles is not straightforward but rather involves a trade-off between reliability and computational cost that in turn increases with the number of parameters. The convergence to the bound will therefore be related also to such choice. Indeed, the number of particles affects the ultimate achievable precision that in a first approximation, and in the absence of resampling, can be expressed as $\left(\frac{2\pi}{n}\right)^{p}$, where *p* indicates the number of parameters being estimated. The performances expressed as both the mean and the median of the obtained quadratic losses are reported for different discretization choices for single- and multi-parameter problems in figure 5. The obtained performances overcome the simple discretization limitation since all the simulations are done employing resampling. The influence of finite discretization results in the attainment of a plateau value, corresponding to the limited sensitivity achievable with such discretization, as it appears in the plots of the first row of figure 5. In these plots the average achieved quadratic losses are reported as a function of the number of injected probe states. The effect of discretization of the integration space is attenuated when considering instead the median of the achieved results, providing also in this case a more robust and reliable measure of

Figure 4. Performance of single parameter estimation achieved with the random adaptive strategy. (a) The performances are reported averaging the obtained variances (dark orange) and quadratic losses (purple) over the 100 different parameter values investigated and the 30 repetitions for each one. The mean of the quadratic loss and variance is reported as a function of the number of probes N employed and it is compared with the CRB (black dashed line). (b) The performances are reported by computing the median over the different parameter values and repetitions of the variances (coral) and of the quadratic losses (violet). While the median of the variances is still bounded by the CRB, the median of the quadratic loss is bounded by a rescaled bound $k \cdot$CRB (red dashed line). The two insets show the same values weighted for the relative number of probes and for the factor k starting from $N = 200$.

the general performances achieved by the estimation algorithm, filtering out fluctuations arising from the finite discretization.

3.2. Multiparameter case

We now consider a second scenario, involving the estimate of two or three phases embedded in a multiarm interferometers injected with two-photon states. This choice is motivated by the fact that such states and platforms are now standard in quantum optics research and are successfully used for quantum multiparameter estimations [24, 26, 43, 50, 51]. Starting from the scheme reported in figure 4, the applied transformation in the considered single and multiparameter scenarios is $U(\vec{\phi}, \vec{\theta}) = e^{i \sum_{k=1}^{p} \hat{n}_k (\phi_k + \theta_k)}$, with $\hat{n}_k$ the particle number operator of the k mode (for the multiparameter scenarios more details on the transformations in the interferometers are given in the appendix). Also in this case, the estimators considered are the mean values of the unknown phases calculates over the Bayesian posterior distribution updated during the estimation process. As for the single parameter case, we first study the strategy where the controls are varied randomly during the estimation process. As observed in figure 5, we find that while the bound is saturated for the single parameter problem, for the analyzed system, this is not the case in the multiparameter framework. The reason why it remains a bias from the achieved quadratic loss and the bound is related to the choice of the adaptive protocol. In the multiparameter scenario, the shape of the likelihood function, and as a consequence the Fisher information landscape, is more complex, and not flat

Figure 5. Comparison of the performances of the single- and multi-parameter estimation algorithms in terms of quadratic loss for different discretizations achieved adopting for all the random adaptive strategy. In each plot, the different curves represent the achieved quadratic losses changing the number of overall particles *n* in the estimation algorithm. The plots in the first row represent the mean of the quadratic loss over the 100 different angles and the 30 repetitions for each with the CRB (black dashed line), while the ones in the second row report the medians and its rescaled CRB (red dashed line).

over the parameter space. Therefore in order to disambiguate among different parameter values in a limited-data regime, a more efficient adaptive strategy involving a feedback loop between the measurement outcomes and the experimental settings is required.

The realm of possible adaptive techniques is varied. Among the others, it includes machine learning-based algorithms [27, 52], variational techniques [53, 54], Bayesian updates based on minimizing the variance of the posterior [24–26], particle swarm optimization [19], genetic algorithms [23], or differential evolution [20, 55]. The choice of one of these efficient adaptive protocols, involving a feedback loop, becomes crucial to achieve the ultimate sensitivity in multiparameter scenarios. This is often necessary to disambiguate multiple parameters in a limited-data regime with increased complexity of the likelihood function. This is evident in figure 6 where, already for the two-parameter problem, it is necessary to adopt a minimization algorithm with a feedback loop to select the measurement settings. This is required to converge towards the ultimate precision bound. Indeed in this regime, the parameters under estimation often exhibit correlations, where changes in one parameter can significantly impact the estimation of others resulting in interdependent parameter uncertainties. Adaptive strategies allow for dynamic adjustments of measurements, taking into account the evolving estimates of all parameters, thereby optimizing resource allocation and minimizing the trade-offs among the uncertainties associated with each parameter [33].

In this scenario, utilizing the median as a metric of merit, as opposed to the mean, offers distinct advantages in benchmarking the performances of an estimation strategy also for a reduced discretization. Specifically, when dealing with a system where the Fisher information exhibits divergences, it becomes increasingly challenging to achieve high-precision estimations for those points in close proximity to the divergent regions, especially when the number of particles is not sufficiently high. As a result, the effectiveness of the minimization algorithm becomes strongly contingent on the specific parameter values being considered. Consequently, if the estimation algorithm does not perform optimally for only a small subset of the 100 phases under investigation, it significantly impacts the overall performance when assessed in terms of the mean. This issue can be mitigated by examining the median, which helps reduce the influence of these outliers, allowing to enhance the robustness of the benchmarking process against the constraints imposed by finite discretization, which are linked only to numerical limitations. Therefore, also in this case, the median provides a more accurate representation of the estimation quality itself.

Note that the goal here is to test estimation techniques in a robust way, mitigating the presence of outliers and instability effects due to numerical discretization. This is a different problem with respect to the case where one considers the performances of a single measurement. Here, the outliers have a non-zero probability to occur, thus having an impact in the single estimation process.
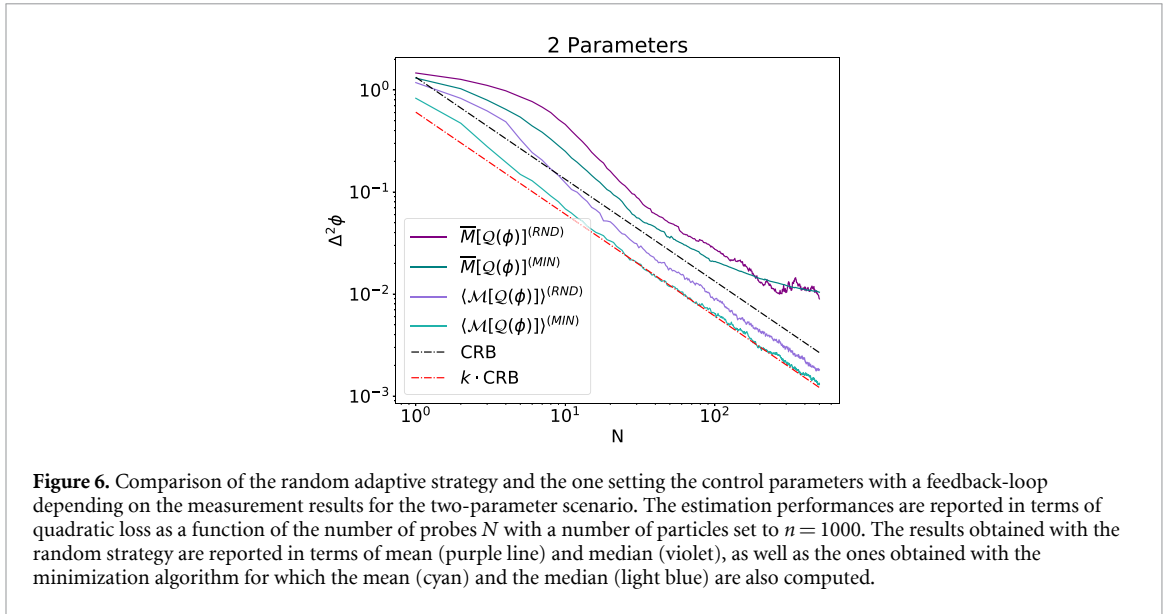
Figure 6. Comparison of the random adaptive strategy and the one setting the control parameters with a feedback-loop depending on the measurement results for the two-parameter scenario. The estimation performances are reported in terms of quadratic loss as a function of the number of probes N with a number of particles set to $n = 1000$. The results obtained with the random strategy are reported in terms of mean (purple line) and median (violet), as well as the ones obtained with the minimization algorithm for which the mean (cyan) and the median (light blue) are also computed.

4. Practical guidelines

Driven by the above results, we discuss below heuristic guidelines for general estimation problems. It becomes beneficial for practical implementations to have a working procedure that allows the identification of the best performing estimation protocol. However, to perform a fair comparison among the adopted approaches it becomes important to avoid alteration due to outliers that are intrinsic to the shot-by-shot Bayesian procedure. Therefore, defining a rule for the considered system that provides a systematic framework for achieving the minimum uncertainty while taking into account constraints such as the number of probes and the level of discretization become crucial in this kind of protocol. This can help in the assessment of resources for specific precision goals, optimizing these parameters to achieve the best trade-off between achievable precision and available resources.

As the number of parameters increases, it becomes important also to deal with the presence of possible correlations among their uncertainties, meaning it becomes more challenging to achieve precise estimates for all the parameters simultaneously. The challenge is compounded by the fact that the number of particles required to show a convergence to the ultimate bound also grows with the number of parameters. The median unique feature of smoothing these trends, reducing the effects of specific phase values, enables the derivation of a generic model describing how the precision is related to the number of parameters and the number of particles employed in the discretization process. The trends retrieved are shown in figure 7 where the relation among the distance from the achieved uncertainty and the relative bound is reported for different numbers of particles n and for problems with a different number of investigated parameters p. Notably, these trends are verified independently from the number of probes employed, provided that their number is sufficiently large to grant the saturation to the bound in the adaptive configuration.
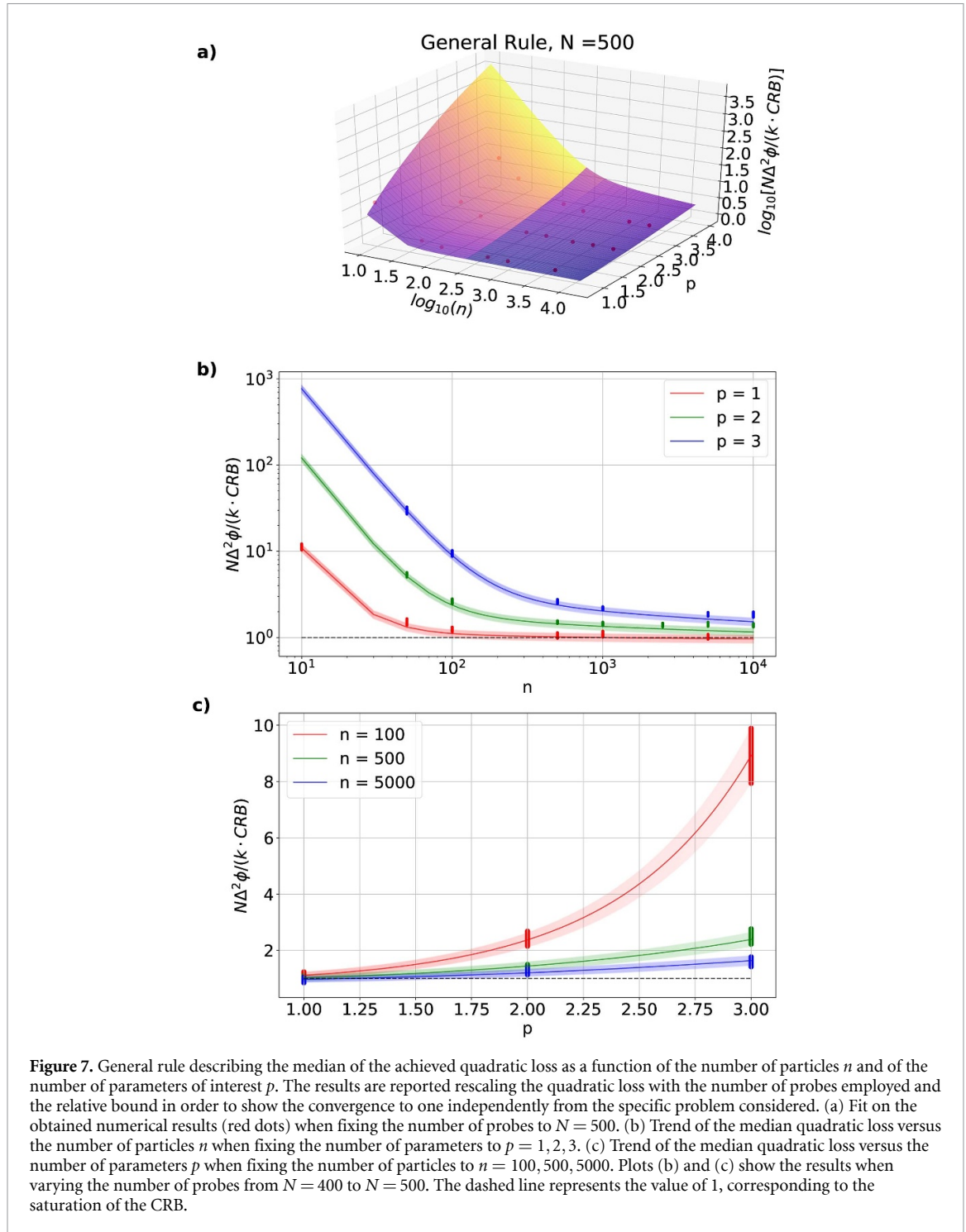
To obtain a rule governing the performance scaling of the considered system, we have performed a multidimensional fit of the median of the obtained quadratic loss over the last 100 probes i.e. $400 \leqslant N \leqslant 500$. The results are then adjusted for the relative bound, which varies depending on the number of parameters being estimated. The considered inputs for the fitting process are the number of investigated parameters and the number of particles employed in the Bayesian estimate. The function retrieved with the heuristic process is:

$$f(n,p) = A\left(1 + \sqrt{p}\right)^2 p\, n^{(-B+Cp)} + D\sqrt{p} + E\frac{p^F}{n^G}, \tag{7}$$

with

$$\begin{aligned} A \simeq 160, \quad B \simeq 5.4, \quad C \simeq 0.8, \quad D \simeq 0.002, \\ E \simeq 0.11, \quad F \simeq 3.5, \quad G \simeq 1.3. \end{aligned} \tag{8}$$

This function has been obtained by making considerations on the exponential scaling of the space of the discretization grid and its relation with the number of parameters i.e. the factor n^{-b+cp} linked with the scaling of the quantum fisher information with the number of parameters [38] giving the factor $a(1 + \sqrt{p})^2 p$.

Figure 7. General rule describing the median of the achieved quadratic loss as a function of the number of particles n and of the number of parameters of interest p. The results are reported rescaling the quadratic loss with the number of probes employed and the relative bound in order to show the convergence to one independently from the specific problem considered. (a) Fit on the obtained numerical results (red dots) when fixing the number of probes to $N = 500$. (b) Trend of the median quadratic loss versus the number of particles n when fixing the number of parameters to $p = 1, 2, 3$. (c) Trend of the median quadratic loss versus the number of parameters p when fixing the number of particles to $n = 100, 500, 5000$. Plots (b) and (c) show the results when varying the number of probes from $N = 400$ to $N = 500$. The dashed line represents the value of 1, corresponding to the saturation of the CRB.

Other terms are instead needed in order to take into account the specific behavior of the likelihood of the investigated system and the impossibility of disambiguating all the phase values with the random strategy.

This heuristic approach can be also used to estimate the computational cost associated to the online data processing required for the protocol in the particle approximation. Given a number of particles n employed, chosen for instance according to the function $f(n, p)$, each step of the protocol requires the estimation of the likelihood $p(m|\vec{\phi}_i, \vec{\theta})$, whose cost $C[p(m|\vec{\phi}_i, \vec{\theta})]$ depends on the specific system. Then, the main computational cost associated to the update of the weights for a single event is thus $O(nC[p(m|\vec{\phi}_i, \vec{\theta})])$, to be multiplied by N (the number of events). This provides the impact of the number of particles employed to the computational cost. Additional costs are provided by those corresponding to the resampling process, and eventual requirements for the application of adaptive strategies.

We note that this work focused on phase estimation protocols. Nevertheless, we expect that the approach can be adapted to general Bayesian estimation problems. The same issues we studied will affect general

Bayesian protocols and the expression in equation (7) can likely describe the performances as a function of the number of particles, parameters and probes, with suitable parameters values, that can be different from those in equation (8).

In conclusion, while some parameters may indeed be related to the investigated system's specific characteristics, the procedure outlined provides a valid approach that can be effectively generalized to retrieve the scaling behavior of Bayesian estimation processes. Indeed, the model described in equation (7) incorporates theoretical consideration on the relationship between parameter space dimensionality and the performance metrics of the estimation process, valid independently of the p value. The described procedure permits to obtain the expected scaling behavior of the specific system of interest, giving an advantage in predicting estimation performance and thereby efficiently allocating the dedicated resources with the potential to advance the understanding and application of these techniques in diverse scientific contexts.

5. Discussions

The aim of this work was to analyze how to robustly benchmark Bayesian estimation protocols, addressing both single and multiparameter scenarios. In the Bayesian framework, an essential role is played by numerical calculations that approximate the belief propagation and allow for optimized protocols such as adaptive ones. Moreover, any practical estimation process has to face finite statistics, in which only a limited amount of resources are used. In this context, a crucial task for quantum metrology is to benchmark the optimality of different estimation protocols. A primary question is to find the right figure of merit able to assess and compare the considered approaches.

In order to highlight the most effective estimation protocol, we have performed a comparative analysis, demonstrating the benefit of using the median of quadratic error as a metric to assess the estimation performances, mitigating effects from numerical discretization, limited data, and outliers inherent in Bayesian estimation. This choice becomes pivotal for benchmarking multiparameter estimation protocols where complexities arise from interdependencies between parameters. Quantifying the relation between uncertainty and number of parameters and particles revealed a heuristic model capturing performance trends. This model aids in resource optimization, enabling trade-offs between achievable precision and computational constraints.

This approach offers a pathway for optimizing resource allocation and achieving optimal precision in estimation processes. In this way, our analysis represents a tool with direct application to disparate metrology problems and platforms: from gravitational wave detection [7], biological sensing [6] to sensing with Gaussian states [56].

Data availability statement

The data cannot be made publicly available upon publication because they are not available in a format that is sufficiently accessible or reusable by other researchers. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

Appendix. Multiphase estimation in multiarm interferometers

As a testbench platform to investigate two- and three-parameter scenarios, we considered multiarm interferometers injected by input states composed of two indistinguishable photons.

For the two-parameter case, we consider a 3-arm interferometer composed of three cascaded transformations: a first unitary tritter transformation U_T the tritter, the phase shifts operator to be estimated, and a second tritter transformation U_T, where:

$$U_\mathrm{T} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{i2/3\pi} & e^{i4/3\pi} \\ 1 & e^{i4/3\pi} & e^{i2/3\pi} \end{bmatrix}. \tag{A1}$$

The overall unitary transformation of the interferometer is this $U^{\text{tot}} = U_{\text{T}} U(\vec{\phi}) U_{\text{T}}$. The photonic states we consider are those generated by two photons injected along the first two arms of the interferometer. More specifically, the probe state is prepared by the first tritter transformation, and for the two-photon input state results to be:

$$
\begin{aligned}
|\psi_0\rangle_{\text{T}} = {} & \frac{\sqrt{2}}{3}\left(|200\rangle + e^{\imath 2\pi/3}|020\rangle + e^{\imath 4\pi/3}|002\rangle\right) \\
& + \frac{1}{3}\Bigg[\left(1 + e^{\imath 2\pi/3}\right)|110\rangle + \left(1 + e^{\imath 4\pi/3}\right)|101\Big\rangle \\
& + \left(e^{\imath 2\pi/3} + e^{\imath 4\pi/3}\right)|011\rangle,
\end{aligned}
\tag{A2}
$$

where $|n_1 n_2 n_2\rangle$ stands for a Fock states with n_i photons in mode i.

Similarly, for the three-parameter scenario, we consider a 4-arm interferometer where both the initial and final unitaries are composed by a quarter described by the following transformation:

$$
U_{\text{Q}} = \frac{1}{2}\begin{bmatrix} 1 & i & i & -1 \\ i & -1 & 1 & i \\ i & 1 & -1 & i \\ -1 & i & i & 1 \end{bmatrix}
\tag{A3}
$$

By injecting two photons in the first two modes ($|1100\rangle$) the probe state generated by the quarter transformation is found to be:

$$
\begin{aligned}
|\psi_0\rangle_{\text{Q}} = {} & \frac{i}{2\sqrt{2}}\left(|2000\rangle - |0200\rangle + |0020\rangle + \right. \\
& \left. -|0002\rangle\right) - \frac{1}{2}\left(|1100\rangle + |0011\rangle\right).
\end{aligned}
\tag{A4}
$$

Such a scenario has been investigated experimentally in [24] to demonstrate quantum multiphase estimation even in adaptive regimes.

ORCID iDs

Valeria Cimini ⓘ https://orcid.org/0000-0001-8513-3766
Emanuele Polino ⓘ https://orcid.org/0000-0003-1438-8942
Nicolò Spagnolo ⓘ https://orcid.org/0000-0003-3471-2252
Fabio Sciarrino ⓘ https://orcid.org/0000-0003-1715-245X

References

[1] Li Y, Pezzé L, Gessner M, Ren Z, Li W and Smerzi A 2018 *Entropy* **20** 628
[2] Lane A S, Braunstein S L and Caves C M 1993 *Phys. Rev. A* **47** 1667
[3] Helstrom C W 1969 *J. Stat. Phys.* **1** 231
[4] Kramer S C and Sorenson H W 1988 *IEEE Trans. Autom. Control* **33** 217
[5] Liu J, Zhang M, Chen H, Wang L and Yuan H 2022 *Adv. Quantum Technol.* **5** 2100080
[6] Taylor M A and Bowen W P 2016 *Phys. Rep.* **615** 1
[7] Abbott B P *et al* 2016 *Phys. Rev. Lett.* **116** 061102
[8] Giovannetti V, Lloyd S and Maccone L 2006 *Phys. Rev. Lett.* **96** 010401
[9] Giovannetti V, Lloyd S and Maccone L 2011 *Nat. Photon.* **5** 222
[10] Demkowicz-Dobrzański R and Maccone L 2014 *Phys. Rev. Lett.* **113** 250801
[11] Polino E, Valeri M, Spagnolo N and Sciarrino F 2020 *AVS Quantum Sci.* **2** 024703
[12] Barbieri M 2022 *PRX Quantum* **3** 010202
[13] Wheatley T A, Berry D W, Yonezawa H, Nakane D, Arao H, Pope D T, Ralph T C, Wiseman H M, Furusawa A and Huntington E H 2010 *Phys. Rev. Lett.* **104** 093601
[14] Sekatski P, Skotiniotis M, Kołodyński J and Dür W 2017 *Quantum* **1** 27
[15] Demkowicz-Dobrzański R, Czajkowski J and Sekatski P 2017 *Phys. Rev. X* **7** 041009
[16] Wiseman H M 1995 *Phys. Rev. Lett.* **75** 4587
[17] Wiseman H M and Milburn G J 2009 *Quantum Measurement and Control* (Cambridge University Press)
[18] Higgins B L, Berry D W, Bartlett S D, Wiseman H M and Pryde G J 2007 *Nature* **450** 393
[19] Hentschel A and Sanders B C 2011 *Phys. Rev. Lett.* **107** 233601
[20] Lovett N B, Crosnier C, Perarnau-Llobet M and Sanders B C 2013 *Phys. Rev. Lett.* **110** 220501
[21] Cimini V, Mellini M, Rampioni G, Sbroscia M, Leoni L, Barbieri M and Gianani I 2019 *Opt. Express* **27** 35245
[22] Lumino A, Polino E, Rab A S, Milani G, Spagnolo N, Wiebe N and Sciarrino F 2018 *Phys. Rev. Appl.* **10** 044033
[23] Rambhatla K, D'Aurelio S E, Valeri M, Polino E, Spagnolo N and Sciarrino F 2020 *Phys. Rev. Res.* **2** 033078
[24] Valeri M *et al* 2023 *Phys. Rev. Res.* **5** 013138

[25] Belliardo F, Cimini V, Polino E, Hoch F, Piccirillo B, Spagnolo N, Giovannetti V and Sciarrino F 2022 Optimizing quantum-enhanced bayesian multiparameter estimation in noisy apparata (arXiv:2211.04747 [quant-ph])

[26] Valeri M, Polino E, Poderini D, Gianani I, Corrielli G, Crespi A, Osellame R, Spagnolo N and Sciarrino F 2020 *npj Quantum Inf.* **6** 92

[27] Cimini V, Valeri M, Polino E, Piacentini S, Ceccarelli F, Corrielli G, Spagnolo N, Osellame R and Sciarrino F 2023 *Adv. Photon.* **5** 016005

[28] Teklu B, Olivares S and Paris M G 2009 Bayesian estimation of one-parameter qubit gates *J Phys. B: Atomic, Molecular and Optical Physics* **42** 035502

[29] Olivares S and Paris M G 2009 Bayesian estimation in homodyne interferometry *J Phys. B: Atomic, Molecular and Optical Physics* **42** 055506

[30] Lee K K, Gagatsos C N, Guha S and Ashok A 2022 *IEEE J. Sel. Top. Signal Process.* **17** 491

[31] Gebhart V, Smerzi A and Pezzè L 2021 *Phys. Rev. Appl.* **16** 014035

[32] Datta A, Demkowicz-Dobrzański R and Liu J 2021 *J. Phys. A: Math. Theor.* **54** 460301

[33] Albarelli F, Barbieri M, Genoni M G and Gianani I 2020 *Phys. Lett.* A **384** 126311

[34] Rubio J and Dunningham J 2019 *New J. Phys.* **21** 043037

[35] Rubio J and Dunningham J 2020 *Phys. Rev.* A **101** 032114

[36] Meyer J J, Khatri S, França D S, Eisert J and Faist P 2023 Quantum metrology in the finite-sample regime (arXiv:2307.06370 [quant-ph])

[37] Paris M G A 2009 *Int. J. Quantum Inf.* **07** 125

[38] Liu J, Yuan H, Lu X-M and Wang X 2020 *J. Phys. A: Math. Theor.* **53** 023001

[39] Goldberg A Z, Sánchez-Soto L L and Ferretti H 2021 *Phys. Rev. Lett.* **127** 110501

[40] Barbieri M, Gianani I and Genoni M G 2023 Fisher information susceptibility for multiparameter quantum estimation (arXiv:2312.02035 [quant-ph])

[41] Trees H L V and Bell K L 2007 *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking* (Wiley-IEEE press)

[42] Tsang M 2012 *Phys. Rev. Lett.* **108** 230401

[43] D'Aurelio S E *et al* 2022 *Quantum Sci. Technol.* **7** 025011

[44] Granade C E, Ferrie C, Wiebe N and Cory D G 2012 *New J. Phys.* **14** 103013

[45] Wiebe N, Granade C, Ferrie C and Cory D G 2014 *Phys. Rev. Lett.* **112** 190501

[46] Huszár F and Houlsby N M 2012 *Phys. Rev.* A **85** 052120

[47] Wiebe N and Granade C 2016 *Phys. Rev. Lett.* **117** 010503

[48] Berry D W and Wiseman H M 2000 *Phys. Rev. Lett.* **85** 5098

[49] Turlach B A *et al* 1993 Bandwidth selection in kernel density estimation: a review *Technical Report* (Humboldt Universitaet Berlin)

[50] Polino E, Riva M, Valeri M, Silvestri R, Corrielli G, Crespi A, Spagnolo N, Osellame R and Sciarrino F 2019 *Optica* **6** 288

[51] Kim D-H, Hong S, Kim Y-S, Kim Y, Lee S-W, Pooser R C, Oh K, Lee S-Y, Lee C and Lim H-T 2024 *Nat. Commun.* **15** 266

[52] Fiderer L J, Schuff J and Braun D 2021 *PRX Quantum* **2** 020303

[53] Cimini V, Valeri M, Piacentini S, Ceccarelli F, Corrielli G, Osellame R, Spagnolo N and Sciarrino F 2024 *npj Quantum Inf.* **10** 26

[54] Meyer J J, Borregaard J and Eisert J 2021 *npj Quantum Inf.* **7** 89

[55] Palittapongarnpim P and Sanders B C 2019 *Phys. Rev.* A **100** 012106

[56] Morelli S, Usui A, Agudelo E and Friis N 2021 *Quantum Sci. Technol.* **6** 025018