Supplementary information for Optimizing quantum-enhanced Bayesian multiparameter estimation in noisy apparata

Federico Belliardo,[1] Valeria Cimini,[2] Emanuele Polino,[2] Francesco Hoch,[2] Bruno Piccirillo,[3] Nicolò Spagnolo,[2] Vittorio Giovannetti,[1, *] and Fabio Sciarrino[2, †]

[1]*NEST, Scuola Normale Superiore and Istituto Nanoscienze-CNR, I-56126 Pisa, Italy*
[2]*Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 5, I-00185 Roma, Italy*
[3]*Department of Physics "E. Pancini", Universitá di Napoli "Federico II",*
Complesso Universitario MSA, via Cintia, 80126, Napoli

Supplementary Note 1.   Derivation of multiparameter precision bounds

In this section, we derive the multiparameter precision bound of Eq. (2) in the main text. In order to do so, we start from the quantity that we simulate, that are the measurements outcomes. The greedy algorithm selects for each photon consumed in the experiment the best value of s among the available ones and the polarization basis b that give us the maximum information gain. We called $\boldsymbol{r}_N$ the list of tuples containing these choices, together with the outcomes o i.e.

$$\boldsymbol{r}_N = \{(s_1, b_1, o_1), (s_2, b_2, o_2), \cdots, (s_K, b_K, o_K)\} , \tag{1}$$

Notice that the number of total resources used N and the number of measurements K, i.e. the number of photons are different. The most widely employed figure of merit for the precision of an estimator is the Mean Squared Error (MSE). Given the estimators $\hat{\theta}^{(\boldsymbol{r}_N)}$ and $\hat{V}_{s_i}^{(\boldsymbol{r}_N)}$ for θ and V_{s_i} respectively, and a weight matrix G that codifies which parameters are of interests, we have introduced in the main text the error quantity

$$\Delta_{\boldsymbol{r}_N,G}^2(\theta) := G_{1,1}|\hat{\theta}^{(\mathbf{r}_N)} - \theta|^2 + \sum_{i=1}^4 G_{i+1,i+1}|\hat{V}_{s_i}^{(\mathbf{r}_N)} - V_{s_i}|^2 , \tag{2}$$

its expectation value on the experimental run is $\Delta_G^2(\theta) := \mathbb{E}_{\boldsymbol{r}_N}[\Delta_{\boldsymbol{r}_N,G}^2(\theta)]$. We then take the expectation value of this precision on the prior distribution for θ and define $\Delta_G^2 := \mathbb{E}_\theta[\Delta_G^2(\theta)]$. We can approximate these expressions by means of M simulations for a discrete and finite number of angles J. So that we can write

$$\Delta_G^2 \simeq \frac{1}{MJ} \sum_{m=1}^M \sum_{j=1}^J \Delta_{\boldsymbol{r}_N^{m,j},G}^2(\theta_j) , \tag{3}$$

where $\boldsymbol{r}_N^{m,j}$ is the string characterizing the m-th experimental run for the j-th angle. The figure of merit in Eq. (1) of the main text is computed by taking the median of the M quantities $\sum_{j=1}^J \Delta_{\boldsymbol{r}_N^{m,j},G}^2(\theta_j)$ instead of the mean. We now see how the Cramèr-Rao (CR) bound sets a limit to Δ_G^2 and how this can become a bound for the median error $\mathcal{M}_G^2$. Depending on the measurement basis chosen, the results $o_1, o_2 \in \{-1, +1\}$ of the two polarization measurements for the $i-$th q-plate are distributed respectively according to

$$p_1(o_1|\theta, V_{s_i}) := \frac{1}{2}\left(1 + o_1 \cdot V_{s_i} \cos 2s_i\theta\right) \quad \text{or} \quad p_2(o_2|\theta, V_{s_i}) := \frac{1}{2}\left(1 + o_2 \cdot V_{s_i} \sin 2s_i\theta\right) . \tag{4}$$

We have ν_i measurements for the q-plate s_i in total, that we assume being evenly split between the two polarization basis. This is not true in the experiment, since the basis is chosen adaptively, however it is an assumption necessary to proceed in the analytical computation of the bound. From these probabilities we can write the 5×5 Fisher information

---

* vittorio.giovannetti@sns.it
† fabio.sciarrino@uniroma1.it

matrix I (FI matrix) for the five parameters $(\theta, V_{s_1}, V_{s_2}, V_{s_3}, V_{s_4})$, whose non-zero elements are

$$I_{11} = \sum_{i=1}^{4} \frac{4s_i^2 V_{s_i}^2 \nu_i \left(-4 + 3V_{s_i}^2 + V_{s_i}^2 \cos 4s_i\theta\right)}{-8 + 8V_{s_i}^2 - V_{s_i}^4 + V_{s_i}^4 \cos 4s_i\theta} \; ,$$

$$I_{i+1,1} = -\frac{4s_i V_{s_i}^3 \nu_i \cot 2s_i\theta}{(V_{s_i}^2 - \csc^2 s_i\theta)(V_{s_i}^2 - \sec^2 s_i\theta)} \; ,$$

$$I_{i+1,i+1} = 2\nu_i \left(\frac{1}{-V_{s_i}^2 + \csc^2 s_i\theta} + \frac{1}{-V_{s_i}^2 + \sec^2 s_i\theta}\right) \; ,$$

for $i = 1, 2, 3, 4$, and $I_{i+1,1} = I_{1,i+1}$ for symmetry. The Cramér-Rao bound, holding true for asymptotically unbiased estimators, is then expressed by the following inequality involving $\Delta_G^2(\theta)$:

$$\Delta_G^2(\theta) \geq \mathrm{Tr}\left(G \cdot I^{-1}\right) \; , \tag{5}$$

by taking the expectation value on the prior on θ we have

$$\Delta_G^2 = \mathbb{E}_\theta[\Delta_G^2(\theta)] \geq \mathbb{E}_\theta[\mathrm{Tr}\left(G \cdot I^{-1}\right)] \geq \mathrm{Tr}\left(G \cdot \mathbb{E}_\theta[I]^{-1}\right) \; . \tag{6}$$

We now want to renormalize the uses of each q-plate ν_i in such a way to highlight the dependence on the total number of resources N, i.e. $\nu_i = x_i N$. The FI matrix I becomes $I = N\widetilde{I}$, where the entries of $\widetilde{I}$ are similar to that of I, only that ν_i is substituted with x_i. The CR bound reads now

$$\Delta_G^2 \geq \frac{\mathrm{Tr}\left(G \cdot \mathbb{E}_\theta[\widetilde{I}]^{-1}\right)}{N} \geq \frac{C_G}{N} \; , \tag{7}$$

where the expectation value of the matrix $\widetilde{I}$ is diagonal with entries

$$\mathbb{E}_\theta[\widetilde{I}_{11}] = 4\sum_{i=1}^{4} x_i s_i^2 \left(1 - \sqrt{1 - V_{s_i}^2}\right) \; , \tag{8}$$

$$\mathbb{E}_\theta[\widetilde{I}_{i+1,i+1}] = \frac{4x_i \left(1 - \sqrt{1 - V_{s_i}^2}\right)}{V_{s_i}^2 \sqrt{1 - V_{s_i}^2}} \; , \quad \text{for } i = 1, \ldots, 4 \; , \tag{9}$$

and C_G is the solution of the following minimization problem:

$$\begin{cases} C_G = \min_{x_i} \mathrm{Tr}\left(G \cdot \mathbb{E}_\theta[\widetilde{I}]^{-1}\right) \\ \text{subject to } \sum_{i=1}^{4} s_i x_i = 1 \\ x_i \geq 0 \end{cases} \; . \tag{10}$$

In order to get a reference value for the median error, we suppose that the estimators $\hat{\theta}^{(\boldsymbol{r}_N)}$ and $\hat{V}_{s_i}^{(\boldsymbol{r}_N)}$ are asymptotically normal and unbiased. Because the square of the deviations $|\hat{\theta}^{(\boldsymbol{r}_N)} - \theta_j|$ and $|\hat{V}_{s_i}^{(\boldsymbol{r}_N)} - V_{s_i}|^2$ are left-skewed and independent variables, we observe that

$$\mathrm{Median}\left[\sum_{j=1}^{J} \Delta_{\boldsymbol{r}_N^j, G}^2(\theta_j)\right] \geq \sum_{j=1}^{J} \mathrm{Median}[|\hat{\theta}^{(\boldsymbol{r}_N)} - \theta_j|^2] + \mathrm{Median}[|\hat{V}_{s_i}^{(\boldsymbol{r}_N)} - V_{s_i}|^2] \; .$$

Under the said hypothesis, the variable $\hat{\theta}^{(\mathbf{r}_N)} - \theta_j$ is a Gaussian centered in zero and the median of its square is proportional to the variance

$$\mathrm{Median}\left[|\hat{\theta}^{(\mathbf{r}_N)} - \theta|^2\right] = \xi \mathbb{E}_{\boldsymbol{r}_N}[|\hat{\theta}^{(\mathbf{r}_N)} - \theta|^2] \; , \tag{11}$$

with a factor $\xi \simeq 0.4549$ that can be estimated numerically. Therefore, the bound on the median error of the estimation is

$$\mathcal{M}_G^2 \gtrsim \frac{\xi C_G}{N} \; . \tag{12}$$

Supplementary Note 2. The Bayesian algorithm

In this section, we present the Bayesian algorithm proposed in [?], with the application to our q-plates setup in mind. With respect to the original formulation, we made a few corrections necessary because of the circular nature of the angular variable that we are going to measure. In every experiment, $n_p = 5000$ particles have been used. The parameters to estimate are collected in the vector $\boldsymbol{x} := (\theta, V_{s_1}, V_{s_2}, V_{s_3}, V_{s_4})$, that contains the phase in the first entry and the four visibilities in the other ones. Being the Granade's method based on a particle filter, it represents internally the posterior probability distribution with the ensemble $\mathcal{E} := \{\boldsymbol{x}^k, w^k\}$, where $\boldsymbol{x}^k$ is the position of the k-th particle and w^k its weight. The j-th component of the k-th particle of the ensemble will be represented as x_j^k, and could correspond to the phase if $j = 0$, that is $x_0^k = \theta^k$ or to one of the visibilities if $j = 1, 2, 3, 4$, that is $x_j^k = V_{s_j}^k$. The mean of the angular values is computed as

$$\hat{\mu}_0 := \arg\left[\sum_{i=1}^{n_p} w^k \exp\left(\theta^k\right)\right] \,, \tag{13}$$

while the mean values of the visibilities are

$$\hat{\mu}_j = \sum_{k=1}^{n_p} w^k V_{s_j}^k \,. \tag{14}$$

Together they form the vectorial mean of the distribution $\hat{\boldsymbol{\mu}} = (\hat{\mu}_0, \hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3, \hat{\mu}_4)$. The covariance matrix is defined as

$$\text{Cov}_{ij} := \sum_{k=1}^{n_p} w^k (x_i^k - \hat{\mu}_i)(x_j^k - \hat{\mu}_j) \,. \tag{15}$$

If $i = 1$ or $j = 1$ then difference in $x_i^k - \hat{\mu}_i$ or $x_j^k - \hat{\mu}_j$ is actually the circular distance

$$d(x_i^k, \hat{\mu}_i) = \pi - |(x_i^k - \hat{\mu}_i) \mod 2\pi - \pi| \,. \tag{16}$$

The Bayesian algorithm tries for each new experiment (each new photon sent with a specific q-plate activated) to minimize the scalar variance of the posterior distribution, that is

$$\sigma^2 := \text{Tr}\left[G \cdot \text{Cov}\right] = \sum_{i,j} G_{ij}\text{Cov}_{ij} \,, \tag{17}$$

by simulating each one of the possible 8 experiments (4 q-plates and 2 polarization basis among which we have to choose) and choosing the one with the lowest expected variance. In this way the algorithm attempts to concentrate as much as possible the distribution around its mean, without however planning for more than one step in the future. The simplest possible non-greedy extension of this would be to simulate two measurement steps, this would mean computing 64 possible expected variances. Simulating many more steps becomes quickly unfeasible. The resampling strategy of the Granade procedure in [?] has also undergone minor changes to adapt it to the phase estimation problem.

Details on the data analysis

At difference with the offline algorithm when we run a simulation with a fixed number of photons we cannot exactly fix the total number N of resources, because the stochastic nature of the measurement outcomes propagates to the choice of the next q-plate and therefore to N. However, we can repeat $M \gg 1$ times the simulation, collecting the precision after each used photon in the tuple $(\Delta_{\boldsymbol{r}_n, G}^2(\theta_j), n)$, where n is the total resource number used up to that point. We will then look at all the point with total resources falling in the interval $[n, n + \Delta n]$ with Δn small, and the median error of this cluster is the error we associate to n. We call this operation clustering of the data, and we chose to cluster only the point having $n > 100$. The chosen interval was $\Delta n = 50$.

Supplementary Note 3. Values of the angles and the visibilities

The value of the rotation angle is known from the mechanical platform on which the receiving end of the apparatus is mounted and the visibilities are computed from the non-Bayesian estimator

$$\hat{V} = \sqrt{\frac{\nu\left[(2f_0 - 1)^2 + (2f_+ - 1)^2\right] - 1}{\nu - 1}} \ , \tag{18}$$

where f_0 and f_+ are the frequencies of the outcomes $o = 0$ and $o = +$ for the two polarization measurements for a fixed phase and q-plate, and ν is the number of experiments executed for each polarization. These estimators for the visibilities are evaluated prior and independently of the Bayesian procedure. The results will be considered the "true" values of the visibilities in evaluating the precision of the Bayesian approach. These are

θ	V_1	V_2	V_3	V_4
0.00235	0.8776	0.9091	0.8445	0.7038
0.06145	0.9085	0.8934	0.8007	0.7611
0.38000	0.9399	0.9153	0.7936	0.7222
0.49620	0.9211	0.9315	0.7261	0.8186
1.6645	0.9331	0.8914	0.8691	0.7312
1.8750	0.9599	0.9081	0.8762	0.6618
2.5900	0.9187	0.9587	0.8775	0.6848
2.9600	0.8986	0.9321	0.8700	0.7528
-	0.9197	0.9174	0.8322	0.7295

The above table contains the values of the eight angles analyzed in the experiment and their corresponding visibilities for each q-plate configuration. The last line reports the mean values of the visibilities.