

Received: 16 April 2019 Accepted: 19 July 2019

Published online: 27 August 2019

OPEN Continuous-variable Quantum Phase Estimation based on **Machine Learning**

Tailong Xiao¹, Jingzheng Huang¹, Jianping Fan² & Guihua Zeng¹

Making use of the general physical model of the Mach-Zehnder interferometer with photon loss which is a fundamental physical issue, we investigate the continuous-variable quantum phase estimation based on machine learning approach, and an efficient recursive Bayesian estimation algorithm for Gaussian states phase estimation has been proposed. With the proposed algorithm, the performance of the phase estimation may be improved distinguishably. For example, the physical limits (i.e., the standard quantum limit and Heisenberg limit) for the phase estimation precision may be reached in more efficient ways especially in the situation of the prior information being employed, the range for the estimated phase parameter can be extended from $[0, \pi/2]$ to $[0, 2\pi]$ compared with the conventional approach, and influences of the photon losses on the output parameter estimation precision may be suppressed dramatically in terms of saturating the lossy bound. In addition, the proposed algorithm can be extended to the time-variable or multi-parameter estimation framework.

Machine learning has become a very promising approach in many fields since it has a strong ability to learn features from large data and subsequently identifies the best strategies with few numbers of trials^{1,2}. In particular, integrating machine learning into quantum physics or quantum information has become a very novel approach to exploit the quantum effects or solve the quantum information processing problems 1,3-8. The quantum phase estimation is an important issue in quantum physics and quantum information⁹⁻¹². It plays significant roles in many fields such as quantum tomography^{13–16}, quantum imaging¹⁷, quantum metrology¹⁸, quantum communication and computation 19,20. To this end, many quantum machine learning algorithms for phase estimation have been investigated theoretically and experimentally for the discrete variable quantum state, e.g., the $(|N0\rangle + |0N\rangle)/\sqrt{2}$ (NOON) state²¹, the optimal entangled state^{3,5} and the product states⁴. These works have played an important role in optimizing the phase estimation precision. However, the discrete variable quantum states are highly unstable and not convenient to be prepared and stored especially for the NOON state and the optimal entangle state²². The measurement scheme of these proposals relies on the high efficient single photon detector. Moreover, the state-of-art machine learning algorithms for quantum phase estimation such as Particle Swarm Optimzation-based algorithm⁵ and Differential Evolution-based algorithm³ present a high complexity and the availability of the algorithms is limited by the average photon numbers. These current difficulties hinder the further precision improvement of discrete variable quantum phase estimation and the wide application of it.

One class of quantum states of particular interest is the Gaussian states²³, which is a kind of continuous-variable (CV) quantum states since they are easy to produce and comparatively robust against losses. For the CV quantum phase estimation in a conventional way, numerous works have been presented. For example, a Bayesian inference-based quantum phase estimation scheme is proposed to achieve the Heisenberg limit (HL) in ref.²⁴. In ref.25, a theoretical HL scaling scheme with two-mode squeezed vacuum state (TMSV) and parity measurement are put forward. In ref.²⁶, an adaptive feedback scheme based on parity measurement and TMSV is studied, which can achieve the HL scaling in phase interval of $[0,\pi]$. In addition, with prior knowledge^{27,28} of the parameter, ref.²⁹ analyzes the Van Trees Bound (VTB) based on Van Trees inequality, which is proven to have higher precision compared with Carmér-Rao Bound (CRB)³⁰.

In this paper, we attempt to obtain a general physical model of quantum phase estimation. We take advantages of the practical benefits of CVs by proposing a machine learning algorithm in infinite-dimensional systems. Thus,

¹State Key Laboratory of Advanced Optical Communication Systems and Networks, and Center of Quantum Information Sensing and Processing, Shanghai Jiao Tong University, Shanghai, 200240, China. ²Department of Computer Science, University of North Carolina-Charlotte, Charlotte, North Carolina, 28223, USA. Correspondence and requests for materials should be addressed to G.Z. (email: qhzeng@sjtu.edu.cn)

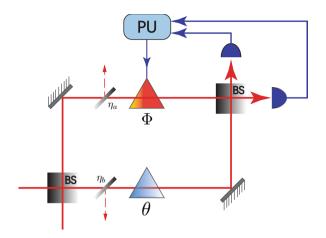


Figure 1. Mach-Zehnder interferometer schematic with an unknown phase difference θ between the two arms and an additional controllable phase shifter Φ . The measurement apparatus is chosen to be balanced homodyne detection. The 50:50 beam-splitter is fictitious with transmissivity η_a in upper arm and η_b in lower arm.

a novel CV quantum phase estimation based on machine learning has been presented. Based on the proposed algorithm, we perform the Monte Carlo simulations and some interesting performances have been obtained. Some physical problems such as the influence of the photon loss, the range limitation of the phase parameter to be estimated can be solved well with the proposed algorithm.

The remainder of this paper is organized as follows. In Sec. 2, we present a continuous-variable quantum machine learning algorithm to phase estimation based on the Mach-Zehnder interferometer and the recursive Bayesian estimation theory. In Sec. 3, we present a performance analysis based on the proposed algorithm for various input Gaussian states, and the effects of the prior information and the photon loss are investigated. The results are summarized in Sec. 4.

Machine Learning for Quantum Phase Estimation

Physical model. Interferometry is an important tool for estimating phase shifts³¹. An interferometer is a physical apparatus that encodes the value of a parameter into a probe state, also called input state. In optical interferometers, phase shifts are generally induced by a lapse in the relative time taken by the light to travel down two distinct paths (the interferometer arms)²⁵. Among the common interferometers such as Michelson-Morley interferometer, Sagna interferometer and Mach-Zehnder interferometer (MZI), they are the effective and equivalent physical models to study the phase estimation problem^{32–34}. However, MZI is a relatively representative one and generally investigated for its mathematical simplicity in describing the the physical process⁵². Therefore, without loss of generality, we employ the MZI to study the parameter model as Fig. 1 shows. In our physical model, MZI consists of two beam splitters (BS), which are assumed to be 50:50, with two arms from the first BS to the second where the difference in the optical lengths is denoted by the phase shift θ . At the output of the second BS are placed by the balanced homodyne detectors (HD) and the single HD will be used to measure the output state. The detection outcomes are transmitted to the processing unit (PU) to obtain the feedback phase Φ by executing machine learning algorithm. All the linear losses are equivalent to the photon loss in two arms, denoted by the loss rate of $\eta_a,~\eta_b.$ We regard Gaussian states as input, and their Wigner characteristic function is given by,

$$\chi(\xi) = \text{Tr}[\hat{\rho} \exp(i\xi^T \hat{R})] = \exp\left(-\frac{1}{2}\xi^T \Sigma \xi + id^T \xi\right),\tag{1}$$

where $\xi \in \mathbb{R}^4$, $\hat{R} = [\hat{x}_a, \hat{p}_a, \hat{x}_b, \hat{p}_b]^T$ is the quadrature operator, $\hat{\rho}$ is the density matrix of CV quantum state, \sum is the second order moment of quantum states $\hat{\rho}$, and d is the first-order moment. The employed \hat{x} and \hat{p} denote the position and momentum operators, respectively. We note here, $\hat{x}_{\ell} = (\hat{\ell} + \hat{\ell}^{\dagger})/\sqrt{2}$, $\hat{p}_{\ell} = (\hat{\ell} - \hat{\ell}^{\dagger})/\sqrt{2}i$, where $\ell = a, b$ denotes two modes in MZI and the $\hat{a}(\hat{b}), \hat{a}^{\dagger}(\hat{b}^{\dagger})$ are the creation and annihilation operators with respect to two modes. While the (i, j)th entry of the \sum is given by,

$$\Sigma_{ij} = \frac{1}{2} \langle \hat{R}_i, \hat{R}_j \rangle - \langle \hat{R}_i \rangle \langle \hat{R}_j \rangle, \tag{2}$$

where $\langle \cdot \rangle$ is the expectation of an operator. The *i*th term of *d* is given by,

$$d_i = \operatorname{Tr}(\hat{\rho}\hat{R}_i). \tag{3}$$

The corresponding Wigner function W(X) of pure Gaussian state is given by ³⁵

$$W(X) = \frac{\exp\left[-\frac{1}{2}(X-d)^T \Sigma^{-1}(X-d)\right]}{(2\pi)^2 \sqrt{\det(\Sigma)}},\tag{4}$$

where each of component in $X = [x_a, p_a, x_b, p_b]^T$ corresponds to the eigenvalue of the operator \hat{R} . The Wigner function of the input state is denoted by $W(X_{\rm in})$. After the input state propagate through MZI, the output state can also be denoted with Wigner function $W(X_{\rm out}|\theta;\Phi)$ with phase shift by Eq. (4), where $X_{\rm out}$ is the eigenvalue vector corresponding to the quadrature operator of the output state. Detailed evolution of first-order and the second order moment of Wigner function corresponding to the input and output state are presented in Appendix (see supplementary material). After measuring the optimal quadrature in one output, the corresponding outcome x has a probability distribution $\mathcal{P}(x|\theta;\Phi)$ which is given by the marginal integral of the Wigner function over the conjugate quadrature x000 for one output port

$$\mathcal{P}(x|\theta;\Phi) = \int_{\mathbb{R}} W(X_{\text{out}}|\theta;\Phi) dp.$$
(5)

This can be generalized to the situation of partially homodyning a multimode bosonic system.

Recursive Bayesian inference for phase estimation. The input of algorithm consists of the CV quantum probe states $|\text{CV}\rangle$, the number of HDs M, the prior distribution $\mathcal{P}(\theta)$, the transmissivity of BS η_a , η_b , and the feedback learning rate λ . If no prior information is obtained, $\mathcal{P}(\theta)$ is set to be a flat distribution. The algorithm is flexible because it adapts to different phase estimation situations by setting the different composition of input parameters as Algorithm 1 presented.

Given an output data set $\mathcal{X} = \{x_1, x_2, \cdots, x_k\}$, the posterior probability for phase shift θ obeys the Bayes formula for which $\mathcal{P}(\theta|\mathcal{X}) = \mathcal{N}^{-1}\mathcal{P}(\mathcal{X}|\theta)\pi(\theta)$, where $\mathcal{N} = \int \mathcal{P}(\mathcal{X}|\theta)\pi(\theta)\mathrm{d}\theta$ is the normalized quantity to legalize the posterior probability distribution, $\pi(\theta)$ denotes the prior probability. Compared with maximum likelihood estimation (MLE), using the Bayesian method mainly has two merits²: (1) Prior information assists one ignore the impossible values the parameter would take. A long tail prior is also beneficial for us to estimate the unknown true phase shift; (2) Bayesian approach uses a set of possible θ generated by posterior and adopts the likelihood as a weight to estimate the parameter rather than using single estimation value of θ . Hence, Bayesian estimation fully utilizes the prior information to estimate the parameter faster. In particular, recursive estimation is applied in adaptive control, where the parameters of a dynamical system working in a closed loop are tracked using measurements of system input and output. We present a recursive estimation policy, where the transition kernel is an identity function since we assume the phase shift is invariant with time. In fact, recursive Bayesian estimation (RBE) is developed by applying the Bayesian view of statistical inference in general to the recursive estimation problem in particular. Moreover, RBE can utilize the outside feedback to automatically adjust the parameter in order to make the optimal policy. By recursive or reinforcement learning from the outside feedback phase, the optimal parameter estimation will finally be learned. However, the standard Bayesian inference cannot make full use of the feedback phase recursively to estimate the parameter.

The likelihood for kth HD is $\mathcal{P}(x_k|\theta; \Phi^{(k)})$, where x_k is the kth measurement outcome, and $\Phi^{(k)}$ is the kth feedback phase. The conceptual solution of this RBE for k number of measurements is given by,

$$\mathcal{P}(\theta|\mathcal{X}_k; \Xi_k) = \frac{\mathcal{P}(x_k|\theta; \Phi^{(k)})\mathcal{P}(\theta|\mathcal{X}_{k-1}; \Xi_{k-1})}{\mathcal{P}(x_k|\mathcal{X}_{k-1}; \Xi_{k-1})},\tag{6}$$

where

$$\mathcal{P}(x_k|\mathcal{X}_{k-1};\Xi_{k-1}) = \int_{\mathbb{R}} \mathcal{P}(x_k|\theta;\Phi^{(k)})\mathcal{P}(\theta|\mathcal{X}_{k-1};\Xi_{k-1})d\theta,$$
(7)

is the overall probability used to normalize the distribution. The recursion is initialized with $\mathcal{P}(\theta|\mathcal{X}_{-1}) = \pi(\theta)$. Note that $\Xi_k = [\Phi^{(1)}, \Phi^{(2)}, \cdots, \Phi^{(k)}]^T$ and $\mathcal{X}_k = [x_1, x_2, \cdots, x_k]^T$ are the columns of historical feedback phase and measurement outcomes for k number of measurements. The recursive process is demonstrated by line 3 to 14 in Algorithm 1, where $d_{\eta}^{o(k)}$, $\Sigma_{\eta}^{o(k)}$ denotes the kth first and second order moment of the quantum state before (k+1) th measurement, respectively. In particular, line $6{\sim}8$ implies a quantum physical interaction process, and costs constant number of operations O(K) in one iteration, where K is a constant value. The complexity for evaluation of the posterior distribution is a constant quantity denoted by O(K) when given the likelihood function and prior distribution. However, the exact normalization requires dividing the phase range into the finer grid which needs at least reaching the ideal precision level of VTB, otherwise the ideal precision can never be obtained under a rough grid (i.e., resolution is less than the ideal precision). Then the normalization process needs to sum over all possible phases within the grid to calculate the normalization constant \mathcal{N} . Hence, the complexity of exact normalization is inversely proportional to the ideal phase precision δ , namely $O(\delta^{-1})$ for each iteration.

malization is inversely proportional to the ideal phase precision δ , namely $O(\delta^{-1})$ for each iteration. We further design a strategy $\Gamma(\mathcal{X})$ to extract the information about θ from \mathcal{X} . $\theta^{(k)} = \Gamma(\mathcal{X}_k)$ is denoted as the kth extracted phase information from k number of measurements. Then, we adopt an updating approach for which $\Phi^{(k)} = (1-\lambda)\Phi^{(k-1)} + \lambda\theta^{(k)}$, where $\lambda \in [0,1]$, $\Phi^{(k)}$ is the feedback phase for kth iteration. This is a common learning (updating) strategy in machine learning algorithms. If $\lambda = 0$, it is equivalent to the case of no feedback control. The feedback process is demonstrated by line 12 and 13 of Algorithm 1. Intuitively, the feedback

phase Φ will approach true phase shift θ by our strategy. Note that if the measurement samples are small and Φ cannot be updated adequately towards to θ , which would cause a large oscillation at the final updating stage. Thus, as a thumb of rule, the chosen magnitude of λ is fairly matched with the measurements trials M to render the posterior distribution converge to a correct stationary distribution. The feedback process only needs to calculate the explicit function such that costs O(K) in one iteration. Thus, the computation complexity of our algorithm is $O(M\delta^{-1})$. Thus, if we input the coherent states, the ideal precision is the standard quantum limit (SQL) scaled by $\delta = (MN)^{-1/2}$ where N is the average photon number. If we input the quantum state, the ideal precision is HL scaled by $\delta = M^{-1/2}N^{-1}$. Therefore, the computation complexity of our algorithm is $O(M\delta^{-1})$. More specifically, the cost for SQL and HL is $O(\sqrt{M^3N})$ and $O(\sqrt{M^3N})$, respectively. Note that the normalization overhead of Algorithm 1 can be further reduced by using the technique of sampling methods^{37,38}.

Algorithm 1. RBE for CV quantum phase estimation.

Input: CV quantum states: $|CV\rangle = |x+ip\rangle$, number of HDs M, transmissivity of BS η_a, η_b , prior distribution $\mathcal{P}(\theta)$, and learning rate λ ;

Output: Phase estimation value of $\hat{\theta}$;

- 1: Initialize $\pi^{(0)} = \mathscr{P}(\theta), \, \mathscr{P}_{pos}^{(0)} = \pi^{(0)}, \, k = 0;$
- 2: Initialize likelihood function $\mathcal{L}^{(k)} = 0$, feedback phase $\Phi^{(0)} = 0$;
- 3: while $(k \le M)$ do
- $\pi^{(k)} = \mathscr{P}_{\mathrm{pos}}^{(k)};$ 4: 5.
 - $\mathscr{L}^{(k)} = 0 \times \pi^{(k)}$:
- Compute first-order moment and covariance matrix of the input CV state d_{in} , Σ_{in} ;
- Propagate the CV state through MZI, and obtain the output state $|CV\rangle_0$; Compute first-order moment and covariance matrix of the output CV state $d_{\eta}^{\text{o}(k)}, \Sigma_{\eta}^{\text{o}(k)}$ in which $\Phi = \Phi^{(k)}$;
- Measure the optimal quadrature of phase variable for $|CV\rangle_0$ by Eq.(5) to obtain the profitability $\mathscr{P}(x_k|\theta;\Phi^{(k)})$; Let $\mathscr{L}^{(k)} = \mathscr{P}(x_k|\theta;\Phi^{(k)})$, and compute the posterior distribution $\mathscr{P}_{pos}^{(k+1)} = \mathscr{L}^{(k)} \times \pi^{(k)}$;
- 10.
- Normalize the posterior distribution by Eq.(7); 11:
- 12. Compute feedback information according to feedback strategy and obtain the kth extracted feedback information $\theta_*^{(k)}$;
- Compute the (k+1)th feedback phase by $\Phi^{(k+1)} = \lambda \theta_*^{(k)} + (1-\lambda)\Phi^{(k)}$; 13: □ Update the feedback phase
- k = k + 1; 14:
- The estimated phase difference is calculated by Eq.(8) since the optimal precision is achieved when $\theta = \Phi^{(M)}$. The estimated phase shift $\hat{\theta}_{\mathrm{MSE}} = |\hat{\epsilon}_{\mathrm{MSE}} + \Phi^{(M)}|;$
- 17: Compare the estimated precision with the theoretical bound, and evaluate the performance of the algorithm.

We adopt the minimum mean-square error (MSE) as the benchmark to evaluate the precision performance. The output is the phase estimation value $\hat{\theta}_{\text{MSE}}$ which can be calculated by $\hat{\theta}_{\text{MSE}} = |\hat{\varepsilon}_{\text{MSE}} + \Phi^{(M)}|$, where $\Phi^{(M)}$ is the Mth (ultimate) feedback phase and $\hat{\varepsilon}_{\text{MSE}}$ is the ultimate phase difference (i.e., $\varepsilon = \theta - \Phi$) between two arms of MZI. Ideally, if $\hat{\varepsilon}_{MSE} = 0$, the visibility of MZI will be the highest, and then we capture the most accurate phase measurement. The phase difference $\hat{\varepsilon}_{\text{MSE}}$ is calculated by,

$$\hat{\varepsilon}_{\text{MSE}} = \int \hat{\varepsilon} \mathcal{P}(\varepsilon | \mathcal{X}_{M}; \Xi_{M}) d\hat{\varepsilon}. \tag{8}$$

▶ Update the prior with posterior distribution

To analyze the optimal performance of MSE benchmark, we assume a general estimator $\Theta(\mathcal{X})$, and it offers a necessarily imperfect estimation of the phase difference ε . Here, we present a Bayesian CRB suitable for finite samples with Van Trees inequality. The VTB in our physical model is given by,

$$\mathcal{V}^{2}[\Theta(\mathcal{X}); \varepsilon; \pi; \Xi] = \iint d\mathcal{X} d\varepsilon \pi \mathcal{P}(\mathcal{X}|\varepsilon; \Xi)(\Theta(\mathcal{X}) - \varepsilon)^{2},$$

$$\geq \mathcal{K}^{-1}(\pi; \Xi), \tag{9}$$

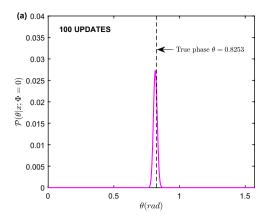
with

$$\mathcal{K}(\pi;\Xi) = \iint d\mathcal{X}d\varepsilon \{\mathcal{G}[\mathcal{P}(\mathcal{X}|\varepsilon;\Xi)] + \mathcal{G}[\pi]\}\mathcal{P}(\mathcal{X}|\varepsilon;\Xi)\pi$$

$$= \mathbb{E}[\mathcal{I}(\varepsilon;\Xi)] + \mathcal{I}(\pi), \tag{10}$$

where $\mathcal{G}[\alpha(\theta)] = (\partial_{\theta} \log \alpha)^2 \alpha$, $\mathcal{I}(\theta)$ is the Fisher information³⁹, Ξ and \mathcal{X} denote the set of feedback phase and measurement outcomes. Note that if the prior distribution is flat, the VTB equals to CRB. In general, we assume an optimal Gaussian prior of width $\sigma > 0$ centered at $\overline{\theta} = \theta_p$. The width of Gaussian prior σ needs to be restricted within an interval since a large width will cause the MSE rule not working well for phase estimation. More specifically, the Gaussian prior width needs at least half of the interval of θ to guarantee the efficiency⁴⁰. For a quantum phase estimation system, the quantum CRB (QCRB) is given by^{41,42},

$$\mathcal{V}^2[\Theta(\mathcal{X}); \theta] \ge \frac{1}{M\mathcal{F}(\theta)}.$$
 (11)



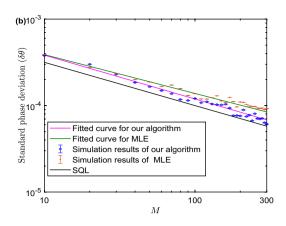


Figure 2. Monte Carlo simulation for coherent states with Algorithm 1 by setting the parameters $\eta_a = \eta_b = 1$, $\lambda = 0$, $\mathcal{P}(\theta) = 2/\pi$ compared with HD by using MLE. (a) The posterior probability distribution of phase shift are sharp with the measurement times M = 100 when $|\alpha|^2 = 100$. (b) The standard phase deviation $\delta\theta$ varies with the number of measurement times M when $|\alpha|^2 = 10^6$. Green and pink lines are fitted by the first typer of power function $y = ax^b$. Each trial is simulated with 10^4 times, and error bars are shown only if they are larger than the size of markers.

QCRB establishes that apart from the statistical scaling (M is the number of measurements), the variance of any unbiased estimator of the phase shift is bounded by the inverse of the quantum Fisher information $\mathcal{F}(\theta)^{30}$. In our physical model, all states and operations are Gaussian, thus $\mathcal{F}(\theta)$ can be calculated by⁴³,

$$\mathcal{F}(\theta) = \left(\frac{\partial d}{\partial \theta}\right)^{T} \Sigma^{-1} \frac{\partial d}{\partial \theta} + \text{Tr} \left[\left(\frac{\partial \Sigma}{\partial \theta} \Sigma^{-1}\right)^{2} \right]. \tag{12}$$

The QCRB is a useful tool to qualify the performance of an interferometer since it applies to all possible quantum measurements on the output modes. Our RBE-based machine learning algorithm develops an integral framework for quantum phase estimation to fast saturate theoretical precision using limited number of measurements for CV quantum phase estimation.

Performance analysis. Coherent state without feedback. According to the proposed algorithm, we investigate the phase estimation problem via Monte Carlo simulation with different typical Gaussian states. We consider a classical scenario where the coherent state $|\alpha\rangle\otimes|0\rangle=\hat{D}(\alpha,\,\theta_0)|0\rangle\otimes|0\rangle$ is regarded as input state. $\hat{D}(\alpha,\,\theta_0)$ is the displacement operator, and its matrix form in phase space is given by $D(\alpha,\,\theta_0)=\mathrm{diag}\{|\alpha|\cos\theta_0,\,|\alpha|\sin\theta_0\}$, where θ_0 is the displacement angle. Let $\theta_0=0$ implying that displace the vacuum state right along the x-axis in phase space. As for the coherent state, the input Wigner function is given by,

$$W(X_{\rm in}) = \frac{\exp[-(X - d_c)^T \Sigma_c^{-1} (X - d_c)]}{(2\pi)^2 \sqrt{\det(\Sigma_c)}},$$
(13)

where $\Sigma_c = \frac{1}{4}\mathbb{I}_4$ is the covariance matrix and $d_0 = \{|\alpha|, \, 0, \, 0, \, 0\}^T$ is the mean vector. When the coherent state passes through the lossless MZI $(\eta_a = \eta_{b_1} = 1)$, the first-order moment and covariance matrix of the output state are calculated, which is given by $d_{\text{opc}} = \frac{1}{2}|\alpha|\{\cos\varepsilon, -\sin\varepsilon, \sin\varepsilon, \cos\varepsilon\}^T$ and $\Sigma_{\text{opc}} = \frac{1}{4}\mathbb{I}_4$. $\varepsilon = \theta - \Phi$ is the phase difference between two arms. Here, we have ignored the global phase difference since it does not impact our results of phase estimation. In this part, let $\Phi = 0$, $\sigma = \infty$ because we do not consider the feedback control and prior information in our algorithm.

The quadrature $\hat{x}_{\psi} = \frac{1}{2}(e^{-i\psi}\hat{a} + e^{i\psi}\hat{a}^{\dagger})$ is measured by means of HD on repeated input state. Here, we find the quadrature \hat{x} measurement for second output port is optimal, thus let $\psi = 0$ and the probability of obtaining x for single measurement given by,

$$\mathcal{P}_{c}(x|\theta) = \sqrt{\frac{\pi}{2}} \exp\left\{-2\left(x - \frac{|\alpha|}{2}\sin\theta\right)^{2}\right\},\tag{14}$$

where $|\alpha|^2$ denotes the average photon number. The Fisher information of Eq. (14) $\mathcal{I}_c(\theta) = |\alpha|^2 \cos^2 \theta$. Thus, the VTB of using HD for M independent identical measurements is $(\delta\theta)_H^2 = 1/(M|\alpha|^2\cos^2\theta)$. Obviously, when $\cos^2\theta = 1$, i.e., $\theta_{\rm opt} = k\pi$, $k \in \mathbb{Z}$, the precision $(\delta\theta)_{\rm opt} = 1/(\sqrt{M}|\alpha|)$ is optimal, namely, the QCRB is calculated by Eq. (12). This bound is also called SQL. For comparison, we adopt the same measurement scheme with the MLE approach to evaluate the performance.

We first simulate Algorithm 1 with Monte Carlo method to evaluate the performance compared with the MLE approach. For simplicity, we set the transmissivity $\eta_a = \eta_b = 1$, $\lambda = 0$ and prior distribution $\pi(\theta) = 2/\pi$. When

 $|\alpha|^2=10$, the dynamic process of Algorithm 1 estimating the phase shift is also presented. The posterior probability is sharp while the measurement trials increase to 100 as Fig. 2(a) shows, which indicates that we accept the estimated phase with high confidence. When $|\alpha|^2=10^6$, our algorithm and MLE can both approach the SQL well as the pink and green fitted line in Fig. 2(b) shows. These two lines are fitted by the first type power function $y=ax^b$ since it fits the expression of $(\delta\theta)_{\rm opt}$. It demonstrates that MLE and our algorithm can both saturate the SQL given a classical input. In particular, our algorithm still presents an advantage of efficiency in terms of the convergence rate given the same measurement trials. That is, our algorithm presents a faster convergence rate for a given precision. Consequently, we can conclude that by using MLE and our algorithm respectively, the precision saturates the SQL for the coherent state and moreover, our algorithm presents an advantage of efficiency in terms of convergence rate given the same measurement trials.

CSV state without feedback. In this part, we discuss the performance of Algorithm 1 under input of quantum CV states. It is believed that given one port be coherent state, the optimal state which can achieve the HL scaling for another port of MZI is squeezed vacuum state 24,44,45 . CSV state can be denoted as $|\text{CSV}\rangle = |s\rangle \otimes |\alpha\rangle$, where $|s\rangle = S(r)|0\rangle$ is the squeezed vacuum state with $S(r) = \exp[r/2(\hat{a}^{\dagger 2} - \hat{a}^2)]$, and r is the squeeze parameter. S(r) can be represented by a linear transformation given by $S(r) = \text{diag}\{e^r, e^{-r}\}$. The covariance matrix of squeeze vacuum state $V = S(r)S(r)^T = \text{diag}\{e^{2r}, e^{-2r}\}$. Therefore, the first-order moment and covariance matrix of the input state are given by $m_{\text{in}}^{\text{CSV}} = \{|\alpha|, 0, 0, 0\}^T, \sum_{\text{CisV}}^{\text{CSV}} = \frac{1}{4}\text{diag}\{1, 1, e^{2r}, e^{-2r}\}$. After CSV state propagating through the MZI and using HD for optimal phase quadrature \hat{x} of the second output port, the mean and variance are given by (here we assume $\Phi = 0$, $\sigma = \infty$),

$$m_{\rm o}^{\rm CSV} = \frac{1}{2} |\alpha| \sin \theta, \tag{15}$$

$$(\sigma_0^{\text{CSV}})^2 = \frac{1}{32} e^{-2r} (2e^{4r} \sin^2 \theta - 8e^r \sinh r \cos \theta + 4e^{2r} + \cos(2\theta) + 3). \tag{16}$$

It is readily to find that the minimum variance is taken at $\theta = 0$ and equals to $\frac{1}{4}e^{-2r}$. The probability of obtaining quadrature \hat{x} is given by,

$$\mathcal{P}_{\text{CSV}}(x|\theta;\Phi) = \frac{1}{\sqrt{2\pi(\sigma_0^{\text{CSV}})^2}} \exp\left(-\frac{(x-m_0^{\text{CSV}})^2}{2(\sigma_0^{\text{CSV}})^2}\right). \tag{17}$$

By Eq. (9), we figure out the phase sensitivity determined by the phase difference for M number of repeated measurements given by,

$$\delta\theta = \sqrt{\frac{2e^{4r}\sin^2\theta - 8e^r \sinh r \cos\theta + 4e^{2r} + \cos 2\theta + 3}{8M|\alpha|^2 e^{2r}\cos^2\theta}}.$$
 (18)

The VTB is achieved at $\theta=k\pi$, $k\in\mathbb{Z}$ (called sub-shot noise limit) for which $(\delta\theta)_{\mathrm{opt}}^{\mathrm{CSV}}=e^{-r}/\sqrt{M}|\alpha|$. By Eq. (12), we calculate the QCRB $(\delta\theta)_{\mathrm{Q}}^{\mathrm{CSV}}=[M(|\alpha|^2e^{2r}+\sinh^2r)]^{-1/2}$. When $|\alpha|^2\gg\sinh^2r$ (i.e. \sinh^2r can be neglected), the sub-shot noise limit is approximately equal to the HL. Given the total average photon number and if $|\alpha|^2\cong\sinh^2r$ (i.e., \sinh^2r cannot be neglected), the optimal phase sensitivity is achieved by parity or photon number detection 24,34 . In particular, photon number detection can avoid the dependence of phase estimation accuracy on the true phase value because it fully utilizes the statistical feature of each photon. However, these measurement methods for Heisenberg scaling are not implemented readily in practical quantum physical systems and more significantly, they are more sensitive to photon loss in lossy MZI^{46,47}.

Similarly, we still adopt the MLE as a comparison to illustrate the superiority of our algorithm. The standard phase deviation with CSV as input can be figured out for different number of measurements when $|\alpha|^2=100,\ r=0.8,\ 1.8$ as Fig. 3(a–c) shows, in which all the lines are fitted by the first type of power function $y=ax^b$ for which is consistent with the form of $(\delta\theta)_{\rm opt}^{\rm CSV}$. Here, we set the parameters $\lambda=0,\ \eta_a=\eta_b=1,\ \mathcal{P}(\theta)=2/\pi$. In our physical model if we set $\lambda=0$, HD without feedback control cannot eliminate the phase dependence of

In our physical model if we set $\lambda=0$, HD without feedback control cannot eliminate the phase dependence of estimation on true phase shift, thus the phase sensitivity relies on the true phase shift in the whole interval of $[0, \pi/2]$. As Fig. 3(a) implies, when $\theta=0.1$, 0.3, respectively, the standard phase deviation $\delta\theta$ changes obviously and small phase shift, i.e. near the optimal phase, will have a better performance. Because if phase shifts severe, the measured fluctuation will be large, which will cause the effectiveness of squeeze operation to enhance precision being not achieved. In addition, we simulate our algorithm and HD with MLE, respectively as Fig. 3(b,c) shows. When $|\alpha|^2=100$, r=0.8 and the phase shift is set to be optimal, we find that HD with MLE cannot saturate sub-shot noise limit (near QCRB) in a few trials (M<300). In particular, when r increases, we find that the gap between the sub-shot noise limit and the simulation results of HD with MLE becomes larger in the early stage of training. It indicates that MLE needs more samples to calculate the likelihood function and then estimate the optimal phase parameter. When the samples size is large, our algorithm will have nearly the same performance with MLE. Our algorithm saturates the sub-shot noise limit efficiently whatever r is larger or small. However, MLE shows a relatively slow convergence rate when r is small since small r indicates the larger fluctuations of output state. In conclusion, our algorithm shows a notable performance when the input is a quantum CV state. More specifically, our algorithm can saturate the sub-shot noise limit (when $|\alpha|^2\gg 1$, near QCRB or HL) effi-

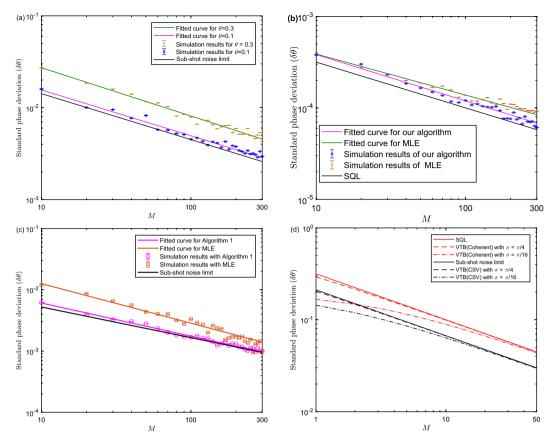


Figure 3. Simulation results for CSV probe states via Algorithm 1 with $\eta_a = \eta_b = 1$, $\lambda = 0$, $\mathcal{P}(\theta) = 2/\pi$. (a) The true phase shift sets to 0.1 and 0.3, respectively. The pink and green curves are fitted by the power function. The true phase shift sets to be optimal for HD with MLE (brown) and our algorithm (pink) for (b,c), and $|\alpha|^2 = 100$, r = 0.8 for (b), $|\alpha|^2 = 100$, r = 1.8 for (c). The pink and green curves are fitted by the first type of power function $y = ax^b$. Each trial is simulated with 10^4 times and error bars are shown only if they are larger than the size of markers (d). The enhancement of VTB with respect to coherent (blue) and CSV (yellow) state compared with conventional bound.

ciently compared with HD by using MLE when the input is CSV state in a situation of limited measurement trials.

Prior information analysis. In this part, we will compare the precision of existing uniform and different width of Gaussian prior distributions under a given phase shift for different probe states. In general, we assume a Gaussian prior of width $\sigma > 0$ centered at $\theta = \theta_p$, that is,

$$\pi(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{\left(\theta - \theta_p\right)^2}{2\sigma^2}\right\},\tag{19}$$

with mean value θ_p and variance σ^2 . The Fisher information of $\pi(\theta)$ can be calculated by Eq. (9) for which $\mathcal{I}[\pi(\theta)] = \sigma^{-2}$. Thus, the VTB for coherent input states with M repeated HDs as Fig. 1 depicted is given by,

$$\mathcal{V}_{c}^{2}[\Theta(\mathcal{X}); \pi] \ge \frac{1}{M|\alpha|^{2} + \sigma^{-2}}.$$
(20)

Note that the width of Gaussian prior σ needs to be restricted within an interval since a large width (variance) will cause the MSE rule not working well for phase estimation⁴⁰. Eq. (20) demonstrates that if we have much information on the phase shift before measurements, i.e., the width of $\sigma \ll 1$, the precision will be highly improved. Similarly, if the CSV states are the input, the VTB with M repeated HDs is given by,

$$\mathcal{V}_{s}^{2}[\Theta(\mathcal{X}); \pi] \ge \frac{1}{M|\alpha|^{2}e^{2r} + \sigma^{-2}}.$$
(21)

We plot the VTB for M repeated measurements with respect to three different Gaussian width $\sigma = \infty$, $\pi/4$, $\pi/16$ when $|\alpha|^2 = 10$ for coherent state and $|\alpha|^2 = 10$, r = 0.8 for CSV state as Fig. 3(d) shows.

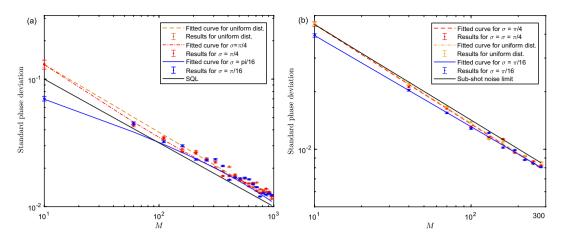


Figure 4. Phase estimation with different Gaussian prior width σ via Algorithm 1 for $\eta=1$, $\lambda=0$, as a function of measurement trials M. The fitted curves adopt the second type of power function $y=ax^b+c$ as the fitting function since the prior information σ is an invariant with M as Eq. (20,21) indicated. Monte Carlo simulation parameters of $|\alpha|^2=10$, $\theta=0.8253$, M=1000 set for left panel of the graph, and $|\alpha|^2=10$, r=0.8, $\theta=0.2352$, M=300 for right panel of the graph.

Theoretically, the VTB is superior to the CRB with a prior information σ according to Eqs (20–21). As expected, the improvement will gradually vanish when measurement times M increases since σ^{-2} can be neglected compared with the Fisher information of the posterior distribution. However, when $M \le 10$, the enhancement of $\delta\theta$ is obvious with a small Gaussian width. Therefore, the VTB can adapt to the situation where the measurements are extremely limited and the prior information is provided.

Without loss of generality, we will set a relatively small phase shift to analyze the impact of prior knowledge on the ultimate estimated precision. For simplicity, we will not consider the case of phase ambiguity, thus the feedback control will not be adopted i.e., $\lambda = 0$. If the probe states are CSV states, we know that the effect of phase shift on precision has been eliminated via feedback control. In Fig. 4(a,b), the fitted lines are fitted by the second type of power function $y = ax^b + c$ on account of the invariant impact of prior information playing the role of constant value c when $\sigma = \pi/4$, $\pi/16$. It is found that the second type of power function has a better fitting fidelity in the simulation since the goodness of fit coefficient $R^2 = 0.9923$ is larger than the situation where the first type of power function is employed whose $R^2 = 0.9732$ when the input state is the coherent state. Similarly, when the input state is CSV state, the goodness of fit coefficient $R^2 = 0.9923$ of the second type of power function is larger than the first type whose $R^2 = 0.9852$ as expected. The difference is not such significant since the prior information σ reduces smaller quantum fluctuations compared to the input being the coherent state. We find that the Gaussian prior knowledge can certainly enhance the precision for coherent and CSV state with small measurement trials (M < 100), but its impact vanishes when trial becomes large because our algorithm uses the history measurement information to update the prior distribution, which also validates the above theoretical analysis. In particular, the improvement when $\sigma = \pi/4$ is not greater than $\sigma = \pi/16$. Therefore, smaller width will have a better performance improvement on precision. The precision of phase estimation can be reached more efficiently when given the prior information.

Feedback strategy and photon loss analysis. In general, if the phase shift is limited to the interval $[0, \pi/2]$, our algorithm with $\lambda = 0$ performs well both for coherent and CSV state as described above. However, in order to eliminate phase ambiguity in $[0, 2\pi]$ which exists in many CV quantum phase estimation problems⁴⁸, we present an explicit feedback strategy which can extract information⁴⁹ about phase shift from each measurement. For kth measurement, we define the information feedback function explicitly for $\Gamma(\mathcal{X}_k)$ which is given by,

$$\theta_*^{(k)} = \Phi^{(k)} - \arcsin\left(\frac{2x_k}{|\alpha|}\right) \zeta\left(\frac{|\alpha|}{2} - |x_k|\right) - \frac{\pi}{2} \operatorname{sgn}(x_k) \zeta\left(|x_k| - \frac{|\alpha|}{2}\right),\tag{22}$$

where $\zeta(z)=1, z\geq 0, \zeta(z)=0, z<0$, and $\operatorname{sgn}(\cdot)$ is the sign function. Eq. (22) denotes that an effective phase feedback occurs when the measured quadrature phase is less than the half of squared average photon number, otherwise executing $\pi/2$ rotation in an opposite direction. It is observed that large quantities of HD outcomes x_k satisfy the condition of $x_k>|\alpha|/2$, and the second case for which $x_k>|\alpha|/2$ is not relevant for highly displaced coherent input states, where such events become very rare. If we set the threshold value of $|\alpha|/2$ even larger, then more samples will be collected to calculate the feedback phase information. This may speed-up the convergence of the feed-back phase to the true phase but is at the cost of reducing the robustness especially in the final stage. The corespondent probability distribution of $x_k>|\alpha|/2$ for coherent and CSV state is given by,

$$\mathcal{P}_{c}(\theta_{*}^{(k)}) = \frac{1}{2} \operatorname{erfc} \left[\frac{|\alpha|}{\sqrt{2}} (1 \pm \sin(\varepsilon^{(k)})) \right], \tag{23}$$

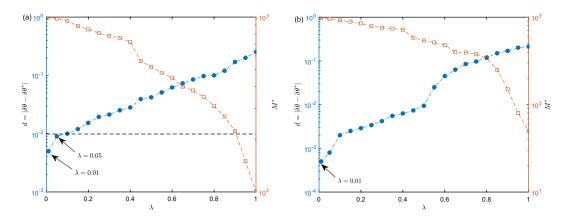


Figure 5. Optimization of the learning rate λ . (a) For the coherent state, the left blue axis is plotted by executing Algorithm 1 when given $|\alpha|^2 = 10$, $M = 10^3$. The right yellow axis is plotted with $|\alpha|^2 = 10$. (b) For CSV state, the left blue axis is plotted with $|\alpha|^2 = 10$, r = 0.8, $M = 10^3$. The right yellow axis is plotted with $|\alpha|^2 = 10$, r = 0.8.

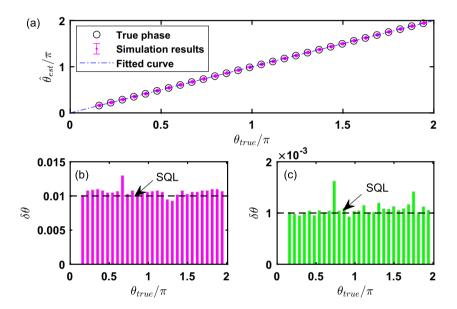


Figure 6. The phase ambiguity of coherent state is eliminated via Algorithm 1 with feedback strategy $\Gamma(\mathcal{X})$ by setting $\mathcal{P}(\theta) = 1/2\pi$, $\eta_a = \eta_b = 1$, $M = 10^3$, $\lambda = 0.01$. (a) The true phase and estimated phase are plotted with $|\alpha|^2 = 10$ and error bars are shown only if they are larger than the size of markers. The bars of (b,c) are simulated with $|\alpha|^2 = 10$, 10^3 , respectively.

$$\mathcal{P}_{s}(\theta_{*}^{(k)}) = \frac{1}{2} \operatorname{erfc} \left[\frac{\sqrt{(\sinh^{2}r + |\alpha|^{2})}}{\sigma_{o}^{CSV}(\varepsilon^{(k)})} (1 \pm \sin(\varepsilon^{(k)})) \right], \tag{24}$$

where $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-t^2) \mathrm{d}t$ denotes the complementary error probability function. Eq. (23) denotes the probability that we cannot extract effective phase information. Here, we provide a heuristic way of trying to extract true phase information among the measured samples. This strategy is a tracking process such that the feedback phase will be updated towards to the true phase shift when one obtains the measurement results. Thus, the true phase shift θ is theoretically broadened to the whole interval of $[0, 2\pi]$ since the feedback phase can always be adjusted to track the unknown phase shift.

In addition, the learning rate λ in our algorithm as a hyperparameter needs to be optimized otherwise will cause the posterior distribution not converged or converged to a wrong position. Intuitively, $\lambda \to 0$ is equivalent to the case of no feedback control, which thus will render the posterior converge to an ideal position by using a large M. However, the dynamical range of phase estimation cannot be broadened under this parameter setting. When $\lambda \to 1$, the posterior distribution will converge fast to a wrong position only by using fewer trials M. Therefore, we need a trade-off between the convergent trials M^* and error precision $d = |\delta \theta - \delta \theta^*|$, where $\delta \theta^*$ is the ideal precision calculated by VTB. A reasonable rule for the acceptance is that $d < \delta \theta^*$. In order to optimize λ ,

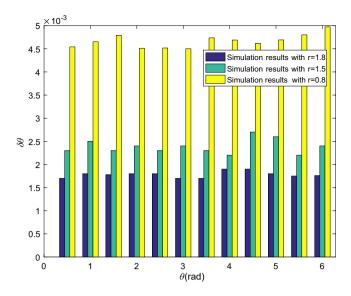


Figure 7. The phase ambiguity is eliminated via Algorithm 1 with $\Gamma(\mathcal{X})$. The three kinds of bars denote the standard phase deviation $\delta\theta$ with respect to $r=1.8,\ 1.5,\ 0.8$ for true phase shift in $[0,2\pi]$. The simulation parameters are set to be $|\alpha|^2=10, M=10^3, \lambda=0.01,\ \eta_a=\eta_b=1$.

we test the error precision d and M^* under a set of different value of λ for the coherent and CSV state, respectively. For the coherent state, we set $|\alpha|^2=10$, $\eta_a=\eta_b=1$, $M=10^3$ in Algorithm 1 to test the error precision d with different learning rate λ . As Fig. 5(a) shows, the error precision preserves a small and an acceptable value when $\lambda=0.01$, 0.05 and M^* approximately equal to M, which indicates that the posterior distribution converges to a correct position. For larger λ , d is unacceptable although the convergent trials M^* becomes smaller. Similar for CSV state, we set $|\alpha|^2=10$, r=0.8, $\eta_a=\eta_b=1$, $M=10^3$ to test the error precision under different value of λ . In Fig. 5(b), the error precision maintains a similar feature with the coherent state and are acceptable when $\lambda=0.01$, $M^*\approx10^3$, which demonstrates that the phase precision is still approximately optimal and not reduced by the feedback learning. Moreover, it is found that $M=10^3$ is matched with the learning rate $\lambda=0.01$ both for the coherent and CSV state. Larger M will also be appropriate for $\lambda=0.01$ for which can render posterior distribution converged correctly, i.e., $M^*=10^3$ is the minimum measurement trials for $\lambda=0.01$.

We analyze two cases for classical input where $|\alpha|^2=10,\ 10^3$ respectively and the phase shift $\theta\in[0,2\pi]$ via Algorithm 1 with feedback function $\Gamma(\mathcal{X})$ as Fig. 6 shows. The optimal choice of λ , M^* can be appropriate for the case of large photon numbers since d is not directly steered by α . Therefore, we set the parameters $\eta_a=\eta_b=1,\ \mathcal{P}(\theta)=1/2\pi,\ \lambda=0.01$ to analyze the performance with the feedback strategy. The results show that for arbitrary phase shift, the precision saturates the SQL as Fig. 6(a-c) shows regardless of small or large $|\alpha|^2$. Thus, it demonstrates that Algorithm 1 with feedback function $\Gamma(\mathcal{X})$ can eliminate the phase ambiguity of CV phase estimation in $[0,2\pi]$. In order to eliminate the impact of true phase shift on the precision of parameter and phase ambiguity for CSV state, we adopt the same feedback strategy as the coherent state does. The precision performance will not be steered by the true phase shift via Algorithm 1 with the proposed feedback strategy $\Gamma(\mathcal{X})$ if we can eliminate the phase ambiguity since the feedback phase can always be adjusted towards the true phase. We simulate Algorithm 1 with the phase shifts sampled from the interval of $[0,2\pi]$ with $|\alpha|^2=10$, r=0.8, 1.5, 1.8 and measurement trials $M=10^3$ for $\lambda=0.01$, $\eta_a=\eta_b=1$ as Fig. 7 shows. The three bars with respect to different squeeze parameter basically reach their sub-shot noise limit $\delta_{\mathrm{Sub-shotnoise}}=0.0045,\ 0.0022,\ 0.0017,\ \text{which demonstrates}$ that the phase ambiguity is also eliminated for CSV state. Our feedback strategy is an explicit function and need not complex optimization process to obtain the feedback information compared with ref. Thus, it can be extended to many real-time phase estimation scenarios.

In our algorithm, we mainly consider the linear photon loss caused by photon loss inside the interferometer and photon loss due to the inefficient detector. Specifically, assuming the losses are existed in two arms and inserted between the first BS and phase shifter ⁴⁷. Here, the photon losses are symmetric in two arms, i.e., $\eta = \eta_a = \eta_b$, and we aim to validate whether our algorithm is robust against linear photon losses. The mean and variance value of the coherent state as input in terms of symmetric lossy MZI are calculated with detection of quadrature \hat{x} of the second port for which $m_{\eta}^{\rm o} = 1/2\sqrt{\eta} |\alpha| \sin \varepsilon$, $(\sigma_{\eta}^{\rm o})^2 = 1/4$. Thus, we can calculate the VTB with a prior information σ under symmetric linear photon loss for M repeated measurements given by,

$$(\delta \varepsilon)_C^{\text{loss}} = \sqrt{\frac{1}{M\eta |\alpha|^2 + \sigma^{-2}}}.$$
 (25)

The optimal phase difference is $\varepsilon_{\mathrm{opt}}^C = k\pi$, $k \in \mathbb{Z}$. Eq. (25) can also be calculated by Eq. (12), implying HD is optimal under linear photon loss. Similarly, for CSV state, we have $m_{\eta}^{\mathrm{oCSV}} = \frac{1}{2} \sqrt{\eta} |\alpha| \sin \varepsilon$, $\left(\sigma_{\eta}^{\mathrm{oCSV}}\right)^2$

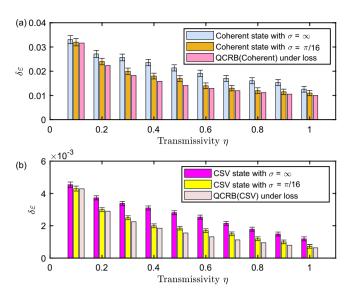


Figure 8. Influence of different transmissivity η on parameter estimation precision. The input parameters $\lambda=0.01, M=10^3, \eta_a=\eta_b, \sigma=\infty, \pi/16$. (a) The blue bars denote the simulation results for the coherent state with $|\alpha|^2=10, \sigma=\infty$ via Algorithm 1 with $\Gamma(\mathcal{X})$, and the orange bars are the simulation results with $|\alpha|^2=10, \sigma=\pi/16$ via Algorithm 1 with $\Gamma(\mathcal{X})$ and the pink bars display the QCRB under the condition of linear photon loss with the flat prior distribution calculated by Eq. (25). (b) The red bars are the simulation results for CSV state with $r=0.8, |\alpha|^2=10, \sigma=\infty$, and the yellow bars denote the simulation results for CSV state with $r=0.8, |\alpha|^2=10, \sigma=\pi/16$, and the pink bars are the QCRB in linear lossy MZI with the flat prior distribution, respectively.

 $= \frac{1}{16} \left(-3\eta + \eta e^{2r} \sin^2 \varepsilon + 4\eta e^{-2r} \cos^4 \left(\frac{\varepsilon}{2} \right) - \eta \cos \varepsilon + 4 \right).$ The VTB of CSV state as input with *a prior* information σ under symmetric linear photon loss for *M* repeated measurements is given by,

$$(\delta \varepsilon)_{\text{CSV}}^{\text{loss}} = \sqrt{\frac{\eta e^{-2r} + 1 - \eta}{M\eta |\alpha|^2 + (\eta e^{-2r} + 1 - \eta)\sigma^{-2}}}.$$
(26)

The optimal phase difference $\varepsilon=k\pi$, $k\in\mathbb{Z}$. Note that $(\delta\varepsilon)_{\mathrm{CSV}}^{\mathrm{loss}}=\sqrt{1/M|\alpha|^2e^{2r}+(1-\eta)/M\eta|\alpha|^2}$ when the prior distribution is uniform, which implies that if we increase $|\alpha|^2$ (coherent power), the effects of photon loss will be suppressed dramatically and ultimately the phase precision approaches to the sub-shot noise limit. The QCRB of CSV state as input in a lossy MZI without prior information for M repeated measurements can be calculated by Eq. (12) given by,

$$(\delta \varepsilon)_{Q}^{loss} = \sqrt{\frac{(1 - \eta)(e^{2r} - 1) + 1}{M\eta\{|\alpha|^{2}e^{2r} + \sinh^{2}r[(1 - \eta)(e^{2r} - 1) + 1]\}}},$$

an equation that can also be found in the literature⁵⁰. We find that when $|\alpha|^2 \gg 1$ and $\sigma = \infty$, $(\delta \varepsilon)_{\text{CSV}}^{\text{loss}} \cong (\delta \varepsilon)_Q^{\text{loss}}$.

We discuss the impact of transmissivity η of BS on phase estimation by Algorithm 1 with $\Gamma(\mathcal{X})$ when given different prior information with respect to the coherent and CSV state. As Fig. 8(a) demonstrates, compared to the case of no prior information assisted, the phase deviation $\delta\varepsilon$ nearly saturates the VTB (QCRB) in lossy MZI when given the prior information $\sigma=\pi/16$, which indicates that a prior integrated algorithm can enhance the robustness of the phase estimation. From Fig. 8(b), compared to the case of $\sigma=\infty$, the standard deviation $\delta\varepsilon$ of the simulation results basically saturates the QCRB when given the prior information $\sigma=\pi/16$. In particular, for small η , the performance is significantly superior to the situation where no prior information is employed. When η becomes large, the enhancement are not such significant, which conversely implies that a prior information can enhance the robustness of the algorithm. As a result, it is implied that the standard phase deviation can saturate the lossy QCRB for the coherent and CSV state by executing Algorithm 1 with feedback strategy when given the prior information, which also demonstrates that the effect of linear photon loss on the phase estimation precision can be suppressed dramatically.

Conclusions

We investigate a quantum phase estimation framework for Gaussian states based on a machine learning approach. We incorporate the recursive Bayesian estimation method and an explicit feedback strategy to the algorithm organically, which renders it more flexible in phase estimation tasks. We have studied two probe practical Gaussian resources that are available in the laboratory and easily produced, for optical quantum metrology. Based on Algorithm 1, compared with the conventional scheme, it is efficient to approach the theoretical bound with feasible CV states. Moreover, by assuming a Gaussian prior being employed, we find that the precision

enhancement corresponds to the different width σ , i.e. prior information assists the phase estimation being more efficient. In particular, the phase ambiguity of coherent and CSV state can be eliminated via Algorithm 1 with the proposed feedback strategy such as to achieve the physical limits with arbitrary phase shift in $[0, 2\pi]$. To validate the robustness of Algorithm 1, we analyze the lossy QCRB in MZI with symmetric linear photon loss for the coherent and CSV state, and the simulation results present negligible precision reduction compared with theoretical lossy QCRB especially for the case of *a prior* information integrated, which implies that the photon losses are suppressed dramatically by Algorithm 1 with our feedback strategy and prior information. Our CV quantum phase estimation framework highlights the machine learning method, studies the CV phase estimation and can be extended to the time-variable or multi-parameter estimation framework.

References

- 1. Lumino, A. et al. Experimental phase estimation enhanced by machine learning. Phys. Rev. Appl. 10, 044033 (2018).
- 2. Nasrabadi, N. M. Pattern recognition and machine learning. J. electronic imaging 16, 049901 (2007).
- 3. Hentschel, A. & Sanders, B. C. Efficient algorithm for optimizing adaptive quantum metrology processes. *Phys. review letters* 107, 233601 (2011).
- 4. Wiebe, N. & Granade, C. Efficient bayesian phase estimation. Phys. review letters 117, 010503 (2016).
- 5. Hentschel, A. & Sanders, B. C. Machine learning for precise quantum measurement. Phys. review letters 104, 063603 (2010).
- Palittapongarnpim, P., Wittek, P. & Sanders, B. C. Robustness of learning-assisted adaptive quantum-enhanced metrology in the presence of noise. In Systems, Man, and Cybernetics (SMC), *IEEE International Conference on*, 294–299 (IEEE, 2017).
- 7. Choo, K., Carleo, G., Regnault, N. & Neupert, T. Symmetries and many-body excitations with neural-network quantum states. *Phys. review letters* 121, 167204 (2018).
- 8. Carleo, G. & Troyer, M. Solving the quantum many-body problem with artificial neural networks. Sci. 355, 602-606 (2017).
- 9. Berni, A. A. et al. Ab initio quantum-enhanced optical phase estimation using real-time feedback control. Nat. Photonics 9, 577 (2015).
- Berry, D. & Wiseman, H. Optimal states and almost optimal adaptive measurements for quantum interferometry. Phys. review letters 85, 5098 (2000).
- 11. Berry, D. W., Wiseman, H. & Breslin, J. Optimal input states and feedback for interferometric phase estimation. *Phys. Rev. A* 63, 053804 (2001).
- 12. Distante, E., Ježek, M. & Andersen, U. L. Deterministic superresolution with coherent states at the shot noise limit. *Phys. review letters* 111, 033603 (2013).
- 13. Torlai, G. et al. Neural-network quantum state tomography. Nat. Phys. 14, 447 (2018).
- 14. Wang, J. et al. Scalable quantum tomography with fidelity estimation. arXiv preprint arXiv:1712.03213 (2017).
- 15. Flammia, S. T., Gross, D., Liu, Y.-K. & Eisert, J. Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators. *New J. Phys.* 14, 095022 (2012).
- 16. Huszár, F. & Houlsby, N. M. Adaptive bayesian quantum tomography. Phys. Rev. A 85, 052120 (2012).
- 17. Lugiato, L., Gatti, A. & Brambilla, E. Quantum imaging. J. Opt. B: Quantum semiclassical optics 4, S176 (2002).
- 18. Giovannetti, V., Lloyd, S. & Maccone, L. Quantum metrology. Phys. review letters 96, 010401 (2006).
- 19. Huang, D., Huang, P., Lin, D. & Zeng, G. Long-distance continuous-variable quantum key distribution by controlling excess noise. *Sci. reports* 6, 19201 (2016).
- 20. Gisin, N. & Thew, R. Quantum communication. Nat. photonics 1, 165 (2007).
- 21. Afek, I., Ambar, O. & Silberberg, Y. High-noon states by mixing quantum and classical light. Sci. 328, 879-881 (2010).
- 22. Schnabel, R. Squeezed states of light and their applications in laser interferometers. Phys. Reports 684, 1-51 (2017).
- 23. Lami, L. et al. Gaussian quantum resource theories. arXiv preprint arXiv:1801.05450 (2018).
- 24. Pezzé, L. & Smerzi, A. Mach-zehnder interferometry at the heisenberg limit with coherent and squeezed-vacuum light. *Phys. review letters* 100, 073601 (2008).
- 25. Anisimov, P. M. *et al.* Quantum metrology with two-mode squeezed vacuum: parity detection beats the heisenberg limit. *Phys. review letters* **104**, 103602 (2010).
- 26. Huang, Z., Motes, K. R., Anisimov, P. M., Dowling, J. P. & Berry, D. W. Adaptive phase estimation with two-mode squeezed vacuum and parity measurement. *Phys. Rev. A* **95**, 053837 (2017).
- 27. Gill, R. D. et al. Applications of the van trees inequality: a bayesian cramér-rao bound. Bernoulli 1, 59-79 (1995).
- 28. Seveso, L., Rossi, M. A. & Paris, M. G. Quantum metrology beyond the quantum cramér-rao theorem. *Phys. Rev. A* **95**, 012111 (2017).
- 29. Martínez-Vargas, E., Pineda, C., Leyvraz, F. & Barberis-Blostein, P. Quantum estimation of unknown parameters. *Phys. Rev. A* **95**, 012136 (2017).
- 30. Chiruvelli, A. & Lee, H. Quantum cramer-rao bound and parity measurement. In *Advances in Photonics of Quantum Computing, Memory, and Communication III*, vol. 7611, 76110S (International Society for Optics and Photonics, 2010).
- 31. Giovannetti, V., Lloyd, S. & Maccone, L. Advances in quantum metrology. *Nat. photonics* 5, 222 (2011).
- 32. Pezze, L. & Smerzi, A. Quantum theory of phase estimation. arXiv preprint arXiv:1411.5164 (2014).
- 33. Dowling, J. P. Quantum optical metrology-the lowdown on high-noon states. Contemp. Phys. 49, 125-143 (2008).
- 34. Gerry, C. C. & Mimih, J. The parity operator in quantum optical metrology. Contemp. Phys. 51, 497-511 (2010).
- 35. Braunstein, S. L. & Van Loock, P. Quantum information with continuous variables. Rev. Mod. Phys. 77, 513 (2005).
- 36. Weedbrook, C. et al. Gaussian quantum information. Rev. Mod. Phys. 84, 621 (2012).
- 37. Chib, S. & Jeliazkov, I. Marginal likelihood from the metropolis-hastings output. J. Am. Stat. Assoc. 96, 270-281 (2001).
- 38. Haario, H. et al. An adaptive metropolis algorithm. Bernoulli 7, 223-242 (2001).
- 39. Gao, Y. & Lee, H. Bounds on quantum multiple-parameter estimation with gaussian state. The Eur. Phys. J. D 68, 347 (2014).
- 40. Friis, N. et al. Flexible resources for quantum metrology. New J. Phys. 19, 063044 (2017).
- 41. Paris, M. G. Quantum estimation for quantum technology. Int. J. Quantum Inf. 7, 125-137 (2009).
- 42. Brody, D. C. & Hughston, L. P. Statistical geometry in quantum mechanics. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, 2445–2475 (The Royal Society, 1998).
- 43. Gard, B. T. et al. Nearly optimal measurement schemes in a noisy mach-zehnder interferometer with coherent and squeezed vacuum. Epj Quantum Technol. 4, 4 (2017).
- 44. Sparaciari, C., Olivares, S. & Paris, M. G. Bounds to precision for quantum interferometry with gaussian states and operations. *JOSA B* **32**, 1354–1359 (2015).
- 45. Lang, M. D. & Caves, C. M. Optimal quantum-enhanced interferometry using a laser power source. *Phys. review letters* 111, 173601 (2013).
- 46. Zhang, X.-X., Yang, Y.-X. & Wang, X.-B. Lossy quantum-optical metrology with squeezed states. Phys. Rev. A 88, 013838 (2013).
- Oh, C., Lee, S.-Y., Nha, H. & Jeong, H. Practical resources and measurements for lossy optical quantum metrology. Phys. Rev. A 96, 062304 (2017).
- 48. Olivares, S. & Paris, M. G. Bayesian estimation in homodyne interferometry. J. Phys. B: At. Mol. Opt. Phys. 42, 055506 (2009).

- 49. D'Ariano, G. M., Paris, M. G. & Seno, R. Feedback-assisted homodyne detection of phase shifts. Phys. Rev. A 54, 4495 (1996).
- 50. Ono, T. & Hofmann, H. F. Effects of photon losses on phase estimation near the heisenberg limit using coherent light and squeezed vacuum. *Phys. Rev. A* 81, 033819 (2010).

Acknowledgements

This work is supported by National Natural Science Foundation of China (Grant No. 61701302 and 61631014).

Author Contributions

T.X. and J.H. conceived the experiments, and T.X. conducted the experiments, J.F. and G.Z. provided the motivation and realization scheme and analysed the results. All the authors also had a fruitful systematic discussion on the progress of the work. The manuscript was prepared with inputs by all the authors. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-48551-0.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019