Efficacy of virtual purification-based error mitigation on quantum metrology

Hyukgun Kwon[1], Changhun Oh[2], Youngrong Lim[3], Hyunseok Jeong[1,*], and Liang Jiang[2,†]

[1]*Department of Physics and Astronomy, Seoul National University, Seoul 08826, Republic of Korea*
[2]*Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA*
[3]*School of Computational Sciences, Korea Institute for Advanced Study, Seoul 02455, Republic of Korea*

Noise is the main source that hinders us from fully exploiting quantum advantages in various quantum informational tasks. However, characterizing and calibrating the effect of noise is not always feasible in practice. Especially for quantum parameter estimation, an estimator constructed without precise knowledge of noise entails an inevitable bias. Recently, virtual purification-based error mitigation (VPEM) has been proposed to apply for quantum metrology to reduce such a bias occurring from unknown noise. While it was demonstrated to work for particular cases, whether VPEM always reduces a bias for general estimation schemes is unclear. For more general applications of VPEM to quantum metrology, we study factors determining whether VPEM can reduce the bias. We find that the closeness between the dominant eigenvector of a noisy state and the ideal quantum probe (without noise) with respect to an observable determines the reducible amount of bias by VPEM. Next, we show that one should carefully choose the reference point of the target parameter, which gives a smaller bias than others because the bias depends on the reference point. Otherwise, even if the dominant eigenvector and the ideal quantum probe are close, the bias of the mitigated case could be larger than the nonmitigated one. Finally, we analyze the error mitigation for a phase estimation scheme under various noises. Based on our analysis, we predict whether VPEM can effectively reduce a bias and numerically verify our results.

I. INTRODUCTION

Quantum metrology is a study of exploiting quantum properties to surpass the classical limit of estimation precision [1–6]. For various cases, quantum resources such as entanglement or squeezing enable us to achieve a better estimation precision that cannot be attained by classical means. Especially for a phase estimation scheme in an interferometer system, which is one of the most important tasks in quantum metrology, the quantum advantage using various nonclassical states, such as the N00N state [7–9] and squeezed state [9–13], have been studied and experimentally demonstrated [14–18].

To consider more practical situations, it is vital to study the effect of noise on estimation performance. Many studies investigated how the performance deteriorates due to noise, mostly focusing on statistical error [19,20]. In practice, however, it is not always feasible to obtain complete information about noise. Such a lack of knowledge of the noise in the system results in a bias of the estimator, which cannot be reduced simply by increasing the number of samples, unlike the statistical error. Especially, when there are sufficiently many samples for an estimation where the bias becomes dominant compared to the statistical error, the bias determines the order of magnitude of the total estimation error. In addition, while quantum error correction can recover the noisy state to its original state [21–28], it requires many resources, such as

midcircuit measurement, a low-noise threshold, and a larger number of required qubits because of encoding, which may not be practically accessible in the current era.

To address the above issues, Ref. [29] has recently proposed applying virtual-purification-based error mitigation (VPEM) [30,31] to quantum metrology to suppress the bias occurring from unknown noise and has shown that it considerably reduces the bias of an estimator of the phase estimation with a GHZ state. The principle behind VPEM is that it purifies the output state, which has undergone an unknown noise, to its dominant eigenstate, so that it effectively enables one to exploit the dominant eigenstate instead of the noisy state [30,31]. While Ref. [29] presented a framework of VPEM for quantum metrology and provided an example in which VPEM effectively reduces a bias, the general applicability of VPEM to quantum metrology has not been fully understood.

In this work we study when VPEM can effectively reduce the bias from unknown noise. In particular, we identify two crucial factors that determine whether VPEM can reduce the bias. The first one is the closeness of the dominant eigenvector of a noisy state and the (noiseless) ideal state with respect to an observable. Because VPEM allows us to exploit only the dominant eigenvector, not necessarily the ideal state, the dominant eigenvector has to be guaranteed to be sufficiently close to the ideal state. We find that the closeness between the expectation values of an observable $\hat{A}$ (used for estimation) of the ideal quantum probe and the dominant eigenvector of a noisy state determines how much bias can be reduced by VPEM. More specifically, the closeness decides the leading order of the reduced bias in the noise strength. In addition,

*h.jeong37@gmail.com
†liangjiang@uchicago.edu

we show that applying VPEM does not necessarily reduce the bias compared to without VPEM. The second one is the reference point of a parameter that we want to estimate. Assuming that the unknown parameter to be estimated is small (often called local parameter estimation [32,33]), the reference point is the one around which the unknown parameter varies.

In local parameter estimation, we show that one has to carefully choose the reference point that gives a smaller bias than others before applying VPEM to quantum parameter estimation. Otherwise, even if the expectation values of $\hat{A}$ over the ideal quantum probe and the dominant eigenvector are the same, the bias of the mitigated case could be larger than the nonmitigated case. We emphasize that considering a reference point is a unique feature of quantum metrology that has not been considered in the previous study [29].

To elaborate on the effect of the two aforementioned factors on VPEM, we apply our analysis to the phase estimation with the parity measurement in an interferometer system and numerically verify our analysis. We consider a N00N state and product of coherent and squeezed state as quantum probes for the phase estimation in the presence of phase diffusion, photon loss, and additive Gaussian noise. For the N00N state with phase diffusion or photon loss and for the product of a coherent and squeezed state in the presence of a special type of additive Gaussian noise, we find that the bias can be reduced by VPEM because the ideal quantum probe and the dominant eigenvector of the noisy state are the same, which is a similar case to the previous study [29]. In addition, we also find that the bias of the mitigation case could be larger than the nonmitigated case if one does not select a proper reference point. In contrast, for the product of coherent and squeezed state in the presence of photon loss, we find that it cannot benefit from VPEM, i.e., the magnitude of bias of the nonmitigated case and mitigated case is similar, even if we adopt a *good* reference point that gives smaller bias than others because the expectation values of the parity over the ideal quantum probe and the dominant eigenvector are not close enough. It indicates that VPEM does not always promise to reduce the bias error.

This paper is organized as follows: In Sec. II we explain how a bias occurs when one disregards unknown noise and introduce VPEM. In Sec. III we show the reducible amount of bias through VPEM. In particular, in Sec. III A we inspect the relation between the dominant eigenvector and the leading order of the bias with respect to noise strength. In Sec. III B we show how the bias depends on the reference point, highlighting the importance of the reference point when one applies VPEM to parameter estimation. In Sec. IV we apply our analysis to the phase estimation in the interferometer system.

II. MSE WITH BIAS

Before introducing our main results, let us explain a basic quantum estimation scheme and how a bias occurs in the presence of an unknown noise.

A. Ideal case

Let us consider a unitary process that encodes an unknown parameter ϕ. To estimate the parameter ϕ, a prepared quantum probe $\hat{\rho}_{\mathrm{id}}^{(0)} = |\psi_{\mathrm{id}}^{(0)}\rangle\langle\psi_{\mathrm{id}}^{(0)}|$, assumed to be a pure state, encodes the parameter by a unitary operation $\hat{\mathcal{U}}(\phi + \phi_0)$, which results in the ideal output state $\hat{\rho}_{\mathrm{id}}(\phi + \phi_0) = |\psi_{\mathrm{id}}(\phi + \phi_0)\rangle\langle\psi_{\mathrm{id}}(\phi + \phi_0)|$. Here ϕ_0 is a reference point of ϕ that one can freely select on purpose. One then measures the state in the eigenbasis of an observable $\hat{A}$ for N_s times, which renders the outcomes $\{A_{\mathrm{id},k}\}_{k=1}^{N_s}$, where $A_{\mathrm{id},k}$ is the kth measurement outcome. Meanwhile, the expectation value of $\hat{A}$ over $\hat{\rho}_{\mathrm{id}}$ is written as

$$\mathrm{Tr}[\hat{A}\hat{\rho}_{\mathrm{id}}(\phi + \phi_0)] = \langle\psi_{\mathrm{id}}(\phi + \phi_0)|\hat{A}|\psi_{\mathrm{id}}(\phi + \phi_0)\rangle \quad (1)$$

$$= x_{\mathrm{id}} + y_{\mathrm{id}}\phi + O(\phi^2), \quad (2)$$

where the expectation value is linearized for small ϕ with the coefficients x_{id} and y_{id}:

$$x_{\mathrm{id}}(\phi_0) \equiv \mathrm{Tr}[\hat{A}\hat{\rho}_{\mathrm{id}}(\phi_0)], \quad y_{\mathrm{id}}(\phi_0) \equiv \left.\frac{\partial\mathrm{Tr}[\hat{A}\hat{\rho}_{\mathrm{id}}(\phi_0 + \phi)]}{\partial\phi}\right|_{\phi=0}. \quad (3)$$

Thus, the following estimator is unbiased for small ϕ,

$$\phi_{\mathrm{id}}^{\mathrm{est}} = \frac{1}{y_{\mathrm{id}}}(\bar{A}_{\mathrm{id}} - x_{\mathrm{id}}), \quad (4)$$

where $\bar{A}_{\mathrm{id}}$ is the average of measurement outcomes $\sum_{k=1}^{N_s} A_{\mathrm{id},k}/N_s$ (see Fig. 1 for the schematic of the estimation).

The mean squared error (MSE) of the estimator is given by

$$\delta^2\phi_{\mathrm{id}} = \frac{1}{y_{\mathrm{id}}^2}\frac{\mathbb{V}[\hat{A}]_{\hat{\rho}_{\mathrm{id}}}}{N_s}, \quad (5)$$

which depends only on the statistical error that originates from the fluctuation of the average of the measurement outcomes due to the finite sample size. Here $\mathbb{V}[\hat{A}]_{\hat{\rho}} \equiv \mathrm{Tr}[\hat{A}^2\hat{\rho}] - \mathrm{Tr}[\hat{A}\hat{\rho}]^2$ is the variance of $\hat{A}$ over the quantum state. Therefore, the MSE can be reduced by increasing the number of samples N_s. It is worth emphasizing that the convergence rate is closely related to Fisher information [32,34], which is the main focus of many studies about quantum metrology [3,4]. Our interest lies in the bias rather than the statistical error, which will be introduced in the following section.

B. Error case

In a more practical scenario where noise occurs during estimation, the ideal state $\hat{\rho}_{\mathrm{id}}$ is replaced by an error state $\hat{\rho}_{\mathrm{e}}$, which depends on the noise. As the ideal case, one prepares N_s numbers of $\hat{\rho}_{\mathrm{e}}$ and obtains measurement outcomes $\{A_{\mathrm{e},k}\}_{k=1}^{N_s}$, where $A_{\mathrm{e},k}$ is the kth measurement outcome from $\hat{\rho}_{\mathrm{e}}$. When one does not know that the noise occurs and exploits the same theoretical value $\mathrm{Tr}[\hat{A}\hat{\rho}_{\mathrm{id}}(\phi + \phi_0)]$ as the ideal case, the estimator is then written as

$$\phi_{\mathrm{e}}^{\mathrm{est}} = \frac{1}{y_{\mathrm{id}}}(\bar{A}_{\mathrm{e}} - x_{\mathrm{id}}), \quad (6)$$

where $\bar{A}_{\mathrm{e}}$ is the average of measurement outcomes $\sum_{k=1}^{N_s} A_{\mathrm{e},k}/N_s$.

Due to this choice, the estimator $\phi_{\mathrm{e}}^{\mathrm{est}}$ generally has a bias [see Fig. 1(d) for the schematic of how bias occurs]

$$\langle\phi_{\mathrm{e}}^{\mathrm{est}}\rangle - \phi = \frac{x_{\mathrm{e}} - x_{\mathrm{id}} + (y_{\mathrm{e}} - y_{\mathrm{id}})\phi}{y_{\mathrm{id}}} + O(\phi^2) \quad (7)$$

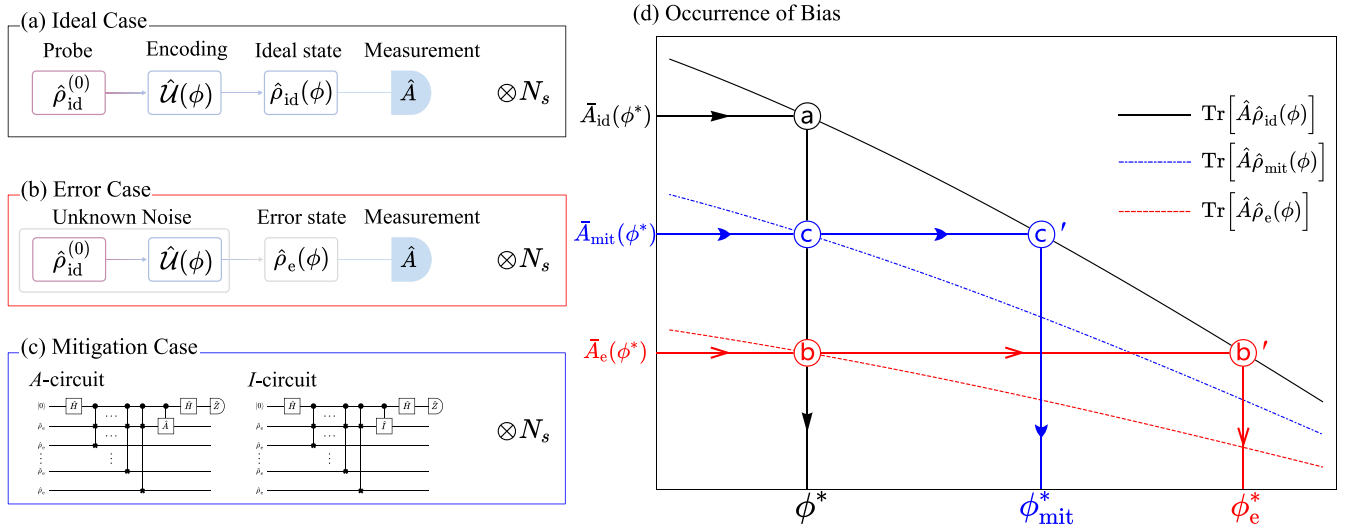$$\equiv B_{\mathrm{e}}(\phi, \phi_0, \Delta), \quad (8)$$

FIG. 1. Schematic diagram of quantum metrology and the effect of VPEM. We ignore the statistical error to focus on the bias such that measurement outcomes $\bar{A}$ is equal to $\text{Tr}[\hat{A}\hat{\rho}(\phi)]$ ($\bar{A}$ and $\hat{\rho}$ indicate $\hat{\rho}_{\text{id/e/mit}}$ and $\bar{A}_{\text{id/e/mit}}$) and set $\phi_0 = 0$ for simplicity. (a) The ϕ estimation process of the ideal case. Let us assume that encoded parameter ϕ is ϕ^* and one gets the average of measurement outcomes $\bar{A}_{\text{id}}(\phi^*)$. One then compares the average with the theoretical value $\text{Tr}[\hat{A}\hat{\rho}_{\text{id}}(\phi^*)]$ and estimates the encoded parameter as ϕ^*; thus there is no bias. [See ⓐ in (d).] (b) The ϕ estimation process in the presence of an unknown noise which renders measurement outcomes $\bar{A}_{\text{e}}(\phi^*)$. [See ⓑ in (d).] Since one still uses the ideal value $\text{Tr}[\hat{A}\hat{\rho}_{\text{id}}]$ for estimation, one estimates ϕ as ϕ_{e}^*, which satisfies $\bar{A}_{\text{e}}(\phi^*) = \text{Tr}[\hat{A}\hat{\rho}_{\text{id}}(\phi_{\text{e}}^*)]$, which results in a bias $\phi_{\text{e}}^* - \phi^*$. [See ⓑ′ in (d).] (c) VPEM effectively replaces the error state $\hat{\rho}_{\text{e}}(\phi)$ with $\hat{\rho}_{\text{mit}}(\phi)$. We assume that the expectation value of $\hat{A}$ over the mitigated state, and the ideal state is close enough, i.e., $\text{Tr}[\hat{A}\hat{\rho}_{\text{mit}}(\phi)] \approx \text{Tr}[\hat{A}\hat{\rho}_{\text{id}}(\phi)]$. Similar to the error case, one estimates the encoded parameter as ϕ_{mit}^* whose bias is $\phi_{\text{mit}}^* - \phi^*$, which is smaller than the error case. Here ϕ_{mit}^* satisfies $\bar{A}_{\text{mit}}(\phi^*) = \text{Tr}[\hat{A}\hat{\rho}_{\text{id}}(\phi_{\text{mit}}^*)]$. [See ⓒ and ⓒ′ in (d).]

where x_{e} and y_{e} are defined by a similar relation to the ideal case

$$\text{Tr}[\hat{A}\hat{\rho}_{\text{e}}(\phi + \phi_0)] = x_{\text{e}} + y_{\text{e}}\phi + O(\phi^2), \tag{9}$$

and $\langle \cdots \rangle$ is an average over all possible measurement outcomes of the quantum state. Suppose one can completely characterize the noise and $\hat{\rho}_{\text{e}}$, which is a typical assumption in many previous studies but may not always be feasible; one can then construct an unbiased estimator using x_{e} and y_{e} as the ideal case. Thus, the bias occurs from the lack of information about the noise. For our case, since $\phi_{\text{e}}^{\text{est}}$ is a biased estimator, the corresponding MSE $\delta^2\phi_{\text{e}}$, contains an additional term due to the bias:

$$\delta^2\phi_{\text{e}} = (B_{\text{e}})^2 + \frac{1}{N_s}\frac{\mathbb{V}[\hat{A}]_{\hat{\rho}_{\text{e}}}}{y_{\text{id}}^2}. \tag{10}$$

For the rest of the paper, we call $(B_{\text{e}})^2$ a bias error of error case. Figure 1 presents the schematic of the occurrence of the bias. In contrast to the statistical error, one cannot reduce the bias by simply increasing the number of samples N_s. We emphasize that especially when there are sufficiently many samples (i.e., large N_s), the bias is more critical than the statistical error, in terms that the MSE is dominated by the bias.

C. Error-mitigated quantum metrology

In this section we introduce the scheme of VPEM quantum metrology, recently proposed in Ref. [29]. The scheme employs $N_{s,\text{mit}}$ copies of a so-called A circuit and I circuit, each of which exploits n copies of an error state [30] (see Fig. 2). Although one might iterate each circuit for different
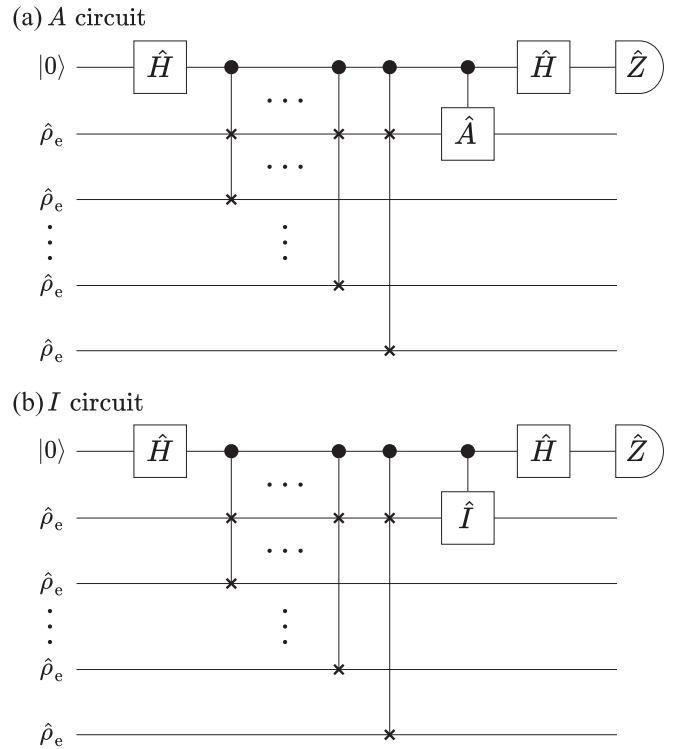

FIG. 2. Error mitigation circuits. After a Hadamard gate on the ancilla qubit, one applies a sequence of n controlled swap gates, followed by a controlled-$\hat{A}$ ($\hat{I}$) and another Hadamard gate, and measures in $\hat{Z}$ basis.

numbers, we equally set $N_{s,\text{mit}}$ as the number of iterations for each circuit for simplicity. The details of the dynamics are described in Refs. [29–31].

By implementing the circuits with an error state with its diagonal form being

$$\hat{\rho}_{\text{e}}(\phi + \phi_0) = \lambda |\psi\rangle\langle\psi| + (1 - \lambda) \sum_{k=1} p_k |\psi_k\rangle\langle\psi_k|, \quad (11)$$

one obtains measurement outcomes $\{z_k^A\}_{k=1}^{N_{s,\text{mit}}}$ and $\{z_k^I\}_{k=1}^{N_{s,\text{mit}}}$ from each circuit, where $z_k^A, z_k^I \in \{1, -1\}$ for all k's. Here $\sum_{k=1} p_k = 1$ and λ is the largest eigenvalue and $|\psi\rangle$ is the corresponding eigenvector, which may be different from the ideal state $|\psi_{\text{id}}\rangle$. Also, the relevant parameters in Eq. (11) are functions of $(\phi + \phi_0)$ and noise strength Δ. After an experiment, we obtain the average values as

$$\bar{Z}^A \equiv \frac{1}{N_{s,\text{mit}}} \sum_{k=1}^{N_{s,\text{mit}}} z_k^A, \quad \bar{Z}^I \equiv \frac{1}{N_{s,\text{mit}}} \sum_{k=1}^{N_{s,\text{mit}}} z_k^I, \quad (12)$$

whose averages over all possible measurement outcomes are given by [31]

$$\langle\bar{Z}^A\rangle = \lambda^n \langle\psi|\hat{A}|\psi\rangle + (1 - \lambda)^n \sum_{k=1} p_k^n \langle\psi_k|\hat{A}|\psi_k\rangle, \quad (13)$$

$$\langle\bar{Z}^I\rangle = \lambda^n + (1 - \lambda)^n \sum_{k=1} p_k^n. \quad (14)$$

One can find that the average of $\bar{Z}^A/\bar{Z}^I$ can be approximated by

$$\left\langle \frac{\bar{Z}^A}{\bar{Z}^I} \right\rangle \approx \frac{\text{Tr}[\hat{A}\hat{\rho}_{\text{e}}^n]}{\text{Tr}[\hat{\rho}_{\text{e}}^n]} \equiv \text{Tr}[\hat{A}\hat{\rho}_{\text{mit}}(\phi + \phi_0)] \quad (15)$$

$$= x_{\text{mit}} + y_{\text{mit}}\phi + O(\phi^2), \quad (16)$$

where $\hat{\rho}_{\text{e}}^n$ is the nth power of $\hat{\rho}_{\text{e}}$ and

$$\hat{\rho}_{\text{mit}} \equiv \lambda' |\psi\rangle\langle\psi| + (1 - \lambda') \sum_{k=1} p_k' |\psi_k\rangle\langle\psi_k|, \quad (17)$$

with

$$\lambda' \equiv \frac{1}{1 + \left(\frac{1-\lambda}{\lambda}\right)^n \sum_{k=1} p_k^n}, \quad p_k' \equiv \frac{p_k^n}{\sum_{l=1} p_l^n}. \quad (18)$$

Note that λ' is the dominant eigenvalue of a mitigated state $\hat{\rho}_{\text{mit}}$. We now set an estimator as

$$\phi_{\text{mit}}^{\text{est}} = \frac{1}{y_{\text{id}}}(\bar{A}_{\text{mit}} - x_{\text{id}}), \quad (19)$$

where $\bar{A}_{\text{mit}} \equiv \bar{Z}^A/\bar{Z}^I$, and we again used the ideal coefficients x_{id} and y_{id} in Eq. (3) because we do not know x_{mit} and y_{mit} for an unknown noise. $\phi_{\text{mit}}^{\text{est}}$ is then still a biased estimator, whose bias is given by (see Fig. 1)

$$\langle\phi_{\text{mit}}^{\text{est}}\rangle - \phi = \frac{x_{\text{mit}} - x_{\text{id}} + (y_{\text{mit}} - y_{\text{id}})\phi}{y_{\text{id}}} + O(\phi^2) \quad (20)$$

$$\equiv B_{\text{mit}}(\phi, \phi_0, \Delta), \quad (21)$$

and the corresponding MSE is [31]

$$\delta^2\phi_{\text{mit}} \approx (B_{\text{mit}})^2 \quad (22)$$

$$+ \frac{1}{N_{s,\text{mit}}y_{\text{id}}^2} \left[\frac{1 - \text{Tr}[\hat{A}\hat{\rho}_{\text{e}}^n]^2}{\text{Tr}[\hat{\rho}_{\text{e}}^n]^2} + \frac{\text{Tr}[\hat{A}\hat{\rho}_{\text{e}}^n]^2 (1 - \text{Tr}[\hat{\rho}_{\text{e}}^n]^2)}{\text{Tr}[\hat{\rho}_{\text{e}}^n]^4} \right]. \quad (23)$$

In Eq. (23) the first term $(B_{\text{mit}})^2$ is the bias error and the second term is the statistical error. We consider $2nN_{s,\text{mit}}$ number of samples for the mitigation case. For a fair comparison between the error case and the mitigation case, we will exploit the same number of samples for an estimation, such that $N_s = 2nN_{s,\text{mit}}$.

To summarize, the effect of VPEM is to replace the error state $\hat{\rho}_{\text{e}}$ with the purified state $\hat{\rho}_{\text{mit}} = \hat{\rho}_{\text{e}}^n/\text{Tr}[\hat{\rho}_{\text{e}}^n]$ for the order n. Since λ is the largest eigenvalue, the λ^n term becomes dominant. In contrast, the coefficient $(1 - \lambda)^n p_k^n$, associated with the orthogonal subspace to the dominant eigenvector, is suppressed as n increases. Therefore, the mitigated state $\hat{\rho}_{\text{mit}}$ in Eq. (17) converges to the dominant eigenvector $|\psi\rangle\langle\psi|$, which results in $\text{Tr}[\hat{A}\hat{\rho}_{\text{mit}}] \approx \langle\psi|\hat{A}|\psi\rangle$. Hence, if y_{id} is not too small and the dominant eigenvector $|\psi\rangle$ and the ideal state $|\psi_{\text{id}}\rangle$ are close enough, such that $(\langle\psi|\hat{A}|\psi\rangle - \langle\psi_{\text{id}}|\hat{A}|\psi_{\text{id}}\rangle)/y_{\text{id}} \approx 0$, the bias error of the mitigation case $(B_{\text{mit}})^2$ would be smaller than that of the error case $(B_{\text{e}})^2$. Figure 1 illustrates how VPEM reduces the bias.

III. CRUCIAL FACTORS FOR REDUCING BIAS VIA VPEM

In the previous section, we observed that VPEM effectively purifies an error state to its dominant eigenvector. In this section we further analyze the efficiency of VPEM concerning the distance between the dominant eigenvector and the ideal state with respect to an observable. Furthermore, we show that when VPEM is applied to quantum metrology, an additional factor must be necessarily considered, which is the reference point.

A. Expectation value of $\hat{A}$ over the dominant eigenvector

We analyze how the difference of the expectation values of the observable $\hat{A}$ over the dominant eigenvector of $\hat{\rho}_{\text{e}}$ and the ideal state affects the efficacy of VPEM. To this end, we find an expression of a bias in the error case:

$$B_{\text{e}}(\phi, \phi_0, \Delta) = \frac{x_{\text{e}}(\phi_0) - x_{\text{id}}(\phi_0) + \phi[y_{\text{e}}(\phi_0) - y_{\text{id}}(\phi_0)]}{y_{\text{id}}(\phi_0)} \quad (24)$$

$$= \frac{1}{y_{\text{id}}(\phi_0)} \sum_{k=1}^{n-1} \left[f_k(\phi_0) + \phi \frac{\partial f_k(\phi)}{\partial \phi}\bigg|_{\phi=\phi_0} \right] \Delta^k + O(\Delta^n), \quad (25)$$

where we defined

$$f_1(\phi) \equiv a_1(\phi) - \lambda_1(\phi)b_0(\phi), \quad (26)$$

$$f_k(\phi) \equiv a_k(\phi) + \sum_{l=1}^{k-1} \lambda_l(\phi)a_{k-l}(\phi) - \sum_{l=1}^{k} \lambda_l(\phi)b_{k-l}(\phi), \quad (27)$$

for $k \geqslant 2$. Here λ_k, a_k, and b_k are coefficients from the following expansions:

$$\lambda = 1 - \sum_{k=1}^{\infty} \lambda_k(\phi + \phi_0)\Delta^k, \qquad (28)$$

$$\langle\psi|\hat{A}|\psi\rangle - \langle\psi_{\mathrm{id}}|\hat{A}|\psi_{\mathrm{id}}\rangle = \sum_{k=1}^{\infty} a_k(\phi + \phi_0)\Delta^k, \qquad (29)$$

$$\sum_{k=1} p_k\langle\psi_k|\hat{A}|\psi_k\rangle - \langle\psi_{\mathrm{id}}|\hat{A}|\psi_{\mathrm{id}}\rangle = \sum_{k=0}^{\infty} b_k(\phi + \phi_0)\Delta^k. \qquad (30)$$

Equation (25) shows that the bias of error case consists of three distinct terms: (1) the difference between the expectation value of $\hat{A}$ over the dominant eigenvector and the ideal state $\langle\psi|\hat{A}|\psi\rangle - \langle\psi_{\mathrm{id}}|\hat{A}|\psi_{\mathrm{id}}\rangle$, characterized by a_k, (2) the difference between the dominant eigenvalue λ and 1, identified by λ_k, and (3) the difference between the expectation value of $\hat{A}$ over the tail terms $\sum_{k=1} p_k\langle\psi_k|\hat{A}|\psi_k\rangle - \langle\psi_{\mathrm{id}}|\hat{A}|\psi_{\mathrm{id}}\rangle$, expressed by b_k. After applying VPEM, the bias becomes

$$B_{\mathrm{mit}}(\phi, \phi_0, \Delta)$$
$$= \frac{x_{\mathrm{mit}}(\phi_0) - x_{\mathrm{id}}(\phi_0) + \phi[y_{\mathrm{mit}}(\phi_0) - y_{\mathrm{id}}(\phi_0)]}{y_{\mathrm{id}}(\phi_0)} \qquad (31)$$

$$= \frac{1}{y_{\mathrm{id}}(\phi_0)} \sum_{k=1}^{n-1} \left[a_k(\phi_0) + \phi\frac{\partial a_k(\phi)}{\partial\phi}\bigg|_{\phi=\phi_0} \right]\Delta^k + O(\Delta^n). \qquad (32)$$

One can easily show Eq. (32) using the definition of mitigated state $\hat{\rho}_{\mathrm{mit}}$ in Eq. (17) and the expansions in Eqs. (29) and (30).

A crucial difference from the error case is that the bias B_{mit} up to the order of Δ^{n-1} consists only of a_k which comes from the difference between the expectation value of $\hat{A}$ over the dominant eigenvector and the ideal state $\langle\psi|\hat{A}|\psi\rangle - \langle\psi_{\mathrm{id}}|\hat{A}|\psi_{\mathrm{id}}\rangle$. Therefore, $\langle\psi|\hat{A}|\psi\rangle - \langle\psi_{\mathrm{id}}|\hat{A}|\psi_{\mathrm{id}}\rangle$ dictates the leading order in Δ of the bias while other contributions from b_k and λ_k are suppressed by VPEM. As a consequence, if $\langle\psi|\hat{A}|\psi\rangle - \langle\psi_{\mathrm{id}}|\hat{A}|\psi_{\mathrm{id}}\rangle = 0$, $B_{\mathrm{mit}} = O(\Delta^n)$ while $B_{\mathrm{e}} = O(\Delta)$, which clearly shows that in the small regime of Δ, the bias can be reduced by applying VPEM. However, if the dominant eigenvector and the ideal state are not close enough, especially when $a_1(\phi_0) \neq 0$, we emphasize that VPEM cannot even guarantee the constant factor reduction of the bias even for small noise. More specifically, for the case when $a_1(\phi_0) \neq 0$, the leading orders of the biases are given by

$$|B_{\mathrm{e}}| \approx \left| \frac{a_1(\phi_0) - \lambda_1(\phi)b_0(\phi_0)}{y_{\mathrm{id}}(\phi_0)} \right|\Delta, \qquad (33)$$

$$|B_{\mathrm{mit}}| \approx \left| \frac{a_1(\phi_0)}{y_{\mathrm{id}}(\phi_0)} \right|\Delta. \qquad (34)$$

Note that since λ is the eigenvalue of the error state, it should be less than 1 for any given Δ, which means the λ_1 in Eq. (28) should be positive for small Δ. In contrast, we emphasize that a_1 and b_0, in Eqs. (29) and (30), can be either positive or negative; as a result, there could be a case where $|B_{\mathrm{e}}| < |B_{\mathrm{mit}}|$. Especially, in Sec. IV B, we study the case where the magnitude of a bias increases even if one applies VPEM.

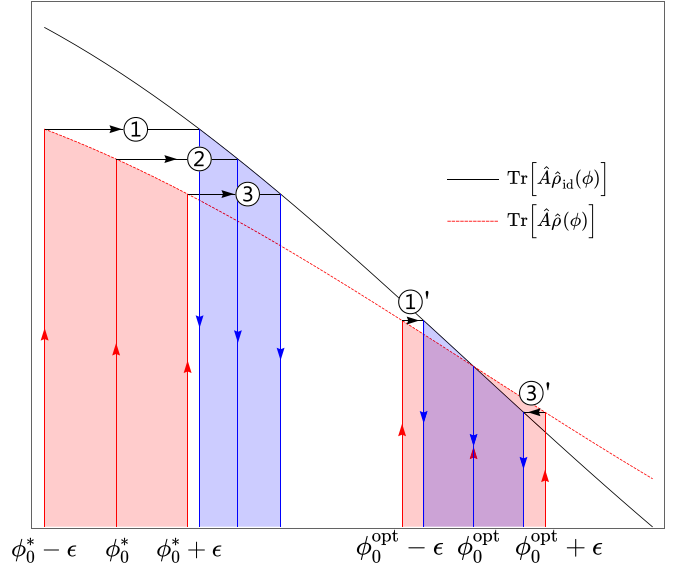We note that there have been analyses about the efficiency of VPEM [30,31,35]. However, their analysis may not be

FIG. 3. Schematic diagram of how the reference point affects the bias. Here $\hat{\rho}$ indicates $\hat{\rho}_{\mathrm{e/mit}}$. We consider local parameter estimation where the parameter is in the range $[-\epsilon, \epsilon]$. The left red and blue regions correspond to the estimation when the reference point is ϕ_0^*, which is not an optimal reference point. The three horizontal black right-headed arrows ①, ②, and ③ are the bias when the true parameter ϕ is $-\epsilon$, 0, and ϵ each. In contrast, the right red and blue regions correspond to the estimation when the optimal reference ϕ_0^{opt} is chosen. When the parameter is 0, there is no bias. The two horizontal black arrows ①′ and ③′ are the bias when the parameter is $-\epsilon$ and ϵ each. One can find that ①′ and ③′ are shorter than ①, ②, and ③, which shows reduced bias by optimal reference point.

directly applicable to VPEM-assisted quantum metrology. In Refs. [30,31,35] they showed that the efficacy of VPEM is often guaranteed in practice, mainly focusing on random states, while quantum metrology often exploits highly structured states.

B. Reference point of the parameter

We now introduce another important factor that plays a crucial role in VPEM for quantum metrology, which is the reference point. Recall that, according to Eqs. (8) and (21), the bias of both error case and mitigation case can be expressed as

$$B(\phi, \phi_0, \Delta) \approx \frac{x(\phi_0) - x_{\mathrm{id}}(\phi_0) + [y(\phi_0) - y_{\mathrm{id}}(\phi_0)]\phi}{y_{\mathrm{id}}(\phi_0)}. \qquad (35)$$

Here $B(\phi, \phi_0, \Delta)$, x, y, and $\hat{\rho}$ stand for $B_{\mathrm{e/mit}}$, $x_{\mathrm{e/mit}}$, $y_{\mathrm{e/mit}}$, and $\hat{\rho}_{\mathrm{e/mit}}$. We note that reference points for a mitigation case and an error case could be different. Because both biases depend on the reference point ϕ_0, one might be able to reduce the bias by simply controlling ϕ_0. In other words, there might be regions or points of ϕ_0 that give a smaller bias than others (see Fig. 3). Therefore, one should consider a reference point when adopting error mitigation to quantum metrology, unlike expectation value estimation [30,31]. Otherwise, even if $\langle\psi|\hat{A}|\psi\rangle = \langle\psi_{\mathrm{id}}|\hat{A}|\psi_{\mathrm{id}}\rangle$, $(B_{\mathrm{mit}})^2$ can be larger than $(B_{\mathrm{e}})^2$ because of an inadequate reference point instead of the failure of error mitigation. We will show corresponding examples

in Sec. IV. Here we suggest several ways to choose a *good* reference point that gives a smaller bias than others to study the importance of a reference point. The reference points that we suggest below are examples that effectively show the importance of a reference point when one applies VPEM to quantum metrology.

First, we can consider a reference point that minimizes the zeroth order of ϕ of a bias. Note that Eq. (35) consists of the zeroth order of ϕ which is $(x - x_{\mathrm{id}})/y_{\mathrm{id}}$ and the first order of ϕ which is $(y - y_{\mathrm{id}})/y_{\mathrm{id}}$. Since we are considering small ϕ, one can alleviate the bias by choosing the reference point ϕ_0 such that

$$\phi_0^{\mathrm{opt}}(\Delta_0) = \arg \min_{\phi_0} \left[\frac{x(\phi_0) - x_{\mathrm{id}}(\phi_0)}{y_{\mathrm{id}}(\phi_0)} \right]^2. \qquad (36)$$

Especially, if there exists ϕ_0^{opt} that makes $(x(\phi_0^{\mathrm{opt}}) - x_{\mathrm{id}}(\phi_0^{\mathrm{opt}}))/y_{\mathrm{id}}(\phi_0^{\mathrm{opt}}) = 0$, the zeroth order of ϕ in Eq. (35) vanishes and the bias contains only ϕ order; equivalently, the bias error has only the ϕ^2 order term. When the $\phi_0^{\mathrm{opt}}(\Delta_0)$ in Eq. (36) does not depend on Δ, one can choose $\phi_0^{\mathrm{opt}}(\Delta_0)$ as a reference point regardless of noise strength. However, considering more general cases, the $\phi_0^{\mathrm{opt}}(\Delta_0)$ depends on noise strength. In that case, a possible and practical choice of a *good* reference point is to consider a reference point that minimizes the average of zeroth order of the bias over the noise strength using some prior knowledge, in such a way that

$$\phi_0^{\mathrm{opt}} = \arg \min_{\phi_0} \int_{\Delta_1}^{\Delta_2} p(\Delta) \left[\frac{x(\phi_0) - x_{\mathrm{id}}(\phi_0)}{y_{\mathrm{id}}(\phi_0)} \right]^2 d\Delta, \qquad (37)$$

where $p(\Delta)$ is a prior distribution of noise strength Δ. For the remainder of this paper, we call both reference points that satisfy Eqs. (36) and (37) an optimal reference point (optimal in terms that it minimizes the zeroth order of ϕ) without confusion; all the examples we suggest clearly show which reference point that we choose. We denote the optimal reference point for an error case as $\phi_{0,\mathrm{e}}^{\mathrm{opt}}$ and a mitigation case as $\phi_{0,\mathrm{mit}}^{\mathrm{opt}}$.

Last, we present three different cases and compare B_{e} and B_{mit} when one chooses the optimal reference points for each case. For simplicity, we assume that there exist optimal reference points $\phi_{0,\mathrm{e}}^{\mathrm{opt}}$ and $\phi_{0,\mathrm{mit}}^{\mathrm{opt}}$ which make the zeroth order of bias in ϕ vanish.

Case 1. The dominant eigenvector is the same as the ideal state, which renders $\langle \psi | \hat{A} | \psi \rangle - \langle \psi_{\mathrm{id}} | \hat{A} | \psi_{\mathrm{id}} \rangle = 0$. In this case, if one applies VPEM, the leading order of a bias of mitigation case B_{mit} in the noise strength is Δ^n. In addition, by choosing $\phi_{0,\mathrm{mit}}^{\mathrm{opt}}$ as a reference point, B_{mit} can be reduced up to $O(\phi \Delta^n)$, which shows the efficacy of VPEM. This case corresponds to the phase estimation in the interferometer system with the N00N state in the presence of phase diffusion or photon loss, studied in Sec. IV A, and with the product of the coherent and squeezed state as a quantum probe in the presence of additive Gaussian noise, which is studied in Sec. IV B.

Case 2. The dominant eigenvector is not close enough to the ideal state such that $\langle \psi | \hat{A} | \psi \rangle - \langle \psi_{\mathrm{id}} | \hat{A} | \psi_{\mathrm{id}} \rangle = O(\phi \Delta)$. In this case B_{e} and B_{mit} have the same order of bias $O(\phi \Delta)$, i.e., VPEM cannot effectively reduce the bias. The phase estimation exploiting the product of a coherent and squeezed
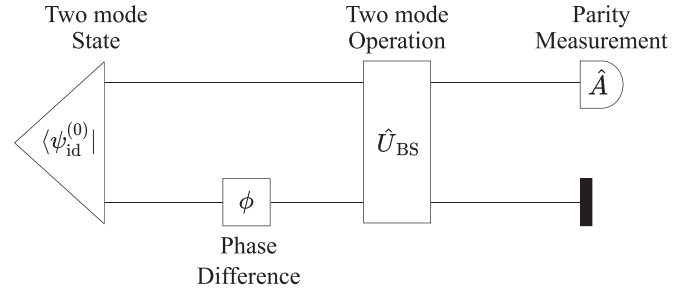

FIG. 4. Interferometer system. See the discussion for more details.

state which suffers from photon loss corresponds to Case 2, which is investigated in Sec. IV B.

Case 3. $\langle \psi | \hat{A} | \psi \rangle - \langle \psi_{\mathrm{id}} | \hat{A} | \psi_{\mathrm{id}} \rangle = O(\phi \Delta^{n'})$ for $n' \geqslant 2$. In this case a bias cannot be suppressed beyond $O(\phi \Delta^{n'})$, no matter how large the order of VPEM n is. Hence, choosing $n > n'$ does not further reduce the leading order of Δ in the bias.

A previous study of phase estimation exploiting the GHZ state in the presence of amplitude-damping noise [29] can be classified as Case 1, where the dominant eigenvector is the same as the ideal state; therefore, the bias of mitigation case B_{mit} is the order of Δ^n, where one can benefit from error mitigation. However, we emphasize that this is a special case of our analysis, and we will investigate other cases in the following sections. In addition, we also note that there is a study that is relevant to our Case 3. In Ref. [36] they study a case that shows increasing mitigation order higher than $n = 2$ does not offer a further enhancement.

IV. NUMERICAL SIMULATION OF VPEM FOR PHASE ESTIMATION SCHEME

In this section we apply VPEM to phase estimation in an interferometer, which is one of the most extensively studied quantum metrological tasks. We exploit a N00N state or the product of a coherent and squeezed state as a quantum probe, which are well-known quantum states that enable quantum-enhanced estimation [8,11]. Based on our analytic study in Sec. III, we inspect whether the phase estimation scheme can benefit from VPEM in the presence of various types of noises. Let us recall a phase estimation in an interferometer system, which is illustrated in Fig. 4. First, one encodes an unknown phase ϕ by a phase shift operator $\hat{\Phi}(\phi) = e^{i\hat{n}_2 \phi}$ onto a prepared two-mode quantum state $|\psi_{\mathrm{id}}^{(0)}\rangle$, where $\hat{n}_i$ is the number operator in the ith mode. After a two-mode balanced beam splitter $\hat{U}_{\mathrm{BS}}$, which renders an ideal output state $|\psi_{\mathrm{id}}(\phi + \phi_0)\rangle \equiv \hat{U}_{\mathrm{BS}} \hat{\Phi}(\phi + \phi_0) |\psi_{\mathrm{id}}^{(0)}\rangle$, one measures the state and estimates the ϕ using the measurement outcomes. Here we choose the parity operator $\hat{A} = (-1)^{\hat{n}_1}$ as a measurement operator. This is a crucial property to apply the error mitigation scheme because the parity operator is both unitary and Hermitian.

In addition, the parity measurement is known to render Heisenberg scaling precision with appropriate input states [37–41].

We will consider three kinds of noise: phase diffusion, photon loss, and additive Gaussian noise, which are dominant noises in phase estimation.

Phase diffusion happens when a random phase x occurs during a phase ϕ encoding. Assuming that x follows the Gaussian distribution, the ideal state then transforms as

$$\hat{\rho}_{\text{id}} \rightarrow \hat{\rho}_{\text{e}} = \frac{1}{\sqrt{2\pi\Delta}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\Delta}} \hat{\rho}_{\text{id}}(\phi_0 + \phi + x) dx, \quad (38)$$

where Δ is a noise strength that characterizes the degree of the phase diffusion.

Photon loss transforms an annihilation operator as $\hat{a} \rightarrow \sqrt{\eta}\hat{a} + \sqrt{1-\eta}\hat{e}$, where $\hat{a}$ and $\hat{e}$ correspond to the annihilation operator for the system and the environment, respectively. The loss rate $0 \leqslant (1-\eta) \equiv \Delta \leqslant 1$ corresponds to noise strength. Since a photon-loss channel occurring on both modes with the same strength commutes with beam splitter and phase shift operation, our analysis covers a photon loss occurring from state preparation to before the measurement.

Finally, additive Gaussian noise is that a random displacement $\vec{x} = (x, p)$ occurs following Gaussian distribution, which transforms a state $\hat{\rho}$ as

$$\frac{1}{2\pi} \frac{1}{\sqrt{\Delta_x \Delta_p}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\Delta_x}} e^{-\frac{p^2}{2\Delta_p}} \hat{D}(\vec{x}) \hat{\rho} \hat{D}^{\dagger}(\vec{x}) \, dx \, dp, \quad (39)$$

where $\hat{D}(\vec{x}) = e^{(x+ip)\hat{a}^{\dagger} - (x-ip)\hat{a}}$ is a displacement operator. Here Δ_x and Δ_p are noise strength. We consider an additive Gaussian noise occurring on the second mode of the initial quantum state $|\psi_{\text{id}}^{(0)}\rangle$.

A. N00N state

1. Ideal case

It is known that a N00N state is one of the representative quantum probes which enables a quantum-enhanced phase estimation [7–9]. In the ideal case, one exploits $|\psi_{\text{id}}^{(0)}\rangle = (|N\rangle_1 |0\rangle_2 + |0\rangle_1 |N\rangle_2)/\sqrt{2}$ as a quantum probe, and the expectation value of parity over $|\psi_{\text{id}}(\phi + \phi_0)\rangle = \hat{U}_{\text{BS}} \hat{\Phi}(\phi + \phi_0)|\psi_{\text{id}}^{(0)}\rangle$ is

$$\text{Tr}[\hat{A}\hat{\rho}_{\text{id}}] = \frac{(-i)^N e^{iN(\phi+\phi_0)} + i^N e^{-iN(\phi+\phi_0)}}{2}. \quad (40)$$

Therefore, $\text{Tr}[\hat{A}\hat{\rho}_{\text{id}}]$ is either $\sin N(\phi + \phi_0)$ or $\cos N(\phi + \phi_0)$ with different signs. Since we assume that we can control the reference point, as a consequence we can always find ϕ_0' that satisfies $\sin N(\phi + \phi_0) = \cos N(\phi + \phi_0')$. Thus, without loss of generality, we consider it as $\sin N(\phi + \phi_0)$. The MSE for the ideal case is given by

$$\delta^2 \phi_{\text{id}} = \frac{1}{y_{\text{id}}^2} \frac{\mathbb{V}[\hat{A}]_{\hat{\rho}_{\text{id}}}}{N_s} = \frac{1}{N_s N^2}, \quad (41)$$

which shows the Heisenberg scaling.

2. Phase diffusion

In the presence of phase diffusion, the ideal state transforms to an error state [which can be found using

Eq. (38)],

$$\hat{\rho}_{\text{e}} = \left(\frac{1 + e^{-\frac{\Delta}{2}N^2}}{2}\right)|\psi_{\text{id}}\rangle\langle\psi_{\text{id}}| + \left(\frac{1 - e^{-\frac{\Delta}{2}N^2}}{2}\right)|\psi_{\perp}\rangle\langle\psi_{\perp}| \quad (42)$$

$$\equiv \lambda|\psi_{\text{id}}\rangle\langle\psi_{\text{id}}| + \lambda_{\perp}|\psi_{\perp}\rangle\langle\psi_{\perp}|, \quad (43)$$

where $|\psi_{\text{id}}\rangle = |\psi_{\text{id}}(\phi + \phi_0)\rangle$ and $|\psi_{\perp}\rangle \equiv \hat{U}_{\text{BS}} \hat{\Phi}(\phi + \phi_0)[(|N\rangle_1|0\rangle_2 - |0\rangle_1|N\rangle_2)/\sqrt{2}]$, which is orthogonal to $|\psi_{\text{id}}\rangle$. In this case the dominant eigenvector is the same as the ideal state, which implies that $\langle\psi|\hat{A}|\psi\rangle = \langle\psi_{\text{id}}|\hat{A}|\psi_{\text{id}}\rangle$. Therefore, according to the analysis in Sec. III, one can expect that VPEM can effectively reduce the bias error. Let us demonstrate this by inspecting the expectation values of $\hat{A}$:

$$\text{Tr}[\hat{A}\hat{\rho}_{\text{id}}] = \sin N(\phi + \phi_0) \quad (44)$$

$$= \sin N\phi_0 + (N\cos N\phi_0)\phi + O(\phi^2), \quad (45)$$

$$\text{Tr}[\hat{A}\hat{\rho}_{\text{e}}] = (\lambda - \lambda_{\perp})\sin N(\phi + \phi_0) \quad (46)$$

$$= (\lambda - \lambda_{\perp})[\sin N\phi_0 + (N\cos N\phi_0)\phi] + O(\phi^2), \quad (47)$$

$$\text{Tr}[\hat{A}\hat{\rho}_{\text{mit}}] = \left(\frac{\lambda^n - \lambda_{\perp}^n}{\lambda^n + \lambda_{\perp}^n}\right)\sin N(\phi + \phi_0) \quad (48)$$

$$= \left(\frac{\lambda^n - \lambda_{\perp}^n}{\lambda^n + \lambda_{\perp}^n}\right)[\sin N\phi_0 + (N\cos N\phi_0)\phi] + O(\phi^2). \quad (49)$$

First, we find that $\text{Tr}[\hat{A}\hat{\rho}_{\text{id}}(0)] = \text{Tr}[\hat{A}\hat{\rho}_{\text{e}}(0)] = \text{Tr}[\hat{A}\hat{\rho}_{\text{mit}}(0)]$, and $\arg\max_{\phi_0} y_{\text{id}}(\phi_0) = 0$, which show that the optimal reference points satisfying Eq. (36) are $\phi_{0,\text{e}}^{\text{opt}} = \phi_{0,\text{mit}}^{\text{opt}} = 0$, for both error and mitigation cases regardless of the noise strength Δ and mitigation order n. Second, once we choose the optimal reference point, the biases are given by

$$B_{\text{e}} = \left(-\frac{N^2}{2}\Delta\right)\phi + O(\Delta^2), \quad (50)$$

$$B_{\text{mit}} = -2\left(\frac{N^2}{4}\Delta\right)^n \phi + O(\Delta^{n+1}), \quad (51)$$

which shows that VPEM reduces the bias in the small range of Δ that satisfies $\frac{N^2}{4}\Delta \gg (\frac{N^2}{4}\Delta)^n$.

We support our results with numerical simulations. Figure 5 exhibits a simulation result of bias errors with the N00N state in the presence of phase diffusion. We show that under the optimal reference point, VPEM can always effectively reduce the bias regardless of noise strength and the value of ϕ. We find that $B_{\text{e}}(\phi, \phi_{0,\text{e}}^{\text{opt}}, \Delta)^2 \geqslant B_{\text{mit}}(\phi, \phi_{0,\text{mit}}^{\text{opt}}, \Delta, n = 2)^2 \geqslant B_{\text{mit}}(\phi, \phi_{0,\text{mit}}^{\text{opt}}, \Delta, n = 3)^2$ as we expected through our analysis since the dominant eigenvector is equal to the ideal state. Here we emphasize that since $\phi_{0,\text{e}}^{\text{opt}} = 0$ and $\phi_{0,\text{mit}}^{\text{opt}} = 0$, both biases disappear at $\phi = 0$. In addition, Fig. 5 also shows the importance of the reference point. It shows that if one does not carefully choose a reference point, an error case could have a smaller bias than a mitigation case even though $\langle\psi|\hat{A}|\psi\rangle = \langle\psi_{\text{id}}|\hat{A}|\psi_{\text{id}}\rangle$. We find that $B_{\text{e}}(\phi, \phi_{0,\text{e}}^{\text{opt}}, \Delta)^2 \leqslant B_{\text{mit}}(\phi, \frac{\pi}{12}, \Delta, n = 3)^2 \leqslant B_{\text{mit}}(\phi, \frac{\pi}{12}, \Delta, n = 2)^2$ when the dominant eigenvalue in the
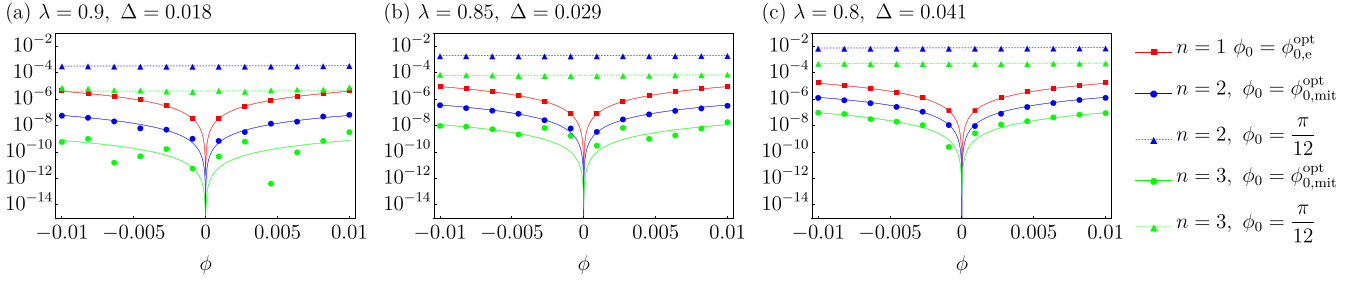
FIG. 5. (a)–(c) Bias errors (with log scale) of the ϕ estimation where $\phi \in [-0.01, 0.01]$, exploiting the N00N state ($N = 5$) as a quantum probe in the presence of phase diffusion with different noise strengths (namely, different dominant eigenvalues). Solid and dashed lines are theoretical values of bias errors. Circle, square, and up-triangle denote squares of the difference between estimated value and real value from simulations, which are $(\phi^{\text{est}} - \phi)^2$ where ϕ^{est}'s are estimators of ϕ [defined in Eqs. (6) and (19)] with $N_s = 10^9$ number of samples. We emphasize that since we exploit a large number of samples which reduces fluctuations of ϕ^{est}'s, $(\phi^{\text{est}} - \phi)^2$'s can be considered as bias errors from simulation. Here we denote the error case as $n = 1$ which corresponds to the red solid lines and red squares.

range $0.8 \leqslant \lambda \leqslant 0.9$, regardless of ϕ. We emphasize that $\pi/12$ is not an optimal reference point.

3. Photon loss

In the presence of the photon loss, a single-mode number state $|N\rangle\langle N|$ transforms to $\sum_{k=0}^{N}\binom{N}{k}\eta^k(1-\eta)^{N-k}|k\rangle\langle k|$, and the component $|N\rangle\langle 0|$ becomes $\sqrt{\eta}^N|N\rangle\langle 0|$, where $(1-\eta)$ is the loss rate and we denote it as noise strength Δ. Since we assume that photon loss occurs on each mode independently with the same noise strength and the loss channel commutes with the beam splitter and phase shift operations, one can easily find that the error state is

$$\hat{\rho}_{\text{e}} = \eta^N|\psi_{\text{id}}\rangle\langle\psi_{\text{id}}| + (1-\eta^N)\sum_{k=1}^{2N}p_k|\psi_k\rangle\langle\psi_k|. \tag{52}$$

Here notice that the dominant eigenvector is equal to the ideal state. In addition,

$$p_k = p_{N+k} = \binom{N}{k-1}\left(\frac{\eta^{k-1}(1-\eta)^{N-k+1}}{2(1-\eta^N)}\right), \tag{53}$$

$$|\psi_k\rangle = \hat{U}_{\text{BS}}\hat{\Phi}(\phi+\phi_0)|k-1,0\rangle_{1,2}, \tag{54}$$

$$|\psi_{k+N}\rangle = \hat{U}_{\text{BS}}\hat{\Phi}(\phi+\phi_0)|0,k-1\rangle_{1,2}, \tag{55}$$

for $k \in \{1, 2, \ldots, N\}$. Note that $\langle\psi_k|\hat{A}|\psi_k\rangle = 0$ for all k. The corresponding expectation values are given by

$$\text{Tr}[\hat{A}\hat{\rho}_{\text{id}}] = \sin N(\phi+\phi_0) \tag{56}$$

$$= \sin N\phi_0 + (N\cos N\phi_0)\phi + O(\phi^2), \tag{57}$$

$$\text{Tr}[\hat{A}\hat{\rho}_{\text{e}}] = (1-\Delta)^N\sin N(\phi+\phi_0) \tag{58}$$

$$= (1-\Delta)^N[\sin N\phi_0 + (N\cos N\phi_0)\phi] + O(\phi^2), \tag{59}$$

$$\text{Tr}[\hat{A}\hat{\rho}_{\text{mit}}] = \left(\frac{(1-\Delta)^{nN}}{\sum_{k=0}^{N}\binom{N}{k}^n(1-\Delta)^{nk}\Delta^{n(N-k)}}\right)\sin N(\phi+\phi_0) \tag{60}$$

$$= \left(\frac{(1-\Delta)^{nN}}{\sum_{k=0}^{N}\binom{N}{k}^n(1-\Delta)^{nk}\Delta^{n(N-k)}}\right) \times [\sin N\phi_0 + (N\cos N\phi_0)\phi] + O(\phi^2). \tag{61}$$

Again, one can find that the optimal reference points satisfying Eq. (36) are $\phi_{0,\text{e}}^{\text{opt}} = \phi_{0,\text{mit}}^{\text{opt}} = 0$. and the corresponding biases are

$$B_{\text{e}} = (-N\Delta)\phi + O(\Delta^2), \tag{62}$$

$$B_{\text{mit}} = -(N\Delta)^n\phi + O(\Delta^3), \tag{63}$$

which clearly shows that the bias is reduced by error mitigation. We also numerically demonstrate that error mitigation can reduce bias in the presence of photon loss and the importance of a reference point, which is shown in Fig. 6. Similarly to the previous case, the VPEM reduces the bias error significantly because the dominant eigenvector is the same as the ideal state. However, if one does not carefully choose a reference point, the bias error of the mitigation case can be larger than the error case.

B. Product of coherent and squeezed states

1. Ideal case

We prepare a coherent state $|\alpha_0\rangle_1$ and a squeezed vacuum state $|r, 0\rangle_2$, and inject the states into a beam splitter which makes the entanglement between the modes. We exploit the state $\hat{U}_{\text{BS}}|\alpha_0\rangle_1|r, 0\rangle_2$ as a quantum probe $|\psi_{\text{id}}^{(0)}\rangle$ which is a two-mode input state in Fig. 4, the ideal state is then $|\psi_{\text{id}}\rangle = \hat{U}_{\text{BS}}\hat{\Phi}(\phi+\phi_0)|\psi_{\text{id}}^{(0)}\rangle$. For the product of coherent state and squeezed state case, we study the effectiveness of error mitigation in the presence of photon loss and additive Gaussian noise. We consider $N_c = N_r = 2.5$. Unlike the previous case, the optimal reference point varies with the noise strength Δ. To choose an appropriate reference point, we set $p(\Delta)$ in Eq. (37) to be uniform $1/(\Delta_2 - \Delta_1)$, in the range $[\Delta_1, \Delta_2]$ such that $\lambda(\Delta_1) = 1$ and $\lambda(\Delta_2) = 0.8$.

2. Photon loss

In the presence of photon loss, the coherent state $|\alpha_0\rangle\langle\alpha_0|$ remains as a coherent state with reduced amplitude coherent state $|\bar{\alpha}_0\rangle\langle\bar{\alpha}_0|$, where $\bar{\alpha}_0 = \sqrt{(1-\Delta)}\alpha_0$. The squeezed vacuum state $|r, 0\rangle\langle r, 0|$, where the r is a squeezing parameter, transforms to a squeezed thermal state [42] $\sum_{k=0}^{\infty}\frac{\bar{N}^k}{(\bar{N}+1)^{k+1}}|\bar{r}, k\rangle\langle\bar{r}, k|_2$ where $\bar{N} \equiv$
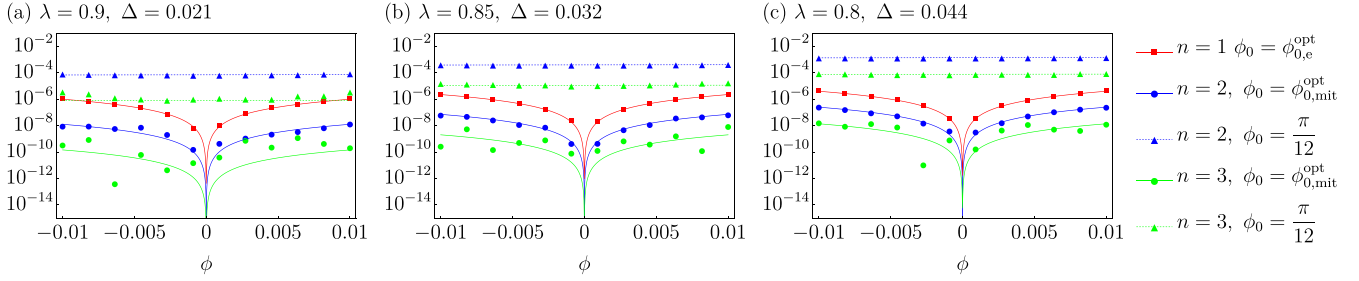
FIG. 6. (a)–(c) Bias errors (with log scale) from simulation, exploiting the N00N state ($N = 5$) as a quantum probe in the presence of photon loss with different noise strengths. Other features of the figure are the same as in Fig. 5.

$\frac{-1+\sqrt{1-4N_r\{(1-\Delta)^2-(1-\Delta)\}}}{2}$ and $\bar{r} \equiv \frac{1}{4}\ln[\frac{(1-\Delta)(2N_r+2\sqrt{N_r^2+N_r})+1}{(1-\Delta)(2N_r-2\sqrt{N_r^2+N_r})+1}]$.
Here $|\bar{r}, k\rangle$ is a squeezed number state with squeezing parameter $\bar{r}$. Again, because we assume that the photon loss occurs on both modes with the same loss rate and the loss channel commutes with linear optical operations which are phase shift and beam splitters, the error state is

$$\hat{\rho}_e = \sum_{k=0}^{\infty} \frac{\bar{N}^k}{(\bar{N}+1)^{k+1}}(\hat{\mathcal{U}}|\bar{\alpha}_0\rangle\langle\bar{\alpha}_0|_1 \otimes |\bar{r}, k\rangle\langle\bar{r}, k|_2\hat{\mathcal{U}}^\dagger), \quad (64)$$

where $\hat{\mathcal{U}} \equiv \hat{U}_{BS}\hat{\Phi}(\phi + \phi_0)\hat{U}_{BS}$.

We emphasize that the dominant eigenvector $|\psi\rangle = \hat{\mathcal{U}}|\bar{\alpha}_0\rangle_1|\bar{r}, k = 0\rangle_2$ is different from the ideal state. According to our analysis in Sec. III A, we can expect that VPEM will not be beneficial since the dominant eigenvector and the ideal state are different.

We numerically show the inefficacy of VPEM. For the numerical simulations, we consider the noise strength in the range $[\Delta_1 = 0, \Delta_2 = 0.146]$ and assume that Δ follows uniform distribution in the range. Here Δ_2 satisfies $\lambda(\Delta_2) = 0.8$ where λ is the dominant eigenvalue of the error state. The optimal reference point ϕ_0^{opt} satisfying Eq. (36) depends on a given Δ for both error and mitigation cases. The relation between Δ and ϕ_0^{opt}'s for the error case and the mitigation case with $n = 2$ and $n = 3$ is provided in Fig. 7. Therefore, we consider an averaged optimal reference point that satisfies

Eq. (37), which are $\phi_{0,e}^{opt} \approx 0.2$, $\phi_{0,mit}^{opt} \approx 0.087$ for $n = 2$, and $\phi_{0,mit}^{opt} \approx 0.037$ for $n = 3$. Under the chosen optimal reference points, we numerically simulate the bias errors of the ϕ estimation exploiting the product of the coherent and the squeezed state in the presence of photon loss. Figure 8 shows the noneffectiveness of VPEM-based quantum metrology as implied by our analysis because of the difference between the dominant eigenvector and the ideal state. For the most of the regions of ϕ, $B_e(\phi, \phi_{0,e}^{opt}, \Delta)^2$, $B_{mit}(\phi, \phi_{0,mit}^{opt}, \Delta, n = 2)^2$, and $B_{mit}(\phi, \phi_{0,mit}^{opt}, \Delta, n = 3)^2$ are similar in magnitude. Moreover, there are combinations of Δ and ϕ where $B_e(\phi, \phi_{0,e}^{opt}, \Delta)^2$ is smaller than the others.

3. Additive Gaussian noise

For pedagogical purposes, we consider an additive Gaussian noise, which might not be directly relevant to experiments. Let us assume that the additive Gaussian noise occurs only on the second mode of the initial quantum probe $|\alpha_0\rangle_1|r, 0\rangle_2$. Furthermore, we assume that Δ_x and Δ_p in Eq. (39) are

$$\Delta_x = \sqrt{\frac{\Delta}{2}}e^{-r}, \quad \Delta_p = \sqrt{\frac{\Delta}{2}}e^r, \quad (65)$$

where r is a squeezing parameter in $|r, 0\rangle_2$ and Δ is a noise strength. We note that the additive Gaussian noise is a Gaussian error such that a Gaussian state remains as
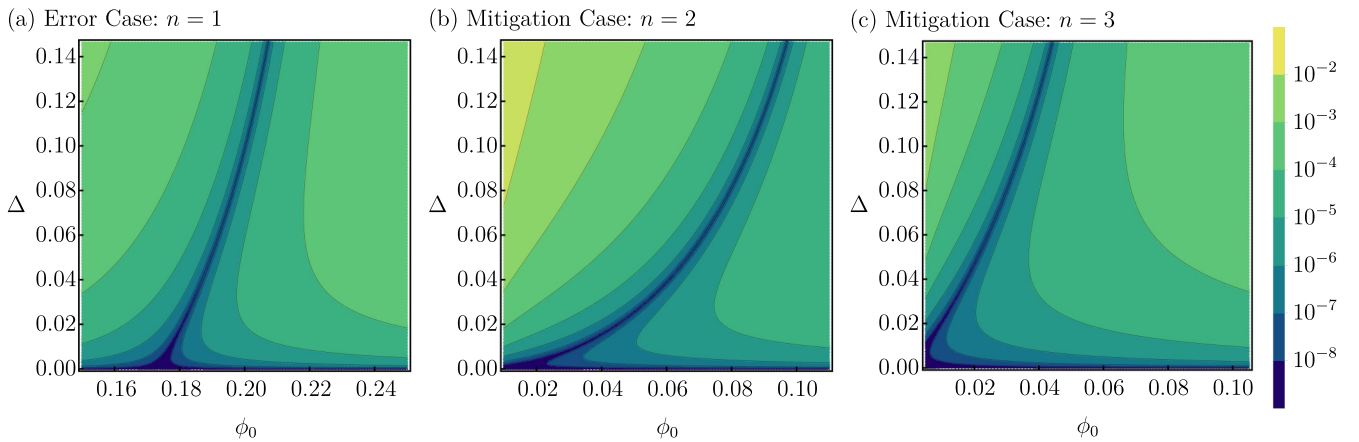

FIG. 7. Contour plots of $[(x(\phi_0) - x_{id}(\phi_0))/y_{id}(\phi_0)]^2$ with log scale, in the presence of photon loss. Here $n = 2$. Panel (a) is for the error case, and (b) is for the mitigation case. The horizontal axis is ϕ_0 and the vertical axis is Δ. The darkest line, where $[(x(\phi_0) - x_{id}(\phi_0)/y_{id}(\phi_0)]^2 = 0$, shows the relation between $\phi_{0,e}^{opt}$ ($\phi_{0,mit}^{opt}$) that satisfies Eq. (36) and Δ.
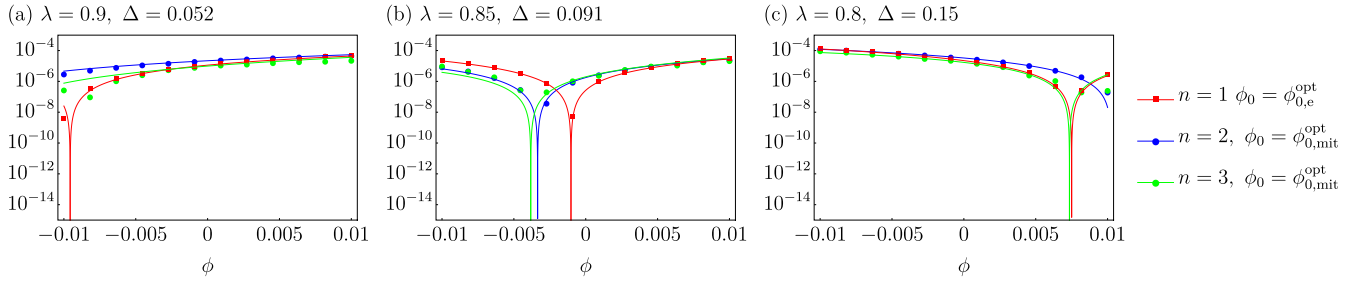
FIG. 8. (a)–(c) Bias errors (with log scale) from simulation using the product of the coherent state ($N_c = 2.5$) and squeezed vacuum state ($N_r = 2.5$) in the presence of photon loss with different noise strengths. The red regions are where the bias of the error case is smaller than the mitigation one, and the blue regions are the opposite. Other features of the figure are the same as in Fig. 5.

a Gaussian state after suffering from the noise. Similar to the photon loss case, the additive Gaussian noise makes a single-mode squeezed state to squeezed thermal state [42] $\sum_{k=0}^{\infty} \frac{\Delta^k}{(\Delta+1)^{k+1}} |r, k\rangle\langle r, k|$. Since we consider the additive Gaussian noise only on the second mode of the quantum probe at the state preparation stage (before passing the first beam splitter), as a consequence, the error state is

$$\hat{\rho}_e = \sum_{k=0}^{\infty} \frac{\Delta^k}{(\Delta+1)^{k+1}} [\hat{\mathcal{U}}|\alpha_0\rangle\langle\alpha_0|_1 \otimes |r, k\rangle\langle r, k|_2 \hat{\mathcal{U}}^\dagger], \quad (66)$$

whose dominant eigenvector is the same as the ideal state, which means $\langle\psi|\hat{A}|\psi\rangle = \langle\psi_{id}|\hat{A}|\psi_{id}\rangle$. In this case, $a_1(\phi)$ and $\frac{\partial a_1(\phi)}{\partial \phi}$ are always 0 regardless of ϕ, therefore, one can expect that error mitigation can efficiently reduce bias error.

We numerically show the efficiency of VPEM. For the numerical simulations, we consider the noise strength in the range $[\Delta_1 = 0, \Delta_2 = 0.25]$ and again assume that $p(\Delta) = 1/(\Delta_2 - \Delta_1)$. Here Δ_2 satisfies $\lambda(\Delta_2) = 0.8$. In an additive Gaussian noise case again, the optimal reference point that satisfies Eq. (36) depends on noise strength (see Fig. 9), and we consider an averaged optimal reference point that satisfies Eq. (37), which is $\phi_{0,e}^{opt} \approx 0.338$ and $\phi_{0,mit}^{opt} \approx 0.318$. Figure 10 is a simulation result with the product of the coherent and squeezed state in the presence of additive Gaussian noise. In Fig. 10 we show that even though one adopts the reference points that satisfy Eq. (37), VPEM can effectively reduce the bias because the additive Gaussian noise lets the dominant eigenvector be the ideal state. For most ϕ's, one can clearly find that $B_e(\phi, \phi_{0,e}^{opt}, \Delta)^2 \geqslant B_{mit}(\phi, \phi_{0,mit}^{opt}, \Delta)^2$. In addition, we

can also find the importance of the reference point in Fig. 10. We find that $B_e(\phi, \phi_{0,e}^{opt}, \Delta)^2 \leqslant B_{mit}(\phi, \frac{\pi}{30}, \Delta)^2$ even though $\langle\psi|\hat{A}|\psi\rangle = \langle\psi_{id}|\hat{A}|\psi_{id}\rangle$.

V. DISCUSSION

We have investigated crucial factors that dictate the efficiency of VPEM when one applies the method to quantum metrology. We find that the dominant eigenvector has to be close to the ideal state, with respect to an observable, to reduce the bias by VPEM successfully. More specifically, $\langle\psi|\hat{A}|\psi\rangle - \langle\psi_{id}|\hat{A}|\psi_{id}\rangle$ determines the reducible amount of bias. In addition, we argue that one should carefully choose an optimal reference point that minimizes bias error. Otherwise, the bias error of the mitigation case could be larger than the error case even if the dominant eigenvector is equal to the ideal state, because of the inadequate reference point. We emphasize that many of the practical estimation scenarios giving an additional parameter for the reference point have already been implemented experimentally, such as phase estimation schemes [17]. Based on the analysis, we analytically and numerically inspect the phase estimation scheme in the interferometer system. We consider the N00N state and the product of the coherent and the squeezed state as quantum probes in the presence of phase diffusion, photon loss, and additive Gaussian noise. We maintain that all the phase estimation schemes that we mentioned above can be explained by our analysis.

Finally, we point out the differences from the recent study [29]. First, while Ref. [29] analyzes bias up to the zeroth order of ϕ, we consider up to the first order of ϕ, which is essential because according to our optimal reference point strategy, the zeroth order of ϕ can vanish for both the error case and mitigation case. Second, we study the cases where VPEM cannot effectively reduce the bias, which has not been studied before, to the best of our knowledge. Last, we finally present the importance of the reference point. In the previous study, because the optimal reference point is fixed for the error case and mitigated case, the role of the reference point was not considered. It would be interesting future work to find a strategy for choosing an optimal reference point other than Eq. (37). For example, one can develop an adaptive strategy that exploits measurement outcomes to estimate a noise strength and updates a reference point using the estimated noise strength. This method may give a tailored reference
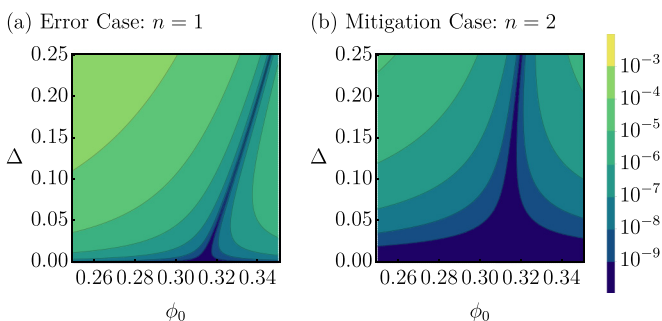

FIG. 9. Contour plots of $[(x(\phi_0) - x_{id}(\phi_0))/y_{id}(\phi_0)]^2$ with log scale, in the presence of additive Gaussian noise. Other features of the figure are similar to those in Fig. 7.
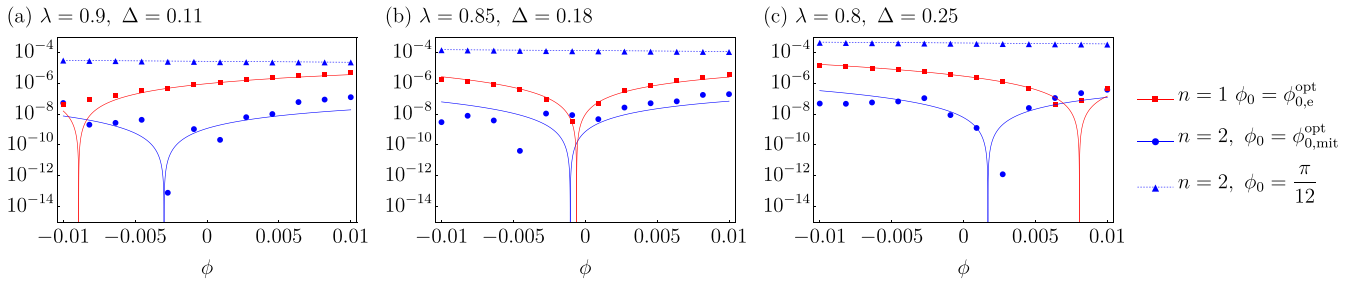
FIG. 10. (a)–(c) The simulations of the bias errors (with log scale) correspond to the product of the coherent state ($N_c = 2.5$) and squeezed vacuum state ($N_r = 2.5$) in the presence of additive Gaussian noise with different noise strengths. Other features of the figure are the same as in Fig. 5.

point for a given noise which results in a more reduced bias than the reference point in Eq. (37).

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

[2] T. Baumgratz and A. Datta, Quantum enhanced estimation of a multidimensional field, Phys. Rev. Lett. **116**, 030801 (2016).

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

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

[5] D. Braun, G. Adesso, F. Benatti, R. Floreanini, U. Marzolino, M. W. Mitchell, and S. Pirandola, Quantum-enhanced measurements without entanglement, Rev. Mod. Phys. **90**, 035006 (2018).

[6] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Advances in photonic quantum sensing, Nature Photon **12**, 724 (2018).

[7] 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).

[8] H. Lee, P. Kok, and J. P. Dowling, A quantum rosetta stone for interferometry, J. Mod. Opt. **49**, 2325 (2002).

[9] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, Quantum limits in optical interferometry, Prog. Opt. **60**, 345 (2015).

[10] C. M. Caves, Quantum-mechanical noise in an interferometer, Phys. Rev. D **23**, 1693 (1981).

[11] L. Pezzé and A. Smerzi, Mach-zehnder interferometry at the Heisenberg limit with coherent and squeezed-vacuum light, Phys. Rev. Lett. **100**, 073601 (2008).

[12] J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott, M. Abernathy, C. Adams, R. Adhikari, C. Affeldt, B. Allen, G. Allen *et al.*, A gravitational wave observatory operating beyond the quantum shot-noise limit, Nat. Phys. **7**, 962 (2011).

[13] J. Aasi, J. Abadie, B. Abbott, R. Abbott, T. Abbott, M. Abernathy, C. Adams, T. Adams, P. Addesso, R. Adhikari *et al.*, Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light, Nat. Photon. **7**, 613 (2013).

[14] T. Nagata, R. Okamoto, J. L. O'brien, K. Sasaki, and S. Takeuchi, Beating the standard quantum limit with four-entangled photons, Science **316**, 726 (2007).

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

[16] H. Yonezawa, D. Nakane, T. A. Wheatley, K. Iwasawa, S. Takeda, H. Arao, K. Ohki, K. Tsumura, D. W. Berry, T. C. Ralph *et al.*, Quantum-enhanced optical-phase tracking, Science **337**, 1514 (2012).

[17] A. A. Berni, T. Gehring, B. M. Nielsen, V. Händchen, M. G. Paris, and U. L. Andersen, *Ab initio* quantum-enhanced optical phase estimation using real-time feedback control, Nature Photon **9**, 577 (2015).

[18] J. Yu, Y. Qin, J. Qin, H. Wang, Z. Yan, X. Jia, and K. Peng, Quantum enhanced optical phase estimation with a squeezed thermal state, Phys. Rev. Appl. **13**, 024037 (2020).

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

[20] M. G. Genoni, S. Olivares, and M. G. A. Paris, Optical phase estimation in the presence of phase diffusion, Phys. Rev. Lett. **106**, 153603 (2011).

[21] W. Dür, M. Skotiniotis, F. Fröwis, and B. Kraus, Improved quantum metrology using quantum error correction, Phys. Rev. Lett. **112**, 080801 (2014).

[22] G. Arrad, Y. Vinkler, D. Aharonov, and A. Retzker, Increasing sensing resolution with error correction, Phys. Rev. Lett. **112**, 150801 (2014).

[23] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin, Quantum error correction for metrology, Phys. Rev. Lett. **112**, 150802 (2014).

[24] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Adaptive quantum metrology under general Markovian noise, Phys. Rev. X **7**, 041009 (2017).

[25] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Achieving the Heisenberg limit in quantum metrology using quantum error correction, Nat. Commun. **9**, 78 (2018).

[26] D. Layden, S. Zhou, P. Cappellaro, and L. Jiang, Ancilla-free quantum error correction codes for quantum metrology, Phys. Rev. Lett. **122**, 040502 (2019).

[27] Q. Zhuang, J. Preskill, and L. Jiang, Distributed quantum sensing enhanced by continuous-variable error correction, New J. Phys. **22**, 022001 (2020).

[28] S. Zhou and L. Jiang, Asymptotic theory of quantum channel estimation, PRX Quantum **2**, 010343 (2021).

[29] K. Yamamoto, S. Endo, H. Hakoshima, Y. Matsuzaki, and Y. Tokunaga, Error-mitigated quantum metrology via virtual purification, Phys. Rev. Lett. **129**, 250503 (2022).

[30] B. Koczor, Exponential error suppression for near-term quantum devices, Phys. Rev. X **11**, 031057 (2021).

[31] W. J. Huggins, S. McArdle, T. E. O'Brien, J. Lee, N. C. Rubin, S. Boixo, K. B. Whaley, R. Babbush, and J. R. McClean, Virtual distillation for quantum error mitigation, Phys. Rev. X **11**, 041036 (2021).

[32] M. G. Paris, Quantum estimation for quantum technology, Int. J. Quantum Inform. **07**, 125 (2009).

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

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

[35] B. Koczor, The dominant eigenvector of a noisy quantum state, New J. Phys. **23**, 123047 (2021).

[36] Y. S. Teo, S. Shin, H. Kwon, S.-H. Lee, and H. Jeong, Virtual distillation with noise dilution, Phys. Rev. A **107**, 022608 (2023).

[37] H. F. Hofmann, All path-symmetric pure states achieve their maximal phase sensitivity in conventional two-path interferometry, Phys. Rev. A **79**, 033822 (2009).

[38] P. M. Anisimov, G. M. Raterman, A. Chiruvelli, W. N. Plick, S. D. Huver, H. Lee, and J. P. Dowling, Quantum metrology with two-mode squeezed vacuum: Parity detection beats the Heisenberg limit, Phys. Rev. Lett. **104**, 103602 (2010).

[39] A. Chiruvelli and H. Lee, Parity measurements in quantum optical metrology, J. Mod. Opt. **58**, 945 (2011).

[40] K. P. Seshadreesan, P. M. Anisimov, H. Lee, and J. P. Dowling, Parity detection achieves the Heisenberg limit in interferometry with coherent mixed with squeezed vacuum light, New J. Phys. **13**, 083026 (2011).

[41] K. P. Seshadreesan, S. Kim, J. P. Dowling, and H. Lee, Phase estimation at the quantum Cramér-Rao bound via parity detection, Phys. Rev. A **87**, 043833 (2013).

[42] A. Serafini, *Quantum Continuous Variables: A Primer of Theoretical Methods* (Taylor & Francis, Oxford, 2017).