Experimental Phase Estimation Enhanced by Machine Learning

Alessandro Lumino,[1] Emanuele Polino,[1] Adil S. Rab,[1] Giorgio Milani,[1] Nicolò Spagnolo,[1]
Nathan Wiebe,[2] and Fabio Sciarrino[1],[*]

[1]*Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro, 5, I-00185 Roma, Italy*
[2]*Quantum Architectures and Computation Group, Microsoft Research, Redmond, Washington 98052, USA*

Phase-estimation protocols provide a fundamental benchmark for the field of quantum metrology. The latter represents one of the most relevant applications of quantum theory, potentially enabling the capability of measuring unknown physical parameters with improved precision over classical strategies. Within this context, most theoretical and experimental studies have focused on determining the fundamental bounds and how to achieve them in the asymptotic regime where a large number of resources are employed. However, in most applications, it is necessary to achieve optimal precision by performing only a limited number of measurements. To this end, machine-learning techniques can be applied as a powerful optimization tool. Here, we implement experimentally single-photon adaptive phase-estimation protocols enhanced by machine learning, showing the capability of reaching optimal precision after a small number of trials. In particular, we introduce an approach for Bayesian estimation that exhibits best performance for a very low number of photons N. Furthermore, we study the resilience to noise of the tested methods, showing that the optimized Bayesian approach is very robust in the presence of imperfections. Application of this methodology can be envisaged in the more general multiparameter case, which represents a paradigmatic scenario for several tasks, including imaging or Hamiltonian learning.

I. INTRODUCTION

Quantum metrology is one of the most promising applications of quantum theory [1–5], in which the aim is to obtain enhanced performance in the estimation of unknown physical parameters by employing quantum resources. Quantum metrology holds a broad area of applications for all those problems that require estimation of one or more parameters [6]. Between some of the most relevant applications, quantum-enhanced performance can be exploited for gravitational-wave detection [7]; quantum microscopy [8] and imaging [9,10]; quantum tomography [11,12]; measurements on biological systems [13]; positioning and clock synchronization [14]; optomechanics [15]; protocols for quantum communication (e.g., discrimination of nonorthogonal states [16] or coordinate transmission [17]); or frequency estimation in spectroscopy and atomic interferometry [18,19].

A notable benchmark for quantum metrology is provided by phase estimation, a task where the parameter to be measured is an optical phase embedded within an interferometric setup. In this scenario, an input probe field is prepared in a suitable state and sent through the system. The value of the phase is retrieved by measuring

the field after the evolution in the interferometer and by repeating the procedure N times to perform statistical analysis. While the ultimate precision achievable with classical resources is known to be bounded by the standard quantum limit (SQL), stating that the achievable error on the unknown phase ϕ scales as $N^{-1/2}$ (where N is the number of photons), the adoption of quantum inputs can, in principle, improve the performance up to the Heisenberg limit (HL) [1,2], scaling as N^{-1}. Several theoretical and experimental studies [20–29] have focused on devising experimental schemes able to reach quantum-enhanced performance. Furthermore, recent advances in integrated photonics have opened up new possibilities for the implementation and development of phase-estimation protocols [30–37]. In parallel, a thorough investigation has been dedicated to identifying the effect of experimental noise and losses [38–41].

Several problems can be cast in a phase-estimation task. Moreover, in some of those scenarios, it is crucial to optimally acquire information on the unknown parameter by performing only a limited number of measurements. For instance, studies of highly photosensitive biological samples are limited by optical damage [13,42] and it is, thus, necessary to impinge on the system with low-energy probes. Other examples lie in quantum communication protocols where single-photon transmission and

measurement are employed [43,44]. Here, long-distance transmission limits the qubit rate to low values and thus optimizing the extraction of information can greatly benefit the task's success.

In the scenario where the parameter to be estimated is a single phase, it is always possible to identify the optimal measurements and a suitable estimator (the latter being essentially a data-processing strategy) to reach the maximum performance achievable with the chosen probe state [3,45,46]. However, those recipes guarantee the capability of reaching the optimal error only in the asymptotic regime, thus requiring us to repeat the estimation process a large number of times. Conversely, in most applications, it is crucial to optimally acquire information on the unknown parameter by performing only a limited number of measurements. The ultimate limits to metrology are given by the quantum Cramér-Rao bound, which yields a lower bound on the variance of any unbiased estimator of the parameter given the ability to perform the optimal POVM. Specifically, when a finite number of probes is employed, higher-order corrections to the achievable lower bounds on estimation precision have to be considered, which reduce to the quantum Cramér-Rao bound for large N [47,48]. The ordinary Cramér-Rao bound is, however, appropriate for the problem of iterative phase estimation since it is purely a classical statistical-inference problem.

Adaptive measurements can be exploited to reduce the number of measurements necessary to reach the ultimate limits permitted by quantum mechanics [49–55]. A promising approach to identify the best strategies capable of reaching optimal precision in phase-estimation protocols with a small number of trials is provided by machine learning [56,57]. The latter refers to the ability of computers to learn and improve their performance without being explicitly programmed and to rely mainly on experience acquired from data. Recently, several studies explored the possibility of merging the machine-learning domain with the quantum world [58,59]. On one side, theoretical investigations have ignited the analysis on whether quantum computing can enhance machine-learning protocols [60–63], for instance, by providing more efficient fundamental routines. On the other side, machine-learning approaches are particularly suitable to handle large amounts of data and complex optimization problems and can thus be potentially applied to improve data processing in quantum-information protocols [64], including phase estimation [65–67].

In this article, we report the experimental implementation of phase-estimation protocols enhanced by machine-learning techniques. We experimentally test an offline adaptive scheme proposed by Hentschel and Sanders in Refs. [68–70], based on a particle-swarm algorithm, able to self-learn the optimal feedback strategy to come close to saturating fundamental limits on the scaling of the uncertainty of any unbiased estimator of the phase with the

number of measurements. We then introduce an optimized version of an adaptive Bayesian approach that sequentially recalculates the feedback phase according to the knowledge acquired in the previous steps and that is tailored for Gaussian prior distributions. These approaches are compared with a previously proposed adaptive Bayesian approach [65], employed as a benchmark for the investigated techniques. Such a Bayesian method based on sequential measurements provides an optimal strategy in terms of the cost associated with making decisions, as already shown in quantum-state estimation for a binary decision scenario [71].

We implement single-photon phase-estimation experiments, showing the capability to reach optimal uncertainty in the parameter after a small number of trials. Furthermore, we also consider the robustness of these methods to the most relevant sources of experimental noise. These approaches can be extended in the general case where many parameters have to be estimated simultaneously, thus representing a benchmark for a significant class of learning scenarios.

II. ADAPTIVE PHASE-ESTIMATION PROTOCOLS

Adaptive protocols represent a general technique to perform phase-estimation experiments starting from an unknown value of the parameter. In nonadaptive estimation protocols, the user sends N times the probe state through the apparatus and finally estimates the value of ϕ after collecting the full set of data. Adaptive protocols exploit additional control on the experimental system. Besides the phase ϕ to be estimated, the user has access to a set of physical parameters (for instance, additional phase shifts) that can be adjusted during the measurement process. At each step, a single instance of the probe state is sent through the system after the set of parameters has been changed by the user according to the previous knowledge acquired on the unknown phase ϕ.

The general scheme for adaptive protocols in a two-mode Mach-Zehnder interferometer (MZI) is shown in Fig. 1(a), while the corresponding experimental apparatus is shown in Fig. 1(b). The parameter to be estimated is the phase ϕ corresponding to one of the two arms, while the additional parameter is provided by a feedback phase Φ inserted on the other arm of the interferometer. After each shot of the experiment, the feedback phase Φ is adjusted according to the previous knowledge acquired on ϕ. The value of the feedback phase at each step can be determined by following either an offline or an online approach. In the first case, a processing unit determines the new value by following a precalculated list of rules. In the second case, the phase Φ at the subsequent stage is directly calculated step by step. In general, the value of the feedback phase Φ_k at step k can be chosen starting

from the results $\{x_m\}$ of all previous measurements ($m = 1, \ldots, k-1$). Hence, it is expressed by a function $\Phi_k = f(x_1, \ldots, x_{k-1})$ whose optimization may be in general a hard task. To this end, because of their capability to handle a large amount of data and high-dimensional systems, machine-learning techniques represent a promising tool to learn the optimal choice of the function $f(x_1, \ldots, x_{k-1})$. This tool would allow us to identify the most efficient strategies able to reach the best precision on ϕ by using a limited number of measurements.

A. Particle swarm optimization

Particle swarm optimization (PSO) is a swarm intelligence algorithm inspired by the social behavior of birds or fishes [72,73]. As birds search for food, particles search for the optimum of an objective function. Particles are represented by a point whose coordinates define a candidate solution to the optimization problem. The algorithm finds the optimal solution by trial and error in a fixed number of steps.

The authors of Refs. [68,70] proposed to employ the PSO technique as a tool for a feedback-based phase-estimation strategy. In their approach, the feedback phase is updated at each step from Φ_{k-1} to Φ_k according to the rule $\Phi_k = \Phi_{k-1} - (-1)^{x_{k-1}} \Delta\Phi_k$, where $x_{k-1} = \{0, 1\}$ is the result of the measurement at step $k-1$. The final estimate for the unknown phase ϕ is provided by the value Φ_N of the adaptive phase at the end of the process. The PSO algorithm is then exploited as an offline resource to determine the set of phase shifts $\{\Delta\Phi_k\}$ to be applied throughout the estimation process. A given sequence of phase shifts $\{\Delta\Phi_k\}$ for $k = 1, \ldots, N$ is named the policy. The objective function optimized by the PSO algorithm measures the goodness of a policy and it is called sharpness $S(\rho) = |\int_{-\pi}^{\pi} P(\theta|\rho)e^{i\theta}d\theta|$, where $P(\theta|\rho)$ represents the probability distribution of the error θ on the estimate and ρ represents a particle position, that is, a policy. This function is related to the Holevo variance $\Delta\theta_\rho^2 = S(\rho)^{-2} - 1$, tailored for cyclic variables such as angles or phases. A maximum for the sharpness corresponds to a minimum in the Holevo variance for a given position of a particle. The PSO algorithm then finds the optimal policy (for a given value of N) by maximizing the associated sharpness [see Fig. 1(c)]. More specifically, with this approach, the process is mapped to the evolution of particles in $\mathbb{R}^N$.

At each iteration, every particle compares the goodness of its position ρ_i with respect to its previous history and to a circular neighborhood comprising the particle itself; these operations end with updating the general best and the local best positions. Then, each particle moves according to the following relations paving the way for a new iteration: $\delta_i \rightarrow \beta_1\xi_1(\hat{\rho}_i - \rho_i) + \alpha_2\xi_2(\hat{\Lambda}_i - \rho_i) + \delta_i$ and $\rho_i \rightarrow \omega\delta_i + \rho_i$, where $\hat{\rho}_i$ is the particle best position; $\hat{\Lambda}_i$ represents the local best of the ith particle in its neighborhood;

and β_1 and α_2 model, respectively, the cognitive and social behavior of the swarm. The other variables in the formulas aid PSO convergence. The operator $\rightarrow$ means that, at each iteration, positions and velocities are first evaluated and then updated with the values on the right member. For further details, we refer to [70] and references therein.

Because of its structure, the PSO is independent of the initial state and allows intensive use of parallel GPU computing, which reduces computational time with respect to a sequential evaluation. Each particle can be associated to a thread, making the entire searching process in $\mathbb{R}^N$ a simultaneous rather than a sequential one. The algorithm searches for the maximum of the sharpness $S(\rho) = |\int_{-\pi}^{\pi} P(\theta|\rho)e^{i\theta}d\theta|$. Evaluating this function requires exponential computing time in the number of photons. Hence, the problem becomes rapidly intractable. To overcome this limitation, one can statistically infer the sharpness of a single particle by randomly choosing K phases, estimating each one of them, and then giving the sharpness estimate as $S(\rho) = |\sum_i^K e^{i\varsigma_k}|/K$, where $\varsigma_k = \varphi_k - \tilde{\varphi}_k$. This quantity is evaluated by the GPU, associating each phase to a block and each particle to a relative thread.

Here, we apply this approach in the scenario where N separable input photons are sent one by one into the input port a of a MZI, thus corresponding to an input state $|1\rangle_a^{\otimes N}$. For each value N, the PSO algorithm determines its own optimal policy.

B. Bayesian phase estimation

Single-photon adaptive Bayesian strategies are an online approach, where the phase Φ is updated with an online method [see Fig. 1(d)]. Bayesian protocols start from a prior probability distribution $p^{(0)}(\phi)$ that quantifies the initial knowledge on the phase ϕ. After each measured event x_k, the conditional probability distribution $p^{(k)}(\phi|\{x_l\}_1^k; \{\Phi_l\}_1^k)$ is updated according to the Bayes rule: $p^{(k)}(\phi|\{x_l\}_1^k; \{\Phi_l\}_1^k) = \mathcal{N}^{-1}p(x_k|\phi; \Phi_k)p^{(k-1)}(\phi|\{x_l\}_1^{k-1}; \{\Phi_l\}_1^{k-1})$, where $\mathcal{N}$ is a normalization constant and $p(x_k|\phi; \Phi_k)$ is the likelihood function expressing the probability of obtaining outcome x_k for a given value of ϕ. For nonadaptive strategies, the phase Φ is kept constant throughout the estimation process. The conditional distribution after N photons then contains all relevant statistical information on the phase. As for the PSO approach, we assume here no *a priori* knowledge on the value of the phase, thus corresponding to a prior distribution $p^{(0)}(\phi) = (2\pi)^{-1}$. The phase Φ is changed by the processing unit at each step k according to a specific rule. The final estimated value after N photons is obtained as the mean of the distribution $p^{(N)}(\phi|\{x_l\}_1^N; \{\Phi_l\}_1^N)$ according to $\phi_{\text{est}} = \int \phi p^{(N)}(\phi|\{x_l\}_1^N; \{\Phi_l\}_1^N)d\phi$. A benchmark choice for the feedback rule is the particle guess heuristic (PGH) approach [65], where the feedback phase Φ is drawn
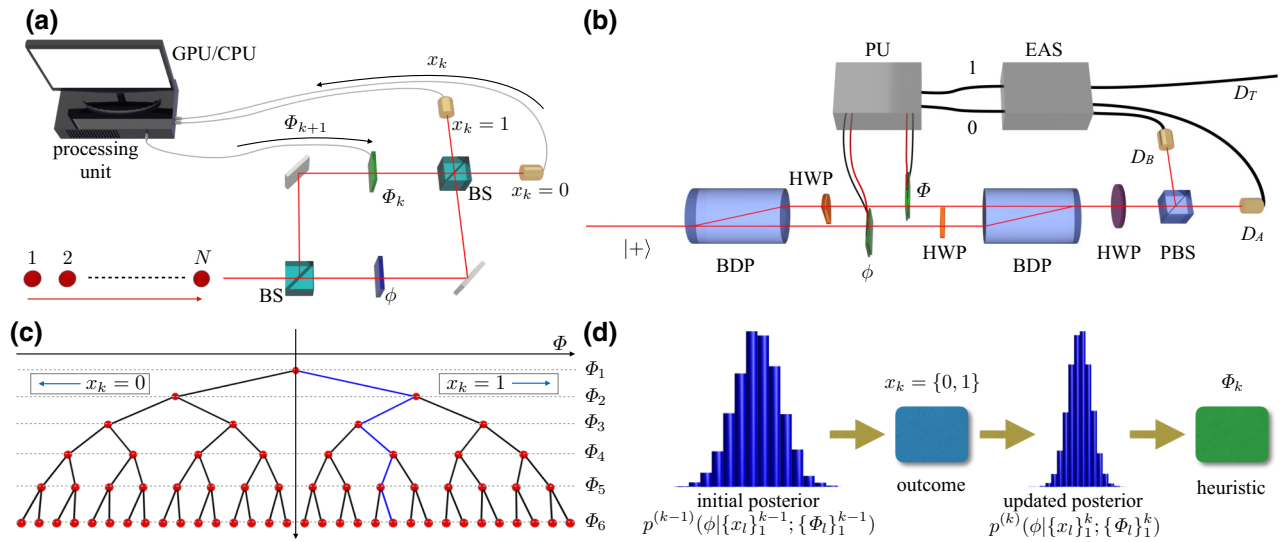
FIG. 1.   Adaptive-feedback scheme for phase estimation. (a) Conceptual scheme. A two-mode Mach-Zehnder interferometer is employed as a benchmark for adaptive phase-estimation protocols. At each step k, a single photon is injected in the interferometer and detected after the evolution. The unknown phase ϕ to be estimated (blue box) is kept constant during the entire experiment, while the phase Φ on the other arm (green box) is changed after each step. According to the results of the measurement x_k, the processing unit (a CPU or a GPU) properly sets a new feedback phase Φ_{k+1} by employing a given strategy, which can be either an offline one (thus following a predetermined set of rules) or an online one (determined by the processing unit throughout the measurement process). The process is iterated until the desired number of measurements N is reached. (b) Experimental apparatus. Experimental implementation of a two-mode MZI in an intrinsically stable configuration. A heralded single photon, triggered by the measurement of a photon in detector D_T and prepared in a $|+\rangle = 2^{-1/2}(|H\rangle + |V\rangle)$ polarization state, is injected in the interferometer. The beam-displacing prisms (BDPs) act as beam splitters by spatially separating and recombining the optical paths depending on the photon polarization state ($|V\rangle$ propagates in the same mode, while $|H\rangle$ is displaced). Phase shifts ϕ and Φ are introduced in the polarization degree of freedom by means of two liquid-crystal (LC) devices driven by an external processing unit (PU), one for each spatial mode. Two half-wave plates (HWPs) at $45°$ are employed within the interferometer to properly recombine the spatial modes. Finally, a HWP at $22.5°$ and a polarizing beam splitter (PBS) separate the two output modes, which are sent to the detection stage [avalanche photodiodes D_A and D_B and an electronic acquisition system (EAS)]. (c) Decision tree for PSO-based estimation strategy. Schematics of the binary tree at the basis of the PSO-based adaptive estimation strategy. As shown in the main text, the phase Φ is updated at each step on the basis of the last measurement result x_k according to $\Phi_k = \Phi_{k-1} - (-1)^{x_{k-1}}\Delta\Phi_k$. Each red point in the binary tree represents a possible feedback action. The blue branch stands for one among the 2^N possible paths. (d) Bayesian phase-estimation protocols. In Bayesian protocols, at each step k, the posterior distribution is updated according to outcome x_k and then the new feedback phase Φ_k is determined from the posterior according to a chosen heuristic.

randomly at each step from the posterior distribution $\Phi \sim p^{(k)}(\phi|\{x_l\}_1^k; \{\Phi_l\}_1^k)$. While this method has been proposed for Heisenberg limited metrology, it does not necessarily follow that such an approach will be appropriate in this setting where the SQL is the ultimate limit.

C. Optimized Gaussian phase estimation

In order to obtain an optimized Bayesian protocol in this regime, we propose below an approach to phase estimation, named Gaussian optimal (GO), that provides the optimal Φ under the assumptions (a) that the prior $P(\phi)$ is Gaussian with mean μ and variance σ^2, (b) that the variance of the posterior distribution obeys $\sigma \ll 0.921$ and has negligible support over the branch cut, (c) the posterior mean is used as an estimator for the phase, and (d) the experimentalist wishes to minimize the expected square error in the estimate.

Under the assumption that $\sigma \ll 0.921$ and that we have negligible support over the branch cut, we can write the given prior mean μ and variance σ^2:

$$P(\phi) \approx \frac{e^{-(\phi-\mu)^2/2\sigma^2}}{\sqrt{2\pi}\sigma}. \tag{1}$$

Now assume that we perform an experiment and measure $x = 0$ for the feedback phase Φ; then, the posterior distribution is

$$P(\phi|0;\Phi) \approx \frac{\cos^2([\Phi-\phi]/2)e^{-(\phi-\mu)^2/2\sigma^2}}{\int_{-\infty}^{\infty} \cos^2([\Phi-\phi]/2)e^{-(\phi-\mu)^2/2\sigma^2}d\phi}. \tag{2}$$

The expected square error can then be found by computing the variance by integrating over the posterior distribution.
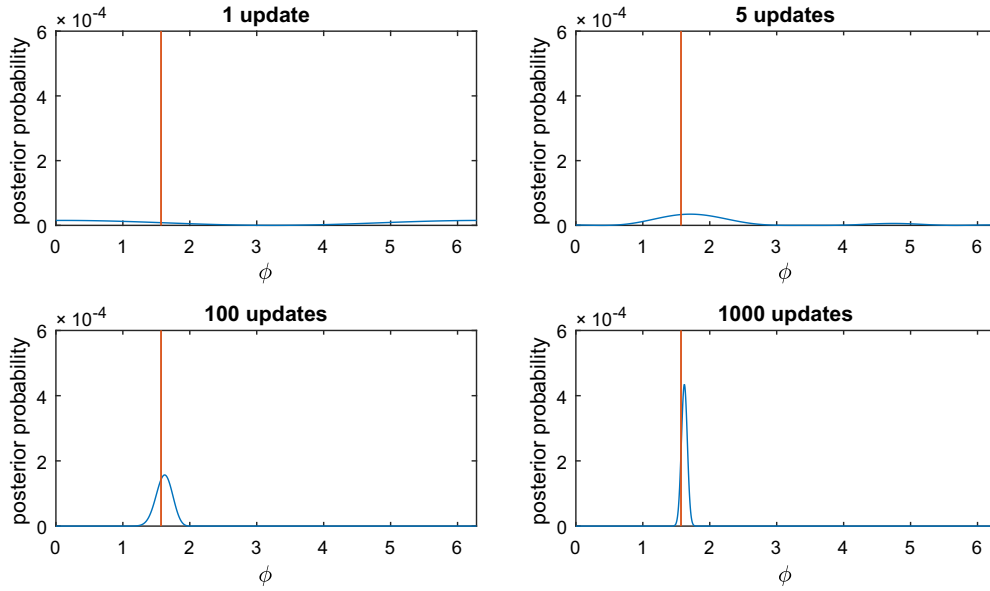
FIG. 2. Numerical verification of Gaussianity assumption. Plot of posterior probability distribution using a 2^{17}-point discretization of the prior probability distribution as a function of the number of measurements using the real component of the optimal Φ for a narrow Gaussian distribution under the assumption of a uniform initial prior. The red line denotes the true phase, which is $\phi = \pi/2$.

Under these approximations, this expression reads

$$\mathbb{V}(\phi|0; \Phi) = \int_{-\infty}^{\infty} \phi^2 P(\phi|0; \Phi)d\phi$$
$$- \left(\int_{-\infty}^{\infty} \phi P(\phi|0; \Phi)d\phi \right)^2. \quad (3)$$

Because the variance is additive, we can follow the same argument and find that the expected posterior variance that we would observe given a feedback phase Φ is chosen as

$$\int_{-\infty}^{\infty} \big\{ \cos^2([\phi - \Phi]/2)\mathbb{V}(\phi|0; \Phi)+$$
$$\sin^2([\phi - \Phi]/2)\mathbb{V}(\phi|1; \Phi) \big\} P(\phi)d\phi, \quad (4)$$

where $\mathbb{V}(\phi|1; \Phi)$ can be found by substituting $\cos^2(\cdot) \rightarrow \sin^2(\cdot)$ in Eqs. (2) and (3).

In order to find the minimum variance, we then simply have to differentiate Eq. (4) with respect to Φ and set the result equal to zero. Fortunately, analytic solutions can be found for this condition. The two corresponding to minima are

$$\Phi = \mu \pm \left[-\pi + \cos^{-1} \left(\frac{e^{\sigma^2/2}(\sigma^2 + \sqrt{(\sigma^2 - 2)(\sigma^2 - 4)} - 2)}{\sigma^2 - 2} \right) \right]. \quad (5)$$

While the precise range of σ for which Φ is real does not have an analytic form, we find numerically that when $\sigma \lesssim 0.921$, then these solutions are real. $\Phi = \mu$ is also a solution, but it yields a maximum, not a minimum, of the posterior variance. An interesting feature of this rule is that, for small σ, the feedback phase obeys $\Phi = \mu \pm [-\pi/2 + \sin^{-1}(\sqrt{2} - 1) + O(\sigma^2)]$, which corresponds to

the phase moving left or right by a constant amount based on the measurement outcome. If the particle guess heuristic (PGH) is used, then we obtain random feedback phases such that $|\Phi - \mu| \sim \sigma$, which is not optimal for minimizing the quadratic loss under the assumption of a Gaussian prior with $\sigma \ll 1$ because the optimal Φ does not converge to μ as $\sigma \rightarrow 0$. Since the branch cut can always be chosen at any point in the phase, the main assumption that needs to be validated is that of the posterior being well modeled by a narrow Gaussian distribution. We provide numerical evidence for this assumption in Fig. 2.

We then perform numerical simulations to address the performance of the method. We observe that, after a few measurements $N \sim 15$–20, both the quadratic loss ($\phi_{\text{est}} - \phi_{\text{true}}$)^2 [Fig. 3(a)] and square root of Holevo variance σ_{est} [Fig. 3(b)] approach the standard quantum limit $N^{-1/2}$. Such scaling is maintained for significantly larger values of N, where assumption (a) is fulfilled. Furthermore, the performance of the algorithm is independent of the true phase ϕ_{true} [Figs. 3(c) and 3(d)], thus showing that the applied algorithm reaches optimal and unbiased performance in the full $[0, 2\pi]$ interval.

III. EXPERIMENTAL SINGLE-PHOTON PHASE ESTIMATION

We then perform adaptive single-photon phase-estimation experiments by employing the different techniques discussed above. The MZI is implemented by exploiting an intrinsically stable configuration as shown in Fig. 1(b). Single-photon states are generated by means of a type-II spontaneous parametric down-conversion process occurring in a 2-mm-long beta-barium borate (BBO) crystal, pumped by a 392.5-nm, 180-fs-long pulsed beam at a 76-MHz repetition rate. The source generates photon
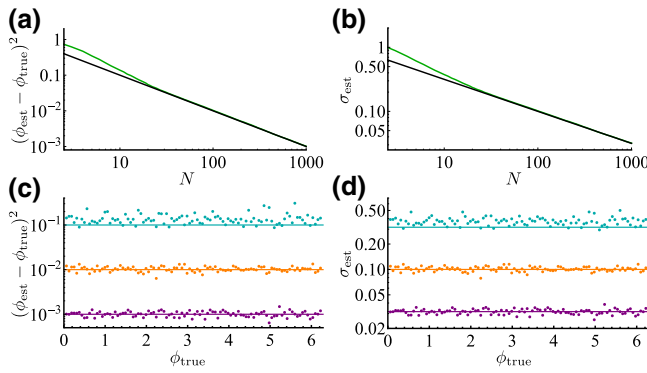
FIG. 3. Numerical simulations with the GO approach in the noiseless regime. $M_{est} = 100$ independent simulated estimation processes are performed for $M_{phases} = 100$ phases distributed in the full $[0, 2\pi]$ interval. (a) Average quadratic loss $(\phi_{est} - \phi_{true})^2$ and (b) square root of Holevo variance σ_{est} as a function of N. In (a),(b), the green lines correspond to numerically simulated data, while the black lines are the standard quantum limit $N^{-1/2}$. (c) Quadratic loss and (d) square root of Holevo variance as a function of the true phase ϕ_{true}, for different values of N. Cyan denotes $N = 10$, orange denotes $N = 100$, and purple denotes $N = 1000$. Solid lines correspond to the standard quantum limit for each value of N.

pairs with orthogonal polarizations at 785 nm. Single-photon inputs are obtained by exploiting the source in a heralded configuration, thus directly detecting one of the two generated photons that acts as a trigger.

A. Acquisition system for adaptive phase estimation

Phases within the interferometer are tuned via liquid-crystal devices that change the relative phase between the horizontal and vertical polarization according to the applied electric voltage. An external processing unit drives two liquid-crystal devices that introduce the phase shifts ϕ (to be estimated) and Φ (the adaptive one) and it is connected to an electronic acquisition system. The latter is connected to the single-photon detectors analyzing the output of the interferometer. This system records a single event and sends the result of the measurement ($x_m = \{0, 1\}$) to the processing unit, which changes the value of Φ according to the chosen rule. More specifically, the electronic acquisition system is a home-built device with two output channels that processes internally the output detected signal. When a single event is recorded, the EAS disables the acquisition process and sends the outcome of the measurement to the processing unit, consisting of a conventional desktop computer. The EAS takes as input the trigger detector D_T and the two detectors placed at the output of the interferometer D_A and D_B. The outcome $x_{k-1} = 0$ ($x_{k-1} = 1$) is obtained when a coincidence D_T/D_A (D_T/D_B) is recorded. A LABVIEW routine then processes the input signal to determine the value of the feedback phase for the next step. In the
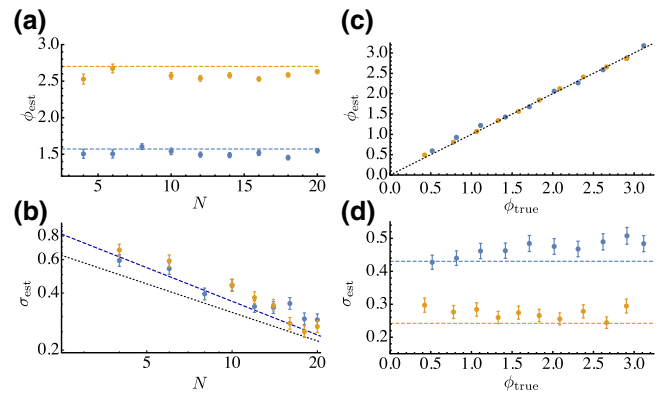

FIG. 4. Single-photon phase-estimation measurements with the PSO approach. Circular mean of estimated values ϕ_{est} and square root of the Holevo variance σ_{est} are obtained from $M = 100$ independent experiments. (a),(b) Measurements for two different phases as a function of N (blue points denote $\phi_{true} = 1.5714$; orange points denote $\phi_{true} = 2.7015$). (c),(d) Measurements as a function of ϕ_{true} (blue points denote $N = 7$; orange points denote $N = 20$). In (a),(c), bars are the errors associated to the circular mean of estimated values, evaluated as $\sigma_{est}M^{-1/2}$, while dashed lines are the true value of the phases. In (b),(d), bars are the errors associated to σ_{est}, evaluated as $\sigma_{est}[2(M-1)]^{-1/2}$, while dashed lines are the theoretical predictions obtained from numerical simulations. In (b), the black dotted line corresponds to the standard quantum limit.

PSO approach, the routine loads the (precalculated) policy $\{\Delta\Phi_k\}$ and then evaluates the new value as $\Phi_k = \Phi_{k-1} - (-1)^{x_{k-1}}\Delta\Phi_k$. A C program is internally loaded to convert the phase value to the corresponding voltage for the liquid-crystal device. For the Bayesian approach, an internally loaded C program updates the conditional distribution $p^{(k-1)}(\phi|\{x_l\}_1^{k-1}; \{\Phi_l\}_1^{k-1})$ and calculates the new value Φ_k as previously discussed. In both cases, the phase is then set via the LABVIEW routine within the interferometer by changing the applied voltage to the liquid-crystal device. This acquisition system permits us to perform an actual N-photon experiment, where the estimate of ϕ is obtained by using N single-photon events and no additional average is performed.

B. Results

The results obtained with the PSO approach are shown in Fig. 4. The performance of the algorithm is verified as a function of the number of photons N and for different values of the phases. We perform $M = 100$ independent N-photon experiments for each tested value of the phase (ϕ_{true}). Indeed, while our experiments yield an estimate of ϕ after using only N photons, it is necessary to repeat the process M times to evaluate the statistical error associated to the protocol (similarly to other approaches such as maximum likelihood). We observe that the experimental results are in good agreement with the predictions obtained from
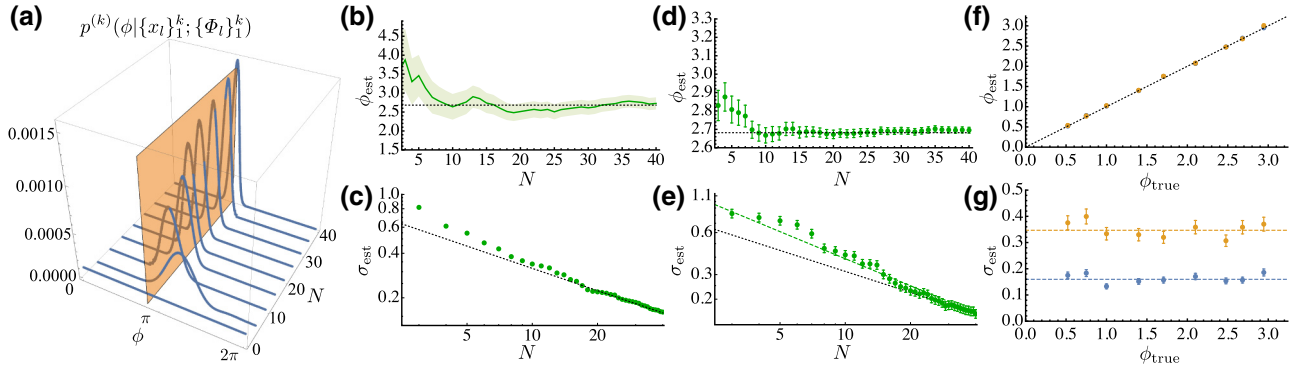
FIG. 5. Single-photon phase-estimation experiment with the Bayesian GO approach. (a)–(c) Single-photon phase-estimation experiment for $\phi_{\text{true}} = 2.6819$, where a sample of $N = 40$ single photons are employed to obtain a final estimate. (a) Evolution of the posterior distribution $p^{(k)}(\phi|\{x_l\}_1^k; \{\Phi_l\}_1^k)$ during the estimation experiment. The orange plane corresponds to value ϕ_{true}. (b) Evolution of the estimated value ϕ_{est} (green solid line) with corresponding confidence interval given by the Holevo variance of the posterior distribution (shaded region). (c) Evolution of the square root of the Holevo variance σ_{est} (green solid line) during the phase estimation experiment, compared with the standard quantum limit $N^{-1/2}$ (black dotted line). (d),(e) Circular mean of estimated value ϕ_{est} and square root of the Holevo variance σ_{est} obtained from $M = 100$ independent experiments for $\phi_{\text{true}} = 2.6819$. In (e), the green dashed line is the prediction obtained from numerical simulations, while the black dotted line corresponds to the standard quantum limit. (f),(g) Measurements as a function of ϕ_{true} (blue points denote $N = 40$; orange points denote $N = 12$). In (d),(f), the error bars are the errors associated to the circular mean of estimated values, evaluated as $\sigma_{\text{est}}M^{-1/2}$, while the dashed lines are the true value of the phases. In (e),(g), the error bars are the error associated to σ_{est}, evaluated as $\sigma_{\text{est}}[2(M-1)]^{-1/2}$.

numerical simulations, thus showing the capability of the PSO approach to perform phase-estimation experiments by using only a limited number of photons N.

We then perform single-photon phase-estimation experiments by employing the Bayesian GO technique. The results are shown in Fig. 5. First, we consider a single-phase estimation experiment obtained by sequentially sending $N = 40$ single photons through the interferometer. One of the main advantages of Bayesian techniques is the capability to provide in a single experiment an estimate of the unknown parameter and a confidence interval for the process. This information is encoded in the posterior distributions, as shown in Figs. 5(a)–5(c). We observe that the algorithm converges to the true value after a few runs ($N \sim 10$) and that the ultimate limit in this scenario, provided by the standard quantum limit, is reached after a limited number of measurements ($N \sim 15$). We then perform $M = 100$ independent phase estimations with $N = 40$ photons for a fixed value of $\phi_{\text{true}} = 2.6819$, showing that the protocol is stable over repeated experimental series [see Figs. 5(d) and 5(e)] and the standard quantum limit can be effectively reached. Finally, the GO method is applied for different phases, showing that this approach provides a similar performance independently of ϕ_{true} [see Figs. 5(f) and 5(g)].

Finally, in Fig. 6, we compare the performances of the PSO and the Bayesian GO approaches (see also Table I). As a benchmark, we consider the recently proposed PGH method as described above. To perform a fair comparison, Bayesian approaches are compared to the PSO one by performing the same statistical analysis on $M = 100$

independent experiments. Indeed, while in the Bayesian case it is possible to associate an estimation error to a single experiment from the distribution $p^{(N)}(\phi|\{x_l\}_1^N; \{\Phi_l\}_1^N)$, we perform a comparison of the two techniques by using the same analysis tools. Our test is also slightly biased in favor of PSO because PSO uses the Holevo variance as its loss function, whereas the Bayesian methods are designed to minimize the quadratic loss function which need not coincide with the Holevo variance for the values of σ tested. We observe that both the PSO approach and the Bayesian GO one outperform the PGH method [see Figs. 6(a)–6(d)]. Finally, the Bayesian GO approach slightly outperforms the PSO one [see Fig. 6(e)].

IV. ROBUSTNESS TO NOISE

After assessing the performances of the different algorithms, we will now examine the robustness to noise of those approaches. We consider the two most relevant sources of noise in experimental interferometric implementations, namely, depolarizing and phase noise. More specifically, depolarizing noise is due to the presence of dark counts and nonunitary visibility. Conversely, phase noise corresponds to random errors on the feedback phase, which can be due to the physical device effectively introducing the phase shift or to phase fluctuations between the two interferometer arms. We perform phase-estimation experiments by adding artificially those two sources of noise. Depolarizing noise can be introduced by adding a single noise parameter p. At each step, with probability $1 - p$, the actual click x_k of the experiment is employed,
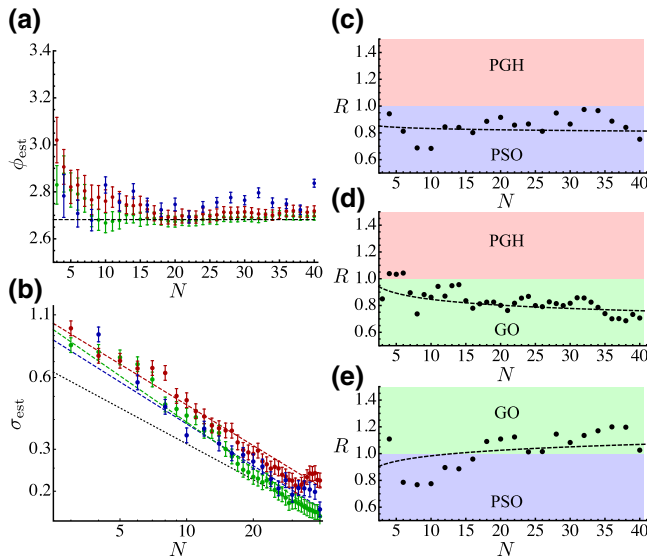
FIG. 6.   Comparison between the different techniques. (a),(b) Circular mean of estimated value ϕ_{est} and square root of the Holevo variance σ_{est} obtained from $M = 100$ independent experiments for $\phi_{\text{true}} = 2.6819$ with the different techniques. Red denotes the Bayesian PGH approach, blue denotes the PSO approach, and green denotes Bayesian GO approach. In (a), the error bars are the errors associated to the circular mean of estimated values, evaluated as $\sigma_{\text{est}} M^{-1/2}$, while the dashed lines are the true values of the phases. In (b), the error bars are the errors associated to σ_{est}, evaluated as $\sigma_{\text{est}}[2(M-1)]^{-1/2}$. (c)–(e) Ratios between the square root of the Holevo variance for the different methods: (c) PSO and PGH, (d) GO and PGH, (e) PSO and GO. The colored region highlights which technique shows improved performance in the comparison. The dashed lines are theoretical predictions obtained from numerical simulations. The black dotted line in (b) corresponds to the standard quantum limit.

while with probability p, the click $x_k = 0, 1$ is randomly drawn. Phase noise can be mimicked by adding to the feedback phase Φ_k a random shift $\delta\Phi$, normally distributed

with zero mean and variance determined by an additional parameter κ^2 representing the noise strength. In both cases, the protocol is not adapted to the presence of noise and, thus, neither the policies (for the PSO method) nor the likelihood function and the heuristic (for the Bayesian approaches) are modified with respect to the ideal case.

A. Cramér-Rao bound in the presence of noise

The ultimate precision achievable with single-photon probes is given by the standard quantum limit, stating that the error in the estimate scales as $\delta\phi \geq N^{-1/2}$, where N is the number of employed probes. In the presence of noise, such a limit cannot be further achieved and has to be modified to take the action of noise into account. In the general single-parameter scenario, given a specific choice of probe, evolution, and measurement, the ultimate achievable precision is provided by the Cramér-Rao bound $\delta\phi \geq [N\mathcal{I}(\phi)]^{-1/2}$, where $\mathcal{I}(\phi) = \mathbb{E}_x[\partial \ln p(x|\phi)/\partial\phi]^2$ is the Fisher information associated to the output probability distribution $p(x|\phi; \Phi)$ at the measurement stage.

We now evaluate the Fisher information for both noise models considered here. As previously discussed, depolarizing noise can be modeled by a parameter p corresponding to the probability of a random click. Such modelization corresponds to having output probability distributions of the form $p_{\text{dep}}(x|\phi; \Phi) = (1-p)p(x|\phi; \Phi) + p/2$, where $x = 0, 1$ corresponds to the two possible outcomes and $p(x|\phi; \Phi)$ is the output probability in the noiseless case, with $p(0|\phi; \Phi) = \cos^2[(\phi - \Phi)/2]$ and $p(1|\phi; \Phi) = \sin^2[(\phi - \Phi)/2]$. By directly applying the definition and by maximizing over (ϕ, Φ), the Fisher information reads $\max \mathcal{I}_{\text{dep}} = (1-p)^2$ and, thus, the minimum achievable precision is modified to $\delta\phi_{\text{dep}} \geq [N(1-p)^2]^{-1/2}$. Phase noise can be modeled by inserting a random phase shift $\delta\Phi$ in the reference arm, normally distributed with mean $\gamma = 0$ and standard deviation κ. The output probability distribution is obtained by averaging the density matrix

TABLE I.   Schematic comparison between the PSO heuristic and discretized Bayesian phase-estimation methods. The PSO-based estimation is characterized by the offline optimization of the best policies, with risk function given by the Holevo variance. The discretized Bayesian approaches are implemented online during the experiments where, at each step, the posterior distribution is updated. The mean-square error represents the corresponding risk function. These methods require different computational resources in N (number of probes). The determination of the optimal policy for the PSO scales as $O(N^6)$ (when an approximation in the Holevo variance is employed). In the discretized Bayesian cases, the number of operations necessary for each step is constant and, thus, an N-photon experiment requires $O(N)$ computational resources. The PSO estimation, unlike the Bayesian approach, does not permit us to evaluate the uncertainty on the estimation in a single experiment. Finally, the PSO approach loses its effectiveness for $N \gtrsim 45$–50.

|	PSO adaptive	Bayesian PGH	Bayesian GO
Type of approach	Offline	Online	Online
Procedure	Policy optimization	Posterior update	Posterior update
Feedback phase	Policy and last click	Random guess	Optimized over variance
Computational resources in N	$O(N^6)$	$O(N)$	$O(N)$
Risk function	Holevo variance	Mean-square error	Mean-square error
Uncertainty estimation in a single experiment	No	Yes	Yes
Range of operation	$N \lesssim 45$–50	No limits	No limits
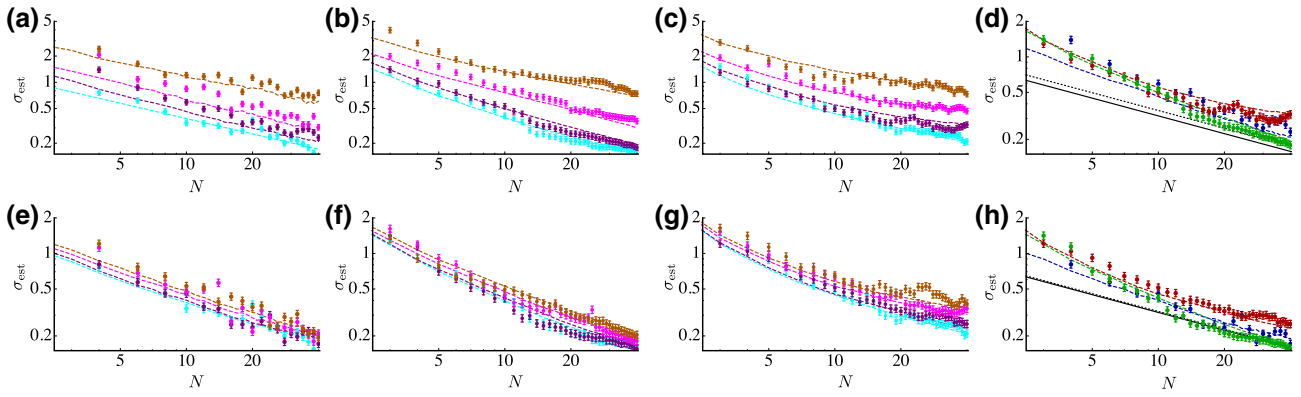
FIG. 7. Robustness to noise. Plots of the square root of the Holevo variance obtained from $M = 100$ independent experiments, estimated by the three approaches in the presence of different values of the two verified noise models. (a)–(d) Results in the presence of the depolarizing noise with the parameter p assuming values 0.1 (purple points), 0.25 (magenta points), and 0.5 (orange points) compared to the noiseless case ($p = 0$, cyan points). (a) PSO approach, (b) Bayesian GO approach, (c) Bayesian PGH approach, and (d) comparison between the three approaches for a depolarizing parameter $p = 0.1$ (blue denotes PSO, red denotes PGH, and green denotes GO). (e)–(h) Results in the presence of the phase noise added on the feedback phases, with the noise strength κ assuming values 0.2 (purple points), 0.4 (magenta points), and 0.5 (orange points) compared to the noiseless case ($\kappa = 0$, cyan points). (e) PSO approach, (f) Bayesian GO approach, (g) Bayesian PGH approach, and (h) comparison between the three approaches for a noise strength $\kappa = 0.2$ (blue denotes PSO, red denotes PGH, and green denotes GO). Dashed lines are theoretical predictions obtained from numerical simulations. In (d) and (h), the black dotted line is the optimal bound in the presence of noise, while the black solid line is the standard quantum limit in the noiseless case.

over the random phase kick $\delta\Phi$, leading to $p_{\text{pha}}(x|\phi; \Phi) = e^{-\kappa^2/2}p(x|\phi; \Phi) + (1 - e^{-\kappa^2/2})/2$. The maximum Fisher information over (ϕ, Φ) reads $\max \mathcal{I}_{\text{pha}} = e^{-\kappa^2}$ and, thus, the minimum achievable uncertainty on any unbiased estimator of ϕ (given no prior information) is modified to $\delta\phi_{\text{pha}} \geq [Ne^{-\kappa^2}]^{-1/2}$.

In both cases, if an efficient estimator exists, then the best-case scenario is that such errors only impact the variance of the optimal estimator by a constant factor. While such reduction of a constant factor is indicative that such errors can be tolerated by our phase estimation protocol, such a lower bound on the variance does not imply that the lower bound is achievable.

B. Results

The experimental results are shown in Fig. 7. For all techniques, we perform phase-estimation experiments by considering different levels of noise for both classes. For depolarizing noise, we observe that both the PSO approach and the Bayesian techniques present a good robustness to this source of error. Furthermore, the PSO and the GO techniques still maintain a better performance than the noiseless PGH method for values of the noise parameter equal to $p \lesssim 0.1$. When phase noise is considered, we observe that these methods are very robust to this error source. Furthermore, the GO approach in the noisy case still permits us to achieve a lower value of the Holevo variance σ_{est} than the noiseless PGH method even when a significant amount of phase noise $\kappa \lesssim 0.5$ is introduced.

For both noise models, we find that the GO approach is more robust than the PSO. This result is also observed experimentally during the measurement process by noting
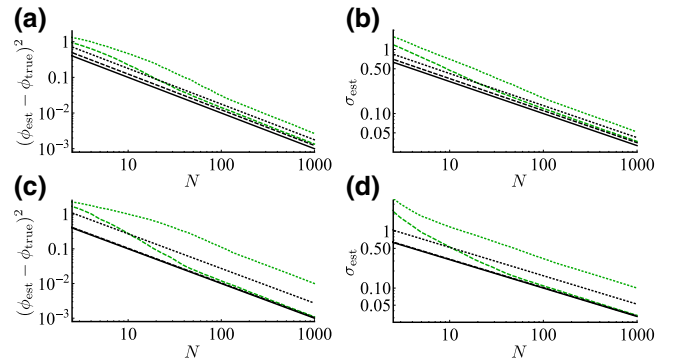

FIG. 8. Numerical simulations with the GO approach in the noisy regime. $M_{\text{est}} = 100$ independent simulated estimation processes are performed for $M_{\text{phases}} = 100$ phases distributed in the full $[0, 2\pi]$ interval. (a) Average quadratic loss $(\phi_{\text{est}} - \phi_{\text{true}})^2$ and (b) square root of Holevo variance σ_{est} as a function of N with depolarizing noise. The dashed lines denote depolarizing noise with $p = 0.1$. The dotted lines denote depolarizing noise with $p = 0.25$. (c) Average quadratic loss $(\phi_{\text{est}} - \phi_{\text{true}})^2$ and (d) square root of Holevo variance σ_{est} as a function of N with phase noise. The dashed lines denote phase noise with $\kappa = 0.2$. The dotted lines denote phase noise with $\kappa = 1$. In all plots: black solid line denotes the standard quantum limit $N^{-1/2}$, black dashed and dotted lines denote the Cramér-Rao bound, green dashed and dotted lines denote numerical simulations with the GO method.
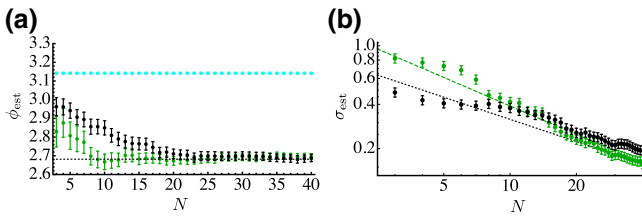
FIG. 9.   Experimental comparison between adaptive and nonadaptive strategies with uniform prior. (a) Circular mean of estimated value ϕ_{est} and (b) square root of the Holevo variance σ_{est} obtained from $M = 100$ experiments for $\phi_{true} = 2.6819$ with the different techniques. In (a), the error bars are the errors associated to the circular mean of estimated values, evaluated as $\sigma_{est}M^{-1/2}$, while the dashed lines are the true value of the phases. Errors on the circular mean in the Bayesian nonadaptive strategy are negligile since the protocol provides the same inconclusive result (π) for all N. In (b), the error bars are the error associated to σ_{est}, evaluated as $\sigma_{est}[2(M-1)]^{-1/2}$, while the dashed lines are the standard quantum limit (black) and the prediction from numerical simulation with the GO strategy (green). Green denotes the GO method. Bayesian-nonadaptive (cyan) and inversion-nonadaptive method (black) points are obtained by generating numerically random sequences from experimentally measured interference fringe data.

that the PSO method is shown to be more sensitive to misalignments of the experimental apparatus. Furthermore, for a moderate amount of noise, the GO approach allows us to reach the optimal bound with single photons when noise is taken into account ($p = 0.1$ and $\kappa = 0.2$; see also numerical simulations shown in Fig. 8).

V. COMPARISON WITH NONADAPTIVE TECHNIQUES

Throughout this article, we discuss the development and implementation of adaptive phase-estimation techniques. Adaptive protocols have to be employed in all situations where the likelihood function presents multiple values of the parameter leading to the same value of the likelihood function. Here, a twofold periodicity is present for the phase ϕ; that is, two different values of ϕ (namely, ϕ_1 and ϕ_2) lead to the same value of the likelihood. Conventional nonadaptive protocols are not able to distinguish between these two values (ϕ_1, ϕ_2), leading to an inconclusive result in the estimation process. Let us then consider a phase-estimation experiment with uniform prior in the interval $[0, 2\pi]$. We compare the GO Bayesian approach with two different nonadaptive methods by analyzing the quadratic loss ($\phi_{est} - \phi_{true})^2$ that quantifies the difference between the estimated and the true value of the phase. The first one is an inversion approach, which relies on sending N single-photon probes, counting the number of times N_0 (N_1) the outcome $x = 0$ ($x = 1$), and estimating the phase by inverting the relation $p(0|\phi; 0) \sim N_0/N$. The second approach is a Bayesian nonadaptive approach, which corresponds to keeping the feedback phase to $\Phi = 0$ when sending N independent single photons. The results are shown in Fig. 9 for $\phi = 2.6819$.

More specifically, we observe that the inversion method has a lower performance with respect to the GO method, while leading to the correct result. Conversely, the nonadaptive Bayesian protocol fails to identify the correct value of the phase even for large values of N. To obtain a more detailed analysis, we perform some numerical
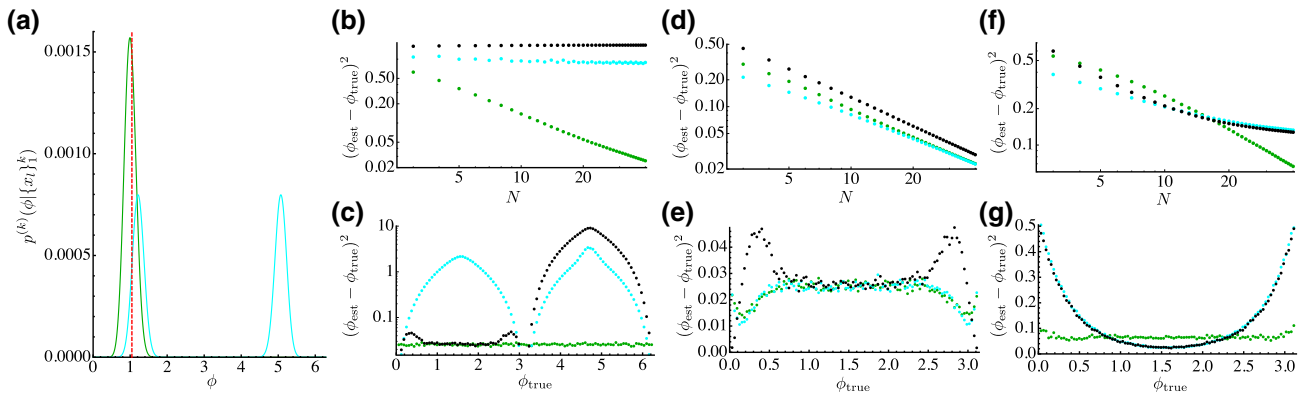

FIG. 10.   Comparison with nonadaptive strategies. (a) Evolution of the posterior distribution of a Bayesian protocol after $N = 40$ events, starting from a uniform prior in the $[0, 2\pi]$ interval, where the vertical dashed line corresponds to the true value. We observe that the nonadaptive strategy (cyan) presents a bimodal distribution, thus leading to an inconclusive result. (b),(c) Quadratic loss for the estimation process in the noiseless case with a uniform prior in $[0, 2\pi]$. (d),(e) Quadratic loss for the estimation process in the noiseless case with a uniform prior in $[0, \pi]$. (f),(g) Quadratic loss in the presence of depolarizing noise with $p = 0.25$. (b),(d),(f) show $(\phi_{est} - \phi_{true})^2$ as a function of N, averaged for 100 different phases ϕ_{true} in the range $[0, 2\pi]$ for (b) and $[0, \pi]$ for (d),(f). (c),(e),(g) show $(\phi_{est} - \phi_{true})^2$ as a function of ϕ_{true} for $N = 40$. In all plots, green denotes the GO method; cyan, the Bayesian nonadaptive method; and black, the inversion nonadaptive method.

simulations in this regime [see Figs. 10(a)–10(c)]. More specifically, in Fig. 10(a), we show the posterior distribution after $N = 40$ events for the Bayesian nonadaptive and the GO approaches. While the GO method results in a Gaussian posterior centered close to the true value, the nonadaptive approach presents a bimodal distribution. In this case, the distribution mean fails to provide an unbiased estimator and the estimation process fails. Note also that the inversion method does not work properly in the full $[0, 2\pi]$ interval and is able to estimate unambiguously the phase value only in $[0, \pi]$ [see Figs. 10(b)–10(c)].

Adaptive strategies lead to an advantage in the estimation process also in smaller intervals where nonadaptive strategies can unambiguously identify the measured phase. To this end, let us consider a uniform prior in $[0, \pi]$ for the single phase-estimation scenario discussed here. In this interval, it is possible to directly apply nonadaptive estimation protocols. In Bayesian approaches, restriction to the range $[0, \pi]$ is included by truncating the prior to have support only in this interval. This latter operation corresponds to the presence of *a priori* knowledge on ϕ. We have then performed some numerical simulation to verify the performance of the different methods. In the noiseless case [see Figs. 10(d) and 10(e)], we observe that the inversion method performs worse than Bayesian approaches, also showing a phase-dependent behavior. The nonadaptive Bayesian and the GO methods present similar performances when $N \geq 10$, by considering that the GO approach is optimal for Gaussian priors. However, the scenario changes significantly when noise is introduced in the estimation process [see Figs. 10(f) and 10(g) for the depolarizing case]. Indeed, we observe that nonadaptive methodologies present a phase-dependent behavior. More specifically, in the depolarizing noise case, the estimation obtained with nonadaptive techniques is biased toward $\phi = \pi/2$ due to the presence of noisy random clicks. Such phase-dependent behavior is not shown by the GO approach, which also obtains a better performance by averaging over ϕ_{true}. This analysis shows that adaptive methodologies perform significantly better than nonadaptive ones.

VI. CONCLUSIONS AND PERSPECTIVES

Quantum-metrology protocols represent one of the most promising applications of quantum theory to improve the sensitivity in measuring unknown parameters. Phase estimation within interferometric setups can be employed as a benchmark to develop suitable techniques to be applied in the general scenario. In this article, we show that machine-learning techniques can provide a powerful tool to design optimized adaptive protocols for this purpose. We implement single-photon phase-estimation

experiments by employing different machine-learning-based techniques, showing the capability to reach almost optimal performance after a limited number of measurements. In particular, we develop a phase-estimation scheme relying on Bayesian inference, which is shown to saturate the standard quantum limit for very low N, allowing us to achieve better performance than the other investigated methods. We also show that those techniques allow for a significant robustness to the most relevant sources of experimental noise and that the optimized method allows us to reach optimal performance in this regime. Hence, this method can be successfully employed in a realistic scenario.

These results highlight that machine-learning methodologies can be applied to optimize the performance of quantum-metrology protocols. Furthermore, those approaches can be extended to protocols exploiting quantum probes [2], permitting us to reach a performance better than the standard quantum limit, and to the general multiparameter scenario [6,34,74–77], opening new perspectives for several applications.

[1] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum-enhanced measurements: Beating the standard quantum limit, Science **306,** 1330 (2004).

[2] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum Metrology, Phys. Rev. Lett. **96,** 010401 (2006).

[3] M. G. A. Paris, Quantum estimation for quantum technology, Int. J. Quantum Inf. **7,** 125 (2009).

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

[5] L. Pezzé and A. Smerzi, in *Proceedings of the International School of Physics "Enrico Fermi"*, edited by G. M. Tino and M. A. Kasevich (IOS Press, Amsterdam, 2014), p. 691.

[6] M. Szczykulska, T. Baumgratz, and A. Datta, Multiparameter quantum metrology, Adv. Phys.: X **1,** 621 (2016).

[7] R. Schnabel, N. Mavalvala, D. E. McClelland and P. K. Lam, Quantum metrology for gravitational wave astronomy, Nat. Commun. **1,** 121 (2010).

[8] U. W. Rathe and M. O. Scully, Theoretical basis for a new subnatural spectroscopy via correlation interferometry, Lett. Math. Phys. **34**, 297 (1995).

[9] M. I. Kolobov, The spatial behavior of nonclassical light, Rev. Mod. Phys. **71**, 1539 (1999).

[10] L. A. Lugiato, A. Gatti, and E. Brambilla, Quantum imaging, J. Opt. B: Quantum Semiclass. Opt. **4**, S176 (2002).

[11] K. Vogel and H Risken, Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase, Phys. Rev. A **40**, 2847(R) (1989).

[12] G. M. D'Ariano, C. Macchiavello, and M. G. A. Paris, Detection of the density matrix through optical homodyne tomography without filtered back projection, Phys. Rev. A **50**, 4298 (1994).

[13] M. A. Taylor, J. Janousek, V. Daria, J. Knittel, B. Hage, H. A. Bachor, and W. P. Bowen, Biological measurement beyond the quantum limit, Nat. Photonics **7**, 229 (2013).

[14] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum-enhanced positioning and clock synchronization, Nature **412**, 417 (2001).

[15] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, Cavity optomechanics, Rev. Mod. Phys. **86**, 1391 (2014).

[16] A. Acín, E. Bagan, M. Baig, L. Masanes, and R. Munoz-Tapia, Multiple-copy two-state discrimination with individual measurements, Phys. Rev. A **71**, 032338 (2005).

[17] A. Peres and P. F. Scudo, Transmission of a Cartesian Frame by a Quantum System, Phys. Rev. Lett. **87**, 167901 (2001).

[18] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, Optimal frequency measurements with maximally correlated states, Phys. Rev. A **54**, R4649(R) (1996).

[19] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Improvement of Frequency Standards with Quantum Entanglement, Phys. Rev. Lett. **79**, 3865 (1997).

[20] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, Super-resolving phase measurements with a multiphoton entangled state, Nature **429**, 161 (2004).

[21] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Entanglement-free Heisenberg-limited phase estimation, Nature **450**, 393 (2007).

[22] L. Pezzé and A. Smerzi, Mach-Zehnder Interferometry at the Heisenberg Limit with Coherent and Squeezed-Vacuum Light, Phys. Rev. Lett. **100**, 073601 (2008).

[23] I. Afek, O. Ambar, and Y. Silberberg, High-noon states by mixing quantum and classical light, Science **328**, 879 (2010).

[24] R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, P. Hyllus, L. Pezzé, and A. Smerzi, Useful Multiparticle Entanglement and Sub-Shot-Noise Sensitivity in Experimental Phase Estimation, Phys. Rev. Lett. **107**, 080504 (2011).

[25] F. Wolfgramm, C. Vitelli, F. A. Beduini, N. Godbout, and M. W. Mitchell, Entanglement-enhanced probing of a delicate material system, Nat. Photonics **7**, 28 (2013).

[26] Z.-E. Su, Y. Li, P. P. Rohde, H.-L. Huang, X.-L. Wang, L. Li, N.-L. Liu, J. P. Dowling, C.-Y. Lu, and J.-W. Pan, Multiphoton Interference in Quantum Fourier Transform Circuits and Applications to Quantum Metrology, Phys. Rev. Lett. **119**, 080502 (2017).

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

[28] S. Slussarenko, M. H. Weston, H. M. Chrzanowski, L. K. Shalm, V. B. Verma, S. W. Nam, and G. J. Pryde, Unconditional violation of the shot-noise limit in photonic quantum metrology, Nat. Photonics **11**, 700 (2017).

[29] K. Rudinger, S. Kimmel, D. Lobser, and P. Maunz, Experimental Demonstration of a Cheap and Accurate Phase Estimation, Phys. Rev. Lett. **118**, 190502 (2017).

[30] J. C. F. Matthews, A. Politi, D. Bonneau, and J. L. O'Brien, Heralding Two-Photon and Four-Photon Path Entanglement on a Chip, Phys. Rev. Lett. **107**, 163602 (2011).

[31] R. Kruse, L. Sansoni, S. Brauner, R. Ricken, C. S. Hamilton, I. Jex, and C. Silberhorn, Dual-path source engineering in integrated quantum optics, Phys. Rev. A **92**, 053841 (2015).

[32] J. W. Silverstone, D. Bonneau, K. Ohira, N. Suzuki, H. Yoshida, N. Iizuka, M. Ezaki, C. M. Natarajan, M. G. Tanner, R. H. Hadfield, V. Zwiller, G. D. Marshall, J. G. Rarity, J. L. O'Brien, and M. G. Thompson, On-chip quantum interference between silicon photon-pair sources, Nat. Photonics **8**, 104 (2015).

[33] Z. Chaboyer, T. Meany, L. G. Helt, M. J. Withford, and M. J. Steel, Tunable quantum interference in a 3D integrated circuit, Sci. Rep. **5**, 9601 (2015).

[34] M. A. Ciampini, N. Spagnolo, C. Vitelli, L. Pezzé, A. Smerzi, and F. Sciarrino, Quantum-enhanced multiparameter estimation in multiarm inteferometers, Sci. Rep. **6**, 28881 (2016).

[35] P. Vergyris, T. Meany, T. Lunghi, G. Sauder, J. Downes, M. J. Steel, M. J. Withford, O. Alibart, and S. Tanzilli, On-chip generation of heralded photon-number states, Sci. Rep. **6**, 35975 (2016).

[36] J. P. Olson, K. R. Motes, P. M. Birchall, N. M. Studer, M. LaBorde, T. Moulder, P. P. Rohde, and J. P. Dowling, Linear optical quantum metrology with single photons: Experimental errors, resource counting, and quantum Cramér-Rao bounds, Phys. Rev. A **96**, 013810 (2017).

[37] S. Atzeni, A. S. Rab, G. Corrielli, E. Polino, M. Valeri, P. Mataloni, N. Spagnolo, A. Crespi, F. Sciarrino, and R. Osellame, Integrated sources of entangled photons at the telecom wavelength in femtosecond-laser-written circuits, Optica **5**, 311 (2018).

[38] U. Dorner, R. Demkowicz-Dobrzanski, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley, Optimal Quantum Phase Estimation, Phys. Rev. Lett. **102**, 040403 (2009).

[39] J. Kolodynski and R. Demkowicz-Dobrzanski, Phase estimation without a priori phase knowledge in the presence of loss, Phys. Rev. A **82**, 053804 (2010).

[40] S. Knysh, V. N. Smelyanskiy, and G. A. Durkin, Scaling laws for precision in quantum interferometry and the bifurcation landscape of the optimal state, Phys. Rev. A **83**, 021804 (2011).

[41] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology, Nat. Phys. **7**, 406 (2011).

[42] M. A. Taylor and W. P. Bowen, Quantum metrology and its application in biology, Phys. Rep. **615**, 1 (2016).

[43] V. D'Ambrosio, N. Spagnolo, L. Del Re, S. Slussarenko, Y. Li, L. C. Kwek, L. Marrucci, S. P. Walborn, L. Aolita, and F. Sciarrino, Photonic polarization gears for ultra-sensitive angular measurements, Nat. Commun. **4,** 2432 (2013).

[44] S.-K. Liao, W.-Q. Cai, J. Handsteiner, B. Liu, J. Yin, L. Zhang, D. Rauch, M. Fink, J.-G. Ren, W.-Y. Liu, Y. Li, Q. Shen, Y. Cao, F.-Z. Li, J.-F. Wang, Y.-M. Huang, L. Deng, T. Xi, L. Ma, T. Hu, L. Li, N.-L. Liu, F. Koidl, P. Wang, Y.-A. Chen, X.-B. Wang, M. Steindorfer, G. Kirchner, C.-Y. Lu, R. Shu, R. Ursin, T. Scheidl, C.-Z. Peng, J.-Y. Wang, A. Zeilinger, and J.-W. Pan, Satellite-Relayed Intercontinental Quantum Network, Phys. Rev. Lett. **120,** 030501 (2018).

[45] S. L. Braunstein and C. M. Caves, Statistical Distance and the Geometry of Quantum States, Phys. Rev. Lett. **72,** 3439 (1994).

[46] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Generalized uncertainty relations: Theory, examples, and lorentz invariance, Ann. Phys. **247,** 135 (1996).

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

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

[49] H. M. Wiseman, Adaptive Phase Measurements of Optical Modes: Going Beyond the Marginal Q Distribution, Phys. Rev. Lett. **75,** 4587 (1995).

[50] D. W. Berry and H. M. Wiseman, Optimal States and Almost Optimal Adaptive Measurements for Quantum Interferometry, Phys. Rev. Lett. **85,** 5098 (2000).

[51] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi, Adaptive Homodyne Measurement of Optical Phase, Phys. Rev. Lett. **89,** 133602 (2002).

[52] R. B. Griffiths and C.-S. Niu, Semiclassical Fourier Transform for Quantum Computation, Phys. Rev. Lett. **76,** 3228 (1996).

[53] O. E. Barndorff-Nielsen and R. D. Gill, Fisher information in quantum statistics, J. Phys. A **30,** 4481 (2000).

[54] N. B. Lovett, C. Crosnier, M. Perarnau-Llobet, and B. C. Sanders, Differential Evolution for Many-Particle Adaptive Quantum Metrology, Phys. Rev. Lett. **110,** 220501 (2013).

[55] C. Bonato, M. S. Blok, H. T. Dinani, D. W. Berry, M. L. Markham, D. J. Twitchen, and R. Hanson, Optimized quantum sensing with a single electron spin using real-time adaptive measurements, Nat. Nanotechnol. **11,** 247 (2016).

[56] K. P. Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press, Boston, 2012).

[57] P. Simon, *Too Big to Ignore: The Business Case for Big Data* (Wiley, New York, 2013).

[58] M. Schuld, I. Sinayskiy, and F. Petruccione, An introduction to quantum machine learning, Contemp. Phys. **56,** 172 (2014).

[59] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, Nature **549,** 195 (2017).

[60] S. Lloyd, M. Mohseni, and P. Rebentrost, Quantum algorithms for supervised and unsupervised machine learning, arXiv:1307.0411.

[61] P. Rebentrost, M. Mohseni, and S. Lloyd, Quantum Support Vector Machine for Big Data Classification, Phys. Rev. Lett. **113,** 130503 (2014).

[62] N. Wiebe, A. Kapoor, and K. M. Svore, Quantum algorithms for nearest-neighbor methods for supervised and unsupervised learning, Quantum Inf. Comput. **15,** 0316 (2015).

[63] X.-D. Cai, D. Wu, Z.-E. Su, M.-C. Chen, X.-L. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, Entanglement-Based Machine Learning on a Quantum Computer, Phys. Rev. Lett. **114,** 110504 (2015).

[64] N. Wiebe, C. Granade, C. Ferrie, and D. G. Cory, Hamiltonian Learning and Certification Using Quantum Resources, Phys. Rev. Lett. **112,** 190501 (2014).

[65] N. Wiebe and C. Granade, Efficient Bayesian Phase Estimation, Phys. Rev. Lett. **117,** 010503 (2016).

[66] N. Wiebe, C. Granade, A. Kapoor, and K. M. Svore, Bayesian inference via rejection filtering, arXiv:1511.06458.

[67] S. Paesani, A. A. Gentile, R. Santagati, J. Wang, N. Wiebe, D. P. Tew, J. L. O'Brien, and M. G. Thompson, Experimental Bayesian Quantum Phase Estimation on a Silicon Photonic Chip, Phys. Rev. Lett. **118,** 100503 (2017).

[68] A. Hentschel and B. C. Sanders, Machine Learning for Precise Quantum Measurement, Phys. Rev. Lett. **104,** 063603 (2010).

[69] A. Hentschel and B. C. Sanders, in *Proc. 7th Int. Conf. Information Technology: New Generations, Las Vegas, 2010* (IEEE, New York, 2010), p. 506.

[70] A. Hentschel and B. C. Sanders, Efficient Algorithm for Optimizing Adaptive Quantum Metrology Processes, Phys. Rev. Lett. **107,** 233601 (2011).

[71] D. C. Brody and B. Meister, Minimum Decision Cost for Quantum Ensembles, Phys. Rev. Lett. **76,** 1 (1996).

[72] R. Eberhart and J. Kennedy, in *Proc. Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, 1995* (IEEE, New York, 1995), p. 39.

[73] A. P. Engelbrecht, *Fundamentals of Computational Swarm Intelligence* (John Wiley & Sons, England, 2006).

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

[75] M. D. Vidrighin, G. Donati, M. G. Genoni, X.-M. Jin, S. W. 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).

[76] 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,** O1LT01 (2017).

[77] L. Pezzé, 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 Estimation of Multiple Phases, Phys. Rev. Lett. **119,** 130504 (2017).