Estimation of Gaussian random displacement using non-Gaussian states

Fumiya Hanamura ®,[1,*] Warit Asavanant ®,[1] Kosuke Fukui,[1] Shunya Konno ®,[1] and Akira Furusawa ®[1,2,†]

[1]*Department of Applied Physics, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*
[2]*Optical Quantum Computing Research Team, RIKEN Center for Quantum Computing, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*

In continuous-variable quantum information processing, quantum error correction of Gaussian errors requires simultaneous estimation of both quadrature components of displacements in phase space. However, quadrature operators x and p are noncommutative conjugate observables, whose simultaneous measurement is prohibited by the uncertainty principle. Gottesman-Kitaev-Preskill (GKP) error correction deals with this problem using complex non-Gaussian states called GKP states. On the other hand, simultaneous estimation of displacement using experimentally feasible non-Gaussian states has not been well studied. In this paper, we consider a multiparameter estimation problem of displacements assuming an isotropic Gaussian prior distribution and allowing postselection of measurement outcomes. We derive a lower bound for the estimation error when only Gaussian operations are used and show that even simple non-Gaussian states such as single-photon states can beat this bound. Based on Ghosh's bound, we also obtain a lower bound for the estimation error when the maximum photon number of the input state is given. Our results reveal the role of non-Gaussianity in the estimation of displacements and pave the way toward the error correction of Gaussian errors using experimentally feasible non-Gaussian states.

I. INTRODUCTION

Continuous-variable optical quantum information processing has attracted much attention due to its extent scalability achieved by large-scale cluster states [1,2] which enable universal Gaussian operations [3,4]. As a next step for the realization of practical quantum information processing, quantum error correction [5] is an essential key. Gaussian error is the most common type of error in optical systems. It includes photon losses and the Gaussian quantum channel [6], which is defined as phase-space displacements following an isotropic Gaussian distribution. It has been proven, however, that Gaussian errors imposed on Gaussian states cannot be corrected using only Gaussian operations [7,8]. Therefore, non-Gaussian states play a crucial role in optical quantum information processing.

Gottesman-Kitaev-Preskill (GKP) error correction [9] is one of the promising ways to correct Gaussian errors using a single-mode states. It uses highly complex non-Gaussian states called GKP states to correct errors. The information regarding displacements on both x and p quadratures imposed on GKP states is extracted by homodyne measurements and ancillary GKP states in the error syndrome measurement [9].

Although x and p are noncommutative conjugate observables whose simultaneous measurement is prohibited by the uncertainty principle, the non-Gaussianity of GKP states and the prior information that the displacement is small enable accurate estimation of both parameters. However, GKP states are difficult to experimentally generate due to their high non-Gaussianity. Although their generation has been reported in other physical systems [10,11], optical generation of GKP states has not been achieved yet. Several other protocols for correcting Gaussian errors using non-Gaussian states (called bosonic codes) are known [12], but it has not been unraveled how experimentally feasible and simpler non-Gaussian states such as Fock states can be exploited for quantum error correction.

In this paper, we investigate a multiparameter quantum estimation [13–18] problem of displacements using non-Gaussian states. We assume an isotropic Gaussian prior distribution of the parameters and postselect measurement outcomes. First, we derive a lower bound for the estimation error when only Gaussian states and Gaussian operations are used. Then we show that this bound can be beaten for some range of the prior variance even with only simple non-Gaussian states such as single-photon states. This result reveals the role of non-Gaussianity in the estimation of displacements and opens up the possibility of correcting Gaussian errors using experimentally feasible non-Gaussian states. We also derive a lower bound of the estimation error depending on the maximum photon number of the input state based on Ghosh's bound [19].

We evaluate the estimation error using the mean square error averaged with respect to the posterior distribution with the measurement outcome fixed. This is because it corresponds
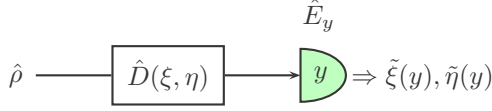
FIG. 1. A schematic of the Gaussian displacement estimation problem. One estimates the amount of displacement ξ, η using the measurement outcome y, which corresponds to the POVM element $\hat{E}_y$. One can use prior information that ξ, η follow an isotropic Gaussian distribution, Eq. (2).

to the natural situation where one knows the prior distribution of the parameters and can postselect "good" measurement outcomes, which reduces the estimation error. Although this criterion is different from the frequentist estimation error (which fixes the true value of the parameter and takes the average with respect to the measurement outcomes) or the Bayesian estimation error (which takes the average with respect to both the true value and the measurement outcomes), it is a commonly used method in Bayesian statistics [20]. Reference [21] also studies the simultaneous displacement estimation problem using non-Gaussian states, although they adopt frequentist estimation error. For Gaussian states, estimation of displacement is a classical problem and has many previous studies [22–24]. For single-parameter estimation using non-Gaussian states, there have been some recent studies [25,26].

In Sec. II, we formulate the Gaussian displacement estimation problem. In Sec. III, we derive lower bounds for the estimation error when only Gaussian (classical) states and Gaussian operations are used. In Sec. IV, we show that the bounds derived in Sec. III can be beaten using non-Gaussian states. In Sec. V, a statistical lower bound on the estimation error which depends on the maximum photon number and the postselection probability is obtained, based on Ghosh's bound. Section VI summarizes our paper and introduces some future works.

II. GAUSSIAN DISPLACEMENT ESTIMATION PROBLEM

Figure 1 depicts the Gaussian displacement estimation problem. We consider a single-mode bosonic state $\hat{\rho}$ represented by quadrature operators $\hat{x}, \hat{p}$ satisfying $[\hat{x}, \hat{p}] = i$. The phase-space displacement operator $\hat{D}(\xi, \eta)$ is defined as

$$\hat{D}(\xi, \eta) = \exp(i\eta\hat{x} - i\xi\hat{p}) \qquad (1)$$

and represents the displacement $(x, p) \rightarrow (x + \xi, p + \eta)$. First, $\hat{D}(\xi, \eta)$ acts on $\hat{\rho}$, then a measurement is performed, and the measurement outcome y is obtained. This measurement corresponds to some POVM (positive operator-valued measure) $\{\hat{E}_y\}$ satisfying $\int \hat{E}_y dy = \hat{I}$, where $\hat{I}$ is an identity operator. We assume that ξ, η are random variables following an isotropic Gaussian distribution with known variance and mean 0:

$$p(\xi, \eta) = \frac{1}{\pi v} \exp\left(-\frac{\xi^2 + \eta^2}{v}\right). \qquad (2)$$

Note that the mean square distance with respect to this distribution is $\langle \xi^2 \rangle + \langle \eta^2 \rangle = v$. Displacement following an isotropic Gaussian distribution corresponds to a common type

of noise in bosonic systems called the Gaussian quantum channel [6] or additive Gaussian noise. The problem is to estimate (ξ, η) from the value of y. We assume that one performs a Bayesian estimation using the *a priori* information, Eq. (2). The conditional probability density of obtaining y, when the values of (ξ, η) are fixed, is given by

$$p(y|\xi, \eta) = \text{Tr}\{\hat{D}(\xi, \eta)\hat{\rho}\hat{D}^\dagger(\xi, \eta)\hat{E}_y\}, \qquad (3)$$

where $\text{Tr}\{\cdot\}$ denotes the trace operation. Defining the Wigner function $W_{\hat{A}}$ of an operator $\hat{A}$ by

$$W_{\hat{A}}(x, p) := \frac{1}{2\pi} \int \exp(ipx') \left\langle x - \frac{x'}{2} \left| \hat{A} \right| x + \frac{x'}{2} \right\rangle dx', \quad (4)$$

Eq. (3) can be expressed in terms of the Wigner functions of $\hat{\rho}$ and $\hat{E}_y$ [27]:

$$p(y|\xi, \eta) = 2\pi \int W_{\hat{\rho}}(x - \xi, p - \eta)W_{\hat{E}_y}(x, p)dxdp. \qquad (5)$$

Once the value of y is known, one gets the corresponding posterior distribution of ξ, η,

$$p(\xi, \eta|y) = \frac{p(y|\xi, \eta)p(\xi, \eta)}{p(y)}, \qquad (6)$$

where $p(y) = \iint p(y|\xi, \eta)p(\xi, \eta)d\xi d\eta$ is the marginal probability distribution of y. Suppose one estimates the values of ξ, η as $\tilde{\xi}(y), \tilde{\eta}(y)$, corresponding to the value of y. To evaluate the accuracy of the estimation, the mean square error with respect to the joint probability distribution is one of the standard choices in Bayesian quantum estimation,

$$v'_{\text{Bayes}} := \int p(\xi, \eta, y)\{(\xi - \tilde{\xi}(y))^2 + (\eta - \tilde{\eta}(y))^2\}d\xi d\eta dy, \qquad (7)$$

where $p(\xi, \eta, y) = p(y|\xi, \eta)p(\xi, \eta)$ is the joint probability distribution of (ξ, η, y).

However, here we consider the postselection of a specific measurement outcome y and take the average only with respect to the true value ξ, η, obtaining

$$v' := \int p(\xi, \eta|y)\{(\xi - \tilde{\xi}(y))^2 + (\eta - \tilde{\eta}(y))^2\}d\xi d\eta. \qquad (8)$$

The meaning of this quantity can be considered the expected amount of error after obtaining the outcome y. Note that more generally there are other options for the estimation error measure (e.g., mean square error with a general weight matrix [28,29]), and here we choose this v' for simplicity. We consider the case where $\tilde{\xi}(y), \tilde{\eta}(y)$ are chosen as the averages with respect to the posterior distribution,

$$\tilde{\xi}(y) = \int p(\xi, \eta|y)\xi d\xi d\eta, \qquad (9)$$

$$\tilde{\eta}(y) = \int p(\xi, \eta|y)\eta d\xi d\eta, \qquad (10)$$

which is the optimal choice to minimize v'. For any probability density function g satisfying $\int g(\xi, \eta)d\xi d\eta = 1$ and $g(\xi, \eta) > 0$, we define $\Sigma[g(\xi, \eta)]$ as the covariance matrix of (ξ, η) with respect to $g(\xi, \eta)$. Then v' can also be expressed as

$$v' = \text{Tr}\{\Sigma[p(\xi, \eta|y)]\}. \qquad (11)$$

Note that v' depends on the measurement outcome y. Usually, in Bayesian quantum estimation, one evaluates the error of estimation by averaging over the measurement outcomes [16]. However, in this paper, we allow postselection of y to get a smaller v'. In Sec. III, we see that one can still obtain lower bounds on v' with this assumption.

We also assume that there exists no initial entanglement between the input state and the measurement system. The case where entanglements exist is studied in, e.g., Ref. [24], and it is shown that the amount of displacement can be estimated with an arbitrary high precision if one uses a two-mode squeezed vacuum with a high squeezing level.

III. CLASSICAL AND GAUSSIAN BOUNDS

In this section, we derive lower bounds on the estimation error v' [Eq. (8)], when only Gaussian states and Gaussian operations are allowed. We only have to consider the case where $\hat{\rho}$ is a pure Gaussian state and $\hat{E}_y \propto |\psi\rangle\langle\psi|$ for some pure Gaussian state $|\psi\rangle$. Indeed, suppose we have $\hat{\rho} = t\hat{\rho}_1 + (1-t)\hat{\rho}_2$, $0 < t < 1$. Then $v'(\hat{\rho}) \geqslant \min\{v'(\hat{\rho}_1), v'(\hat{\rho}_2)\}$ because the variance is concave with respect to the probability distribution. In the same way, one can also show that mixing $\hat{E}_y$ does not reduce v' either. Wigner functions of pure Gaussian states can be written as

$$W(x, p) = \frac{1}{\pi} \exp\left(-\frac{1}{2}(\boldsymbol{q} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{q} - \boldsymbol{\mu})\right), \quad (12)$$

where $\boldsymbol{q} := (x, p)^T$, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are parameters representing the mean and the covariance matrix of the Gaussian function, and $\boldsymbol{\Sigma}$ satisfies $\det \boldsymbol{\Sigma} = \frac{1}{4}$. Note that this constraint means that Heisenberg's uncertainty is saturated by pure Gaussian states [30]. Thus, the Wigner functions of $\hat{\rho}$ and $\hat{E}_y$ are both Gaussian functions and satisfy

$$\det \Sigma[W_{\hat{\rho}}(\xi, \eta)] = \det \Sigma[W_{\hat{E}_y}(\xi, \eta)] = \frac{1}{4}. \quad (13)$$

Here we extend our definition of Σ in the last section to unnormalized distribution functions by defining $\Sigma[g(\xi, \eta)] := \Sigma[\bar{g}(\xi, \eta)]$ for any unnormalized distribution function $g(\xi, \eta)$ satisfying $\int g(\xi, \eta)d\xi d\eta < \infty$ and $g(\xi, \eta) > 0$, where $\bar{g}(\xi, \eta) = Ng(\xi, \eta)$ is the normalization of g which satisfies $\int \bar{g}(\xi, \eta)d\xi d\eta = 1$. Note that although generally Wigner functions can have negative parts, Gaussian Wigner functions are positive and thus can be regarded as probability distributions. Using Eq. (5) and a property of convolution, $p(y|\xi, \eta)$ is also Gaussian with respect to (ξ, η), and its covariance matrix is

$$\Sigma[p(y|\xi, \eta)] = \Sigma[W_{\hat{\rho}}(\xi, \eta)] + \Sigma[W_{\hat{E}_y}(\xi, \eta)]. \quad (14)$$

Furthermore, from Eq. (6), $p(\xi, \eta|y)$ is also Gaussian, and

$$\Sigma[p(\xi, \eta|y)]^{-1} = \Sigma[p(y|\xi, \eta)]^{-1} + \Sigma[p(\xi, \eta)]^{-1} \quad (15)$$

holds.

A. Classical bound

First, we consider the case where only classical states, i.e., coherent states, are available as $\hat{\rho}$. In this case, we have $\Sigma[W_{\hat{\rho}}(\xi, \eta)] = \frac{1}{2}I$, where I is the identity matrix. From Eq. (13), two eigenvalues of $\Sigma[W_{\hat{E}_y}(\xi, \eta)]$ can be written as

$a/2$, $1/(2a)$ for $a > 0$. Then, from Eqs. (11), (14), and (15), one obtains

$$v' = \frac{1}{2/v + 2/(1+a)} + \frac{1}{2/v + 2a/(1+a)}. \quad (16)$$

This is minimized when $a = 1$, regardless of the value of v. Therefore, in this case, the optimal $\hat{E}_y$ is a projection to a classical state, and the classical lower bound of v' is given by

$$v'_{\mathrm{C}} = \frac{2v}{v + 2}. \quad (17)$$

B. Gaussian bound

Next we consider the case where arbitrary Gaussian states and operations can be used. For positive definite 2×2 matrices A, B,

$$\det(A + B) \geqslant \det A + \det B + 2\sqrt{\det A \det B} \quad (18)$$

holds (see Appendix B for the derivation). Using this fact and Eqs. (13) and (14), one can show that

$$\det \Sigma[p(y|\xi, \eta)] \geqslant 1. \quad (19)$$

Conversely, any Gaussian $p(y|\xi, \eta)$ satisfying Eq. (19) is possible by taking $\Sigma[W_{\hat{\rho}}(\xi, \eta)] = \Sigma[W_{\hat{E}_y}(\xi, \eta)] = \Sigma[p(y|\xi, \eta)]/2$. Because $p(y|\xi, \eta)$ minimizing v' must saturate Eq. (19), its two eigenvalues can be written as a, $1/a$ for some $a > 0$, and

$$v' = \frac{1}{2/v + a} + \frac{1}{2/v + 1/a} \quad (20)$$

holds. The value of a which minimizes v' depends on the value of v. When $v > 2$, $a = 1$ minimizes v'. This means that choosing vacuum (or coherent) states for both $\hat{\rho}$ and $\hat{E}_y$ is optimal, hence the minimum of v' coincides with the classical bound, $\frac{2v}{v+2}$. On the other hand, when $v < 2$, by taking $a \to 0$ or $a \to \infty$, i.e., taking both $\hat{\rho}$ and $\hat{E}_y$ to be infinitely squeezed states, v' approaches the lower bound $v/2$. It is precisely half of the original variance v, because variance along the squeezed direction is made 0, and the variance along the antisqueezed direction remains $v/2$. To summarize, the Gaussian lower bound of v' is given by

$$v'_{\mathrm{G}} = \begin{cases} \frac{v}{2} & (v < 2), \\ \frac{2v}{v+2} & (v > 2). \end{cases} \quad (21)$$

Note that the lower bounds of the Bayesian estimation error v'_{Bayes} [Eq. (7)] in these settings are also given by v'_{C} and v'_{G} in Eqs. (17) and (21), because these bounds are independent of the measurement outcome y. (The classical bound v'_{C} matches the result in Ref. [24], which is derived as the Bayesian right logarithmic derivative quantum Cramér-Rao bound. The authors also show the tightness of the bound.) This means that postselection does not improve the estimation accuracy in these settings, which we show is not true for non-Gaussian cases.

IV. ESTIMATION USING NON-GAUSSIAN STATES

In this section, we show that the bounds introduced in Sec. III can be beaten using non-Gaussian states and a
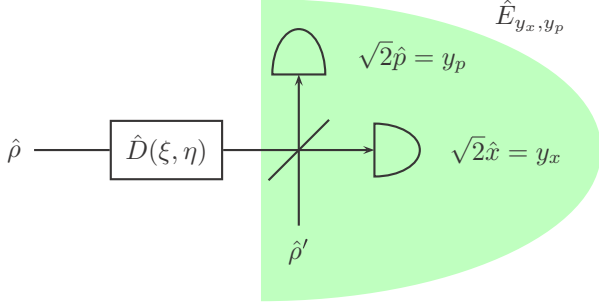
FIG. 2. The case where one uses heterodyne measurement as the measurement in Fig. 1. First, the state to be measured is combined with an ancillary state by a half-beam splitter, then $\sqrt{2}\hat{x}$ and $\sqrt{2}\hat{p}$ are measured in each mode. The corresponding POVM is expressed by Eq. (22).

heterodyne measurement, which is widely used for simultaneous measurements of both quadratures [31–34]. As examples of non-Gaussian states, we consider the GKP state [9], which is expected to give a high estimation accuracy as in Ref. [21], and Fock states, which are experimentally feasible for small photon numbers.

A. Heterodyne measurement

As the measurement $\{\hat{E}_y\}$ for the system in Fig. 1, we consider a setup called heterodyne measurement shown in Fig. 2, which consists of a half-beam splitter, an ancillary state $\hat{\rho}'$, and two homodyne measurements. The input state after the displacement is combined with $\hat{\rho}'$ by the half-beam splitter, then $\sqrt{2}\hat{x}$ and $\sqrt{2}\hat{p}$ are measured in each mode. The case where $\hat{\rho}'$ is a vacuum state can be seen as a sampling from the Husimi Q function and has wide applications, e.g., state verifications [31,32]. Here we consider a more general case where any single-mode state can be used as $\hat{\rho}'$ [33,34]. Denoting two measurement outcomes $y = (y_x, y_p)$, the corresponding POVM element is [27]

$$\hat{E}_{y_x, y_p} = \frac{1}{2\pi} \hat{D}(y_x, y_p) \hat{\rho}'^* \hat{D}^{\dagger}(y_x, y_p), \quad (22)$$

where $\hat{\rho}'^*$ is the operator whose matrix elements are the complex conjugate of those of $\hat{\rho}'$. Because taking the complex conjugate corresponds to the time reversal operation, the Wigner function of $\hat{\rho}'^*$ is given by $W_{\hat{\rho}'^*}(x, p) = W_{\hat{\rho}'}(x, -p)$. Therefore, substituting Eq. (22) into Eq. (3), one can see that the conditional probability density $p(y_x, y_p | \xi, \eta)$ only depends on $(\xi - y_x, \eta - y_p)$,

$$p(y_x, y_p | \xi, \eta)$$
$$= \frac{1}{2\pi} \text{Tr}\left\{\hat{D}(\xi - y_x, \eta - y_p)\hat{\rho}\hat{D}^{\dagger}(\xi - y_x, \eta - y_p)\hat{\rho}'^*\right\}$$
$$= f(\xi - y_x, \eta - y_p), \quad (23)$$

where the "filter function" f is determined by $\hat{\rho}$ and $\hat{\rho}'$ as

$$f(x, p) := \int W_{\hat{\rho}}(x' - x, p' - p) W_{\hat{\rho}'}(x', -p') dx' dp' \quad (24)$$

using Eq. (5). From Eq. (6), the posterior distribution of ξ, η is obtained by multiplying the prior distribution by the filter
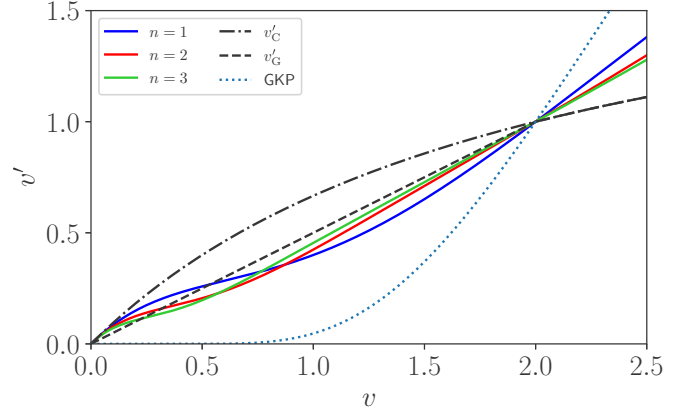

FIG. 3. The relation between v and v'. "$n = 1$," "$n = 2$," and "$n = 3$": $\hat{\rho} = \hat{\rho}' = |n\rangle\langle n|$, and $(y_x, y_p) = (0, 0)$ is postselected. "GKP": $\hat{\rho} = \hat{\rho}' = |GKP\rangle\langle GKP|$, and $(y_x, y_p) = (0, 0)$ is postselected. "$v'_{\text{Bayes}}(n = 1)$": $\hat{\rho} = \hat{\rho}' = |1\rangle\langle 1|$, without postselection. v'_C and v'_G are the classical and Gaussian bounds derived in Sec. III, respectively.

function displaced by the measurement outcome y_x, y_p:

$$p(\xi, \eta | y_x, y_p) \propto p(\xi, \eta) f(\xi - y_x, \eta - y_p). \quad (25)$$

Note that this form of measurement is not only experimentally feasible, but also quite general in the presence of postselection, because having any POVM element $\hat{E}$ is equivalent to taking $\hat{\rho}'^* \propto \hat{E}$ and postselecting $y_x = y_p = 0$ in this heterodyne setting.

B. Estimation using the GKP state

The GKP state, or grid state [21], is a non-Gaussian state used for GKP error correction [9] and is defined as

$$|GKP\rangle \propto \sum_{s=-\infty}^{\infty} |x = \sqrt{2\pi}s\rangle. \quad (26)$$

Note that this state is not normalizable and is, therefore, an unphysical state. The Wigner function of the GKP state is a sum of Dirac delta functions:

$$W_{GKP}(x, p) \propto \sum_{s=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} (-1)^{st} \delta(x - \sqrt{\pi/2}s)\delta(p - \sqrt{\pi/2}t). \quad (27)$$

If one takes $\hat{\rho} = \hat{\rho}' = |GKP\rangle\langle GKP|$, the filter function, Eq. (24), becomes

$$f_{GKP}(x, p) = \sum_{s=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} \delta(x - \sqrt{2\pi}s)\delta(p - \sqrt{2\pi}t). \quad (28)$$

Reference [21] points out that this (unphysical) GKP state can achieve 0 estimation error in the non-Bayesian setting with the local unbiasedness condition. It is expected to achieve a good estimation accuracy also in our setting, where we consider Bayesian estimation with postselection. This is because the prior information that (ξ, η) is near the origin and the locally sharp structure of the filter function circumvents the uncertainty principle. The line labeled "GKP" in Fig. 3 shows the relation between v and v' when one takes $\hat{\rho} =$
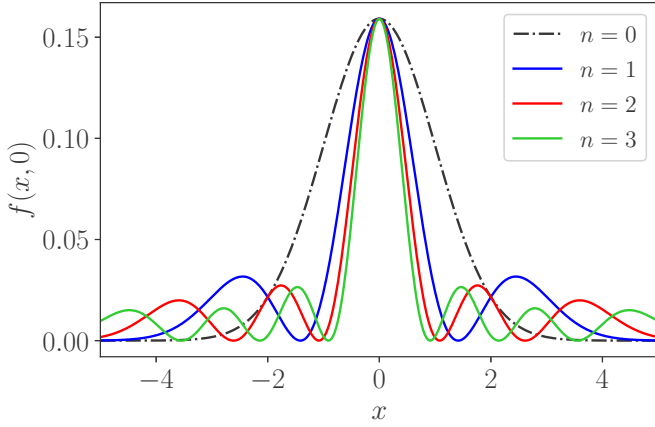
FIG. 4. The filter function $f_n(x, 0)$ when $\hat{\rho} = \hat{\rho}' = |n\rangle\langle n|$ for $n = 0, 1, 2, 3$ [Eq. (29)].

FIG. 5. The relation between v and v', when $\hat{\rho} = \hat{\rho}' = (1 - l)|1\rangle\langle 1| + l|0\rangle\langle 0|$, where l corresponds to the photon loss rate.

$\hat{\rho}' = |\text{GKP}\rangle\langle\text{GKP}|$ and $(y_x, y_p) = (0, 0)$ is postselected. One can see that v' is largely reduced compared to the Gaussian and classical bounds for small v. Note that even with the ideal GKP state, the Bayesian estimation error v' remains finite due to the finite tail of the Gaussian prior distribution, unlike in frequentist's locally unbiased estimation schemes such as in Ref. [21]. On the other hand, v' becomes larger than the classical bound in the region $v \geqslant 2$. This may be due to the fact that the filter function $f_{\text{GKP}}(x, p)$ has a large variance. Note also the similarity of this setting to the error syndrome measurement for the GKP code [9], where the SUM gate is used instead of the beam splitter.

C. Estimation using Fock states

While the GKP state is obviously useful for the estimation of displacement, experimental generation of (an approximation with the physical state of) it has not been realized in optics. We now discuss whether the Gaussian bound can be beaten using experimentally feasible non-Gaussian states. Fock states $|n\rangle$ are the simplest examples of non-Gaussian states and, also, experimentally feasible when n is small [35,36], although they are still harder than Gaussian states. If one takes $\hat{\rho} = \hat{\rho}' = |n\rangle\langle n|$, the filter function, Eq. (24), becomes

$$f_n(x, p) = \frac{1}{2\pi}\left[L_n\left(\frac{x^2 + p^2}{2}\right)\right]^2 \exp\left(-\frac{x^2 + p^2}{2}\right), \quad (29)$$

where $L_n(\cdot)$ is the Laguerre polynomial. $f(x, 0)$ for $n = 0, 1, 2, 3$ is shown in Fig. 4. The distance from the origin to the first zero of f_n scales as $n^{-1/2}$ [37]. Thus, roughly speaking, for sufficiently small v and when $(y_x, y_p) = (0, 0)$ is postselected, f_n acts like a Gaussian function of variance $\sim n^{-1}$, therefore Fock states are candidates for good input states. Note that when $\hat{\rho} = \hat{\rho}'^*$, $f_n(x, p)$ always has a maximum value of $1/(2\pi)$ at the origin.

Figure 3 shows the relation between v and v' for $n = 1, 2, 3$. For comparison, the Bayesian estimation error v'_{Bayes} without postselection [Eq. (7)] for $n = 1$ is also shown. One can see that the estimation accuracy is improved by postselection, and even when $n = 1$, the Gaussian bound is beaten in some range of v. We also calculate the effect of photon
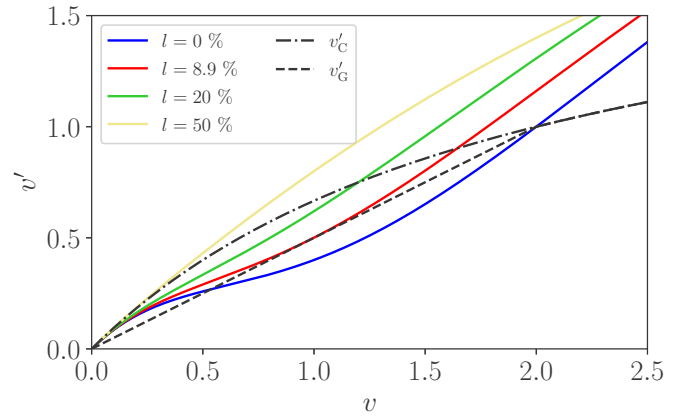
losses on both input states for the case $\hat{\rho} = \hat{\rho}' = |1\rangle\langle 1|$, as is important for an actual experimental realization (Fig. 5). We assume the same amount of losses for $\hat{\rho}$ and $\hat{\rho}'$. Losses of up to 8.9% and 50% are allowed for beating the Gaussian bound and the classical bound, respectively.

Note that practically postselecting a single value of (y_x, y_p) is impossible in continuous-variable measurements such as the heterodyne measurement, and one has to select events in some range of (y_x, y_p) to have a sufficiently high postselection probability. In Appendix C, we include an analysis of the case selecting a finite range of outcomes with a finite probability in the single-photon case. It shows that the classical and Gaussian bounds are still beaten in some range of v, with finite postselection probability.

It is interesting to observe that for all GKP-state and Fock-state cases, v' becomes 1 at $v = 2$. In fact, one can explicitly show that when $v = 2$, $v' \geqslant 1$ holds for arbitrary $\hat{\rho}$ and $\hat{E}_y$, and the equality holds when $\hat{\rho} \propto \hat{E}_y \propto |\psi\rangle\langle\psi|$ for some $|\psi\rangle$. See Appendix D for the derivation.

V. ANALYSIS OF LOWER BOUNDS FOR NON-GAUSSIAN STATES

We return to the general setting in Sec. II. One natural matter of interest is to find better input states and find the limit of estimation using non-Gaussian states. Generally, preparation of non-Gaussian states with large photon numbers is difficult, and experimentally feasible states lie in a subspace with some maximum photon number [36]. In this section, we discuss lower bounds on v', when $\hat{\rho}$ is a superposition of Fock states up to the maximum photon number n.

In quantum estimation theory, the Cramér-Rao bound [16–18,38,39] is often used to obtain a lower bound on the estimation error. Because we have prior information and postselect a single measurement outcome in our case, we use an inequality called Ghosh's bound instead [19]. See Appendix A for the detailed derivation of Ghosh's bound and its evaluation. According to the inequality, a lower bound for v', when the measurement outcome is y, is given by

$$v' \geqslant \frac{4}{F_0(y) + F(y)}, \quad (30)$$

where $F_0(y)$ and $F(y)$ are defined as

$$F_0(y) := \sum_i \int p(\boldsymbol{\theta}|y)\left(-\frac{\partial^2}{\partial \theta_i^2}\ln p(\boldsymbol{\theta})\right)d\boldsymbol{\theta}, \qquad (31)$$

$$F(y) := \sum_i \int p(\boldsymbol{\theta}|y)\left(-\frac{\partial^2}{\partial \theta_i^2}\ln p(y|\boldsymbol{\theta})\right)d\boldsymbol{\theta}. \qquad (32)$$

Here we write $\boldsymbol{\theta} := (\theta_1, \theta_2) := (\xi, \eta)$, $(\hat{q}_1, \hat{q}_2) := (\hat{x}, \hat{p})$, and $d\boldsymbol{\theta} := d\theta_1 d\theta_2$. The quantities $F(y)$ and $F_0(y)$ can be considered variants of Fisher information. There is an upper bound on $F(y)$,

$$F(y) \leqslant \sum_i \int d\boldsymbol{\theta}\frac{p(\boldsymbol{\theta})}{p(y)}\mathrm{Tr}\left\{\left(2\hat{q}_i\hat{\rho}\hat{q}_i + \hat{q}_i^2\hat{\rho} + \hat{\rho}\hat{q}_i^2\right)\hat{E}_{y|\boldsymbol{\theta}}\right\}, \qquad (33)$$

where

$$\hat{E}_{y|\boldsymbol{\theta}} := \hat{D}^\dagger(\boldsymbol{\theta})\hat{E}_y\hat{D}(\boldsymbol{\theta}) \qquad (34)$$

(see Appendix A for the derivation). For the Gaussian prior distribution, Eq. (2), $F_0(y)$ is simply a constant:

$$F_0(y) = \frac{4}{v}. \qquad (35)$$

Therefore, inequality Eq. (30) can be transformed as

$$\frac{1}{v'} \leqslant \frac{1}{v} + \frac{F(y)}{4}. \qquad (36)$$

An upper bound for $F(y)$, in the case where the maximum photon number of $\hat{\rho}$ is n, can be obtained using the Schwarz inequality. For pure $\hat{\rho} = |\psi\rangle\langle\psi|$, we have

$$\begin{aligned}
\sum_i \mathrm{Tr}\{\hat{q}_i\hat{\rho}\hat{q}_i\hat{E}_{y|\boldsymbol{\theta}}\} &\leqslant \sqrt{\sum_i \mathrm{Tr}\{\hat{q}_i^2\hat{\rho}\hat{q}_i^2\hat{\rho}\}}\sqrt{2\mathrm{Tr}\,\hat{E}_y^2} \\
&= \sqrt{\sum_i (\langle\psi|\hat{q}_i^2|\psi\rangle)^2}\sqrt{2\mathrm{Tr}\,\hat{E}_y^2} \\
&\leqslant \sum_i \langle\psi|\hat{q}_i^2|\psi\rangle\sqrt{2\mathrm{Tr}\,\hat{E}_y^2} \\
&= \langle\psi|(2\hat{n}+1)|\psi\rangle\sqrt{2\mathrm{Tr}\,\hat{E}_y^2} \\
&\leqslant \sqrt{2}(2n+1)\sqrt{\mathrm{Tr}\,\hat{E}_y^2}.
\end{aligned} \qquad (37)$$

Here we have used the Schwarz inequality with respect to the inner product $\langle A, B\rangle := \sum_i \mathrm{Tr}\{\hat{A}_i\hat{B}_i^\dagger\}$ in the first inequality, taking $\hat{A}_i = \hat{q}_i\hat{\rho}\hat{q}_i$, $\hat{B}_i = \hat{E}_{y|\boldsymbol{\theta}}$. The second inequality follows from the fact that $x^2 + y^2 \leqslant (x+y)^2$ for $x, y > 0$. In the same way, considering an inner product $\langle A, B\rangle := \mathrm{Tr}\{\hat{A}\hat{B}^\dagger\}$ and taking $\hat{A} = \sum_i \hat{q}_i^2\hat{\rho}$ and $\hat{B} = \hat{E}_{y|\boldsymbol{\theta}}$, we get

$$\begin{aligned}
\sum_i \mathrm{Tr}\{\hat{q}_i^2\hat{\rho}\hat{E}_{y|\boldsymbol{\theta}}\} &\leqslant \sqrt{\mathrm{Tr}\left\{\left(\sum_i \hat{q}_i^2\right)^2\hat{\rho}^2\right\}}\sqrt{\mathrm{Tr}\,\hat{E}_y^2} \\
&\leqslant \sqrt{\langle\psi|\left(\sum_i \hat{q}_i^2\right)^2|\psi\rangle}\sqrt{\mathrm{Tr}\,\hat{E}_y^2} \\
&= \sqrt{\langle\psi|(2\hat{n}+1)^2|\psi\rangle}\sqrt{\mathrm{Tr}\,\hat{E}_y^2} \\
&\leqslant (2n+1)\sqrt{\mathrm{Tr}\,\hat{E}_y^2}.
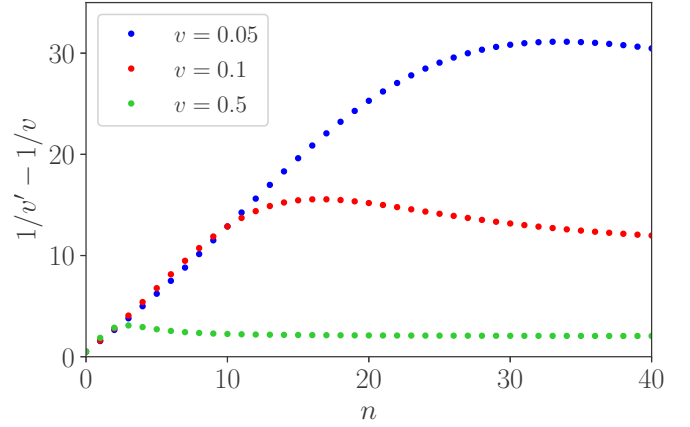\end{aligned} \qquad (38)$$

FIG. 6. The relation between n and $1/v' - 1/v$ for fixed values of v, when $\hat{\rho} = \hat{E}_y = |n\rangle\langle n|$.

Therefore, from Eq. (33), we obtain

$$F(y) \leqslant 2(\sqrt{2}+1)(2n+1)\frac{\sqrt{\mathrm{Tr}\,\hat{E}_y^2}}{p(y)}. \qquad (39)$$

Since the left-hand sides of Eqs. (37) and (38) are linear in $\hat{\rho}$, the bound, Eq. (39), also holds for arbitrary mixed $\hat{\rho}$. From Eqs. (39) and (36), one obtains a lower bound of v':

$$\frac{1}{v'} \leqslant \frac{1}{v} + (\sqrt{2}+1)(n+1/2)\frac{\sqrt{\mathrm{Tr}\,\hat{E}_y^2}}{p(y)}. \qquad (40)$$

The $O(n)$ upper bound of $F(y)$ is similar to the result in Ref. [21], which studies the case with no prior information and derives an $O(n)$ upper bound for the Fisher information. The authors state that whether or not the n scaling of the Fisher information is achievable is an open problem and argue that Fock states $|n\rangle$ do not achieve it due to the fringe of the likelihood function [the filter function $f_n(x, p)$ in our setting] [21]. Similarly, in our problem, if we take $\hat{\rho} = \hat{E}_y = |n\rangle\langle n|$ and increase n for fixed v, $1/v' - 1/v$ only behaves $\sim n$ up to some maximum n (which depends on v) and then decreases, while $p(y)$ also gradually decreases (Figs. 6 and 7). This
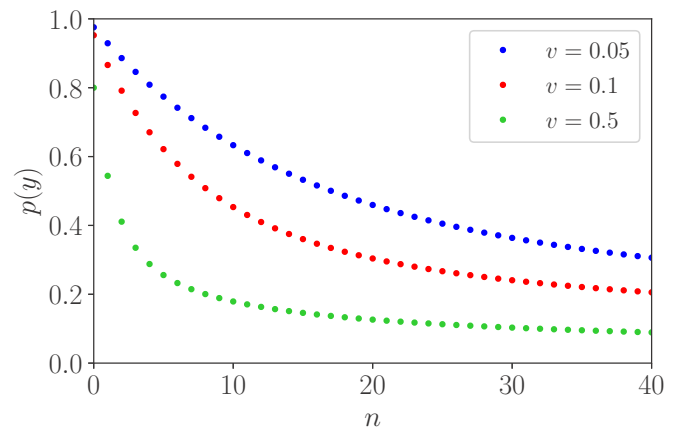

FIG. 7. The relation between n and $p(y)$ for fixed values of v, when $\hat{\rho} = \hat{E}_y = |n\rangle\langle n|$.

result is consistent with the observation that the filter function, Eq. (29), behaves like a Gaussian with variance $\sim n^{-1}$ only when v is smaller than this variance.

The bound, Eq. (40), also depends on the normalized post-selection probability density $\frac{p(y)}{\sqrt{\mathrm{Tr}\,\hat{E}_y^2}}$. This implies that one can possibly obtain a large $F(y)$ at the expense of a small postselection probability density. In fact, even when $n = 1$, one can construct an example where $1/v' - 1/v$ is inversely proportional to $\frac{p(y)}{\sqrt{\mathrm{Tr}\,\hat{E}_y^2}}$ when $\frac{p(y)}{\sqrt{\mathrm{Tr}\,\hat{E}_y^2}} \to 0$. See Appendix E for details.

VI. CONCLUSION

We have considered the problem of estimating the amount of random displacement that follows an isotropic Gaussian distribution. We have shown that there is a lower bound on the estimation error v' when only Gaussian states and Gaussian operations are used. This bound can be beaten for some range of the prior variance v, using only linear optics and simple non-Gaussian states such as single-photon states. When the maximum photon number of the input state is n, $1/v' - 1/v$ has an $O(n)$ upper bound which is also inversely proportional to the postselection probability.

Because of the similarity of the estimation of displacement to the error syndrome measurement in the GKP code [9], extending the method proposed in this paper may lead to an error correcting code using experimentally feasible non-Gaussian states such as single-photon states, which is important from a practical point of view. The Gaussian bound derived in this paper seems to correspond to the impossibility of correcting Gaussian errors using Gaussian states [8]. Finding a more direct connection between the estimation of displacement and the quantum error correction is a promising future study.

Another possible future subject is to find the best input state and POVM for a given constraint, e.g., maximum photon number. The bound, Eq. (40), is not necessarily tight, but for quantum estimation of displacement, many kinds of lower bounds for the estimation error are known [29,40,41], and some of them have been shown to be efficiently solvable [42,43]. It is possible that they can be extended to our setting.

APPENDIX A: DERIVATION OF GHOSH'S BOUND AND AN UPPER BOUND OF FISHER INFORMATION

We consider the estimation problem described in Sec. II. We write ξ, η as θ_1, θ_2 and $\hat{x}, \hat{p}$ as $\hat{q}_1, \hat{q}_2$ in this section, for convenience. We also just write (θ_1, θ_2) as $\boldsymbol{\theta}$ and $d\theta_1 d\theta_2$ as $d\boldsymbol{\theta}$.

Let us denote the estimated value of θ_i by $\tilde{\theta}_i(y)$. The Van Trees inequality [16,28,44], also called the Bayesian Cramér-Rao inequality, gives a lower bound for the Bayesian

estimation error. In a weaker form than the original matrix inequality, it can be written as

$$\sum_i \int dy \int d\boldsymbol{\theta} (\tilde{\theta}_i(y) - \theta_i)^2 p(y, \boldsymbol{\theta}) \geqslant \frac{4}{F_0^{(\mathrm{VT})} + F^{(\mathrm{VT})}}, \quad \text{(A1)}$$

where

$$
\begin{aligned}
F_0^{(\mathrm{VT})} &:= \sum_i \int d\boldsymbol{\theta} p(\boldsymbol{\theta}) \left(-\frac{\partial^2}{\partial \theta_i^2} \ln p(\boldsymbol{\theta}) \right), \\
F^{(\mathrm{VT})} &:= \sum_i \int dy \int d\boldsymbol{\theta} p(y, \boldsymbol{\theta}) \left(-\frac{\partial^2}{\partial \theta_i^2} \ln p(y|\boldsymbol{\theta}) \right).
\end{aligned}
\quad \text{(A2)}
$$

$F_0^{(\mathrm{VT})}$ is the *a priori* Fisher information of the prior distribution $p(\theta)$, and $F^{(\mathrm{VT})}$ is the expectation value of the Fisher information with respect to the prior distribution of θ, which can be considered as the average information obtained by the measurement. Equation (A1) can be used to evaluate the mean square error of the estimator, averaged over the measurement outcome y and the parameter θ. However, the Van Trees inequality cannot be directly applied to our case, because we allow postselection of y. To generalize the Van Trees inequality to the case including postselection, consider the equation

$$\sum_i \int (\tilde{\theta}_i(y) - \theta_i) \frac{\partial p(\boldsymbol{\theta}|y)}{\partial \theta_i} d\boldsymbol{\theta} = 2, \quad \text{(A3)}$$

which can be shown using a partial integration. Using the Schwarz inequality, one obtains

$$\sum_i \int (\tilde{\theta}_i(y) - \theta_i)^2 p(\boldsymbol{\theta}|y) d\boldsymbol{\theta} \geqslant \frac{4}{\tilde{F}(y)}, \quad \text{(A4)}$$

where $\tilde{F}(y)$ is

$$
\begin{aligned}
\tilde{F}(y) &:= \sum_i \int \frac{1}{p(\boldsymbol{\theta}|y)} \left(\frac{\partial p(\boldsymbol{\theta}|y)}{\partial \theta_i} \right)^2 d\boldsymbol{\theta} \\
&= \sum_i \int \left(\frac{\partial p(\boldsymbol{\theta}|y)}{\partial \theta_i} \right) \left(\frac{\partial}{\partial \theta_i} \ln p(\boldsymbol{\theta}|y) \right) d\boldsymbol{\theta} \\
&= \sum_i \int p(\boldsymbol{\theta}|y) \left(-\frac{\partial^2}{\partial \theta_i^2} \ln p(\boldsymbol{\theta}|y) \right) d\boldsymbol{\theta}.
\end{aligned}
\quad \text{(A5)}
$$

Equation (A4) is called Ghosh's bound [19] in statistics.

$\tilde{F}(y)$ can be further decomposed as

$$\tilde{F}(y) = F_0(y) + F(y), \quad \text{(A6)}$$

$$F_0(y) := \sum_i \int p(\boldsymbol{\theta}|y) \left(-\frac{\partial^2}{\partial \theta_i^2} \ln p(\boldsymbol{\theta}) \right) d\boldsymbol{\theta}, \quad \text{(A7)}$$

$$F(y) := \sum_i \int p(\boldsymbol{\theta}|y) \left(-\frac{\partial^2}{\partial \theta_i^2} \ln p(y|\boldsymbol{\theta}) \right) d\boldsymbol{\theta}. \quad \text{(A8)}$$

Equation (A4), together with Eqs. (A6), (A7), and (A8), gives the variant of the Van Trees inequality with postselection. $F_0^{(\mathrm{VT})}$ and $F^{(\mathrm{VT})}$ can be obtained as the expectation values of $F_0(y)$ and $F(y)$ with respect to $p(y)$. Therefore, if one takes the expectation values of both sides of Eq. (A4) with respect

to $p(y)$ and uses the convexity of the function $1/t$, one obtains

$$\sum_i \int dy \int d\boldsymbol{\theta} \big(\tilde{\theta}_i(y) - \theta_i\big)^2 p(y, \boldsymbol{\theta}) \geqslant \frac{4}{\int p(y)\tilde{F}(y)dy}, \quad \text{(A9)}$$

which reproduces Eq. (A1). To find an expression for $F(y)$ in our case, one transforms Eq. (A8) as

$$F(y) = \sum_i \int p(\boldsymbol{\theta}|y)\left(-\frac{\partial^2}{\partial\theta_i^2}\ln p(y|\boldsymbol{\theta})\right)d\boldsymbol{\theta}$$

$$= \sum_i \int \frac{p(\boldsymbol{\theta})}{p(y)}\left(\frac{1}{p(y|\boldsymbol{\theta})}\left(\frac{\partial p(y|\boldsymbol{\theta})}{\partial\theta_i}\right)^2 - \frac{\partial^2 p(y|\boldsymbol{\theta})}{\partial\theta_i^2}\right)d\boldsymbol{\theta}. \quad \text{(A10)}$$

Using that from Eq. (3),

$$p(y|\boldsymbol{\theta}) = \text{Tr}\{\hat{\rho}\hat{E}_{y|\boldsymbol{\theta}}\}, \quad \text{(A11)}$$

where

$$\hat{E}_{y|\boldsymbol{\theta}} := \hat{D}^\dagger(\boldsymbol{\theta})\hat{E}_y\hat{D}(\boldsymbol{\theta}), \quad \text{(A12)}$$

one can explicitly calculate terms in Eq. (A10) as

$$\sum_i \left(\frac{\partial}{\partial\theta_i}\text{Tr}\{\hat{\rho}\hat{E}_{y|\boldsymbol{\theta}}\}\right)^2 = -\sum_i \text{Tr}\{(\hat{q}_i\hat{\rho} - \hat{\rho}\hat{q}_i)\hat{E}_{y|\boldsymbol{\theta}}\}^2, \quad \text{(A13)}$$

$$\sum_i \frac{\partial^2}{\partial\theta_i^2}\text{Tr}\{\hat{\rho}\hat{E}_{y|\boldsymbol{\theta}}\} = \sum_i \text{Tr}\{\big(2\hat{q}_i\hat{\rho}\hat{q}_i - \hat{q}_i^2\hat{\rho} - \hat{\rho}\hat{q}_i^2\big)\hat{E}_{y|\boldsymbol{\theta}}\}, \quad \text{(A14)}$$

resulting in

$$F(y) = \sum_i \int d\boldsymbol{\theta}\frac{p(\boldsymbol{\theta})}{p(y)}\left[-\frac{\text{Tr}\{(\hat{q}_i\hat{\rho} - \hat{\rho}\hat{q}_i)\hat{E}_{y|\boldsymbol{\theta}}\}^2}{\text{Tr}\{\hat{\rho}\hat{E}_{y|\boldsymbol{\theta}}\}}\right.$$

$$\left. + \text{Tr}\{\big(-2\hat{q}_i\hat{\rho}\hat{q}_i + \hat{q}_i^2\hat{\rho} + \hat{\rho}\hat{q}_i^2\big)\hat{E}_{y|\boldsymbol{\theta}}\}\right]. \quad \text{(A15)}$$

An upper bound of $F(y)$ with a slightly simpler form can be obtained by applying the Schwarz inequality to

$$\{\hat{q}_i\hat{\rho}\hat{E}_{y|\boldsymbol{\theta}}\} = \text{Tr}\{\big(\hat{E}_{y|\boldsymbol{\theta}}^{1/2}\hat{q}_i\hat{\rho}^{1/2}\big)\cdot\big(\hat{\rho}^{1/2}\hat{E}_{y|\boldsymbol{\theta}}^{1/2}\big)\}, \quad \text{(A16)}$$

obtaining

$$-\frac{\text{Tr}\{(\hat{q}_i\hat{\rho} - \hat{\rho}\hat{q}_i)\hat{E}_{y|\boldsymbol{\theta}}\}^2}{\text{Tr}\{\hat{\rho}\hat{E}_{y|\boldsymbol{\theta}}\}} \leqslant \frac{4|\text{Tr}\{\hat{q}_i\hat{\rho}\hat{E}_{y|\boldsymbol{\theta}}\}|^2}{\text{Tr}\{\hat{\rho}\hat{E}_{y|\boldsymbol{\theta}}\}} \quad \text{(A17)}$$

$$\leqslant 4\text{Tr}\{\hat{q}_i\hat{\rho}\hat{q}_i\hat{E}_{y|\boldsymbol{\theta}}\},$$

which leads to

$$F(y) \leqslant \sum_i \int d\boldsymbol{\theta}\frac{p(\boldsymbol{\theta})}{p(y)}\text{Tr}\{\big(2\hat{q}_i\hat{\rho}\hat{q}_i + \hat{q}_i^2\hat{\rho} + \hat{\rho}\hat{q}_i^2\big)\hat{E}_{y|\boldsymbol{\theta}}\}. \quad \text{(A18)}$$

APPENDIX B: DERIVATION OF EQ. (18)

Suppose A, B are positive definite 2×2 matrices. Then we have

$$\det(A + B) = \det A \det(I + A^{-1/2}BA^{-1/2}), \quad \text{(B1)}$$

where I is the identity matrix. Since $A^{-1/2}BA^{-1/2}$ is positive definite and $\det(A^{-1/2}BA^{-1/2}) = \frac{\det B}{\det A}$, we can denote two

eigenvalues of $A^{-1/2}BA^{-1/2}$ as x and $\frac{1}{x}\frac{\det B}{\det A}$. Then we get

$$\det(I + A^{-1/2}BA^{-1/2}) = (1 + x)\left(1 + \frac{1}{x}\frac{\det B}{\det A}\right)$$

$$= 1 + \frac{\det B}{\det A} + x + \frac{1}{x}\frac{\det B}{\det A} \quad \text{(B2)}$$

$$\geqslant 1 + \frac{\det B}{\det A} + 2\sqrt{\frac{\det B}{\det A}}.$$
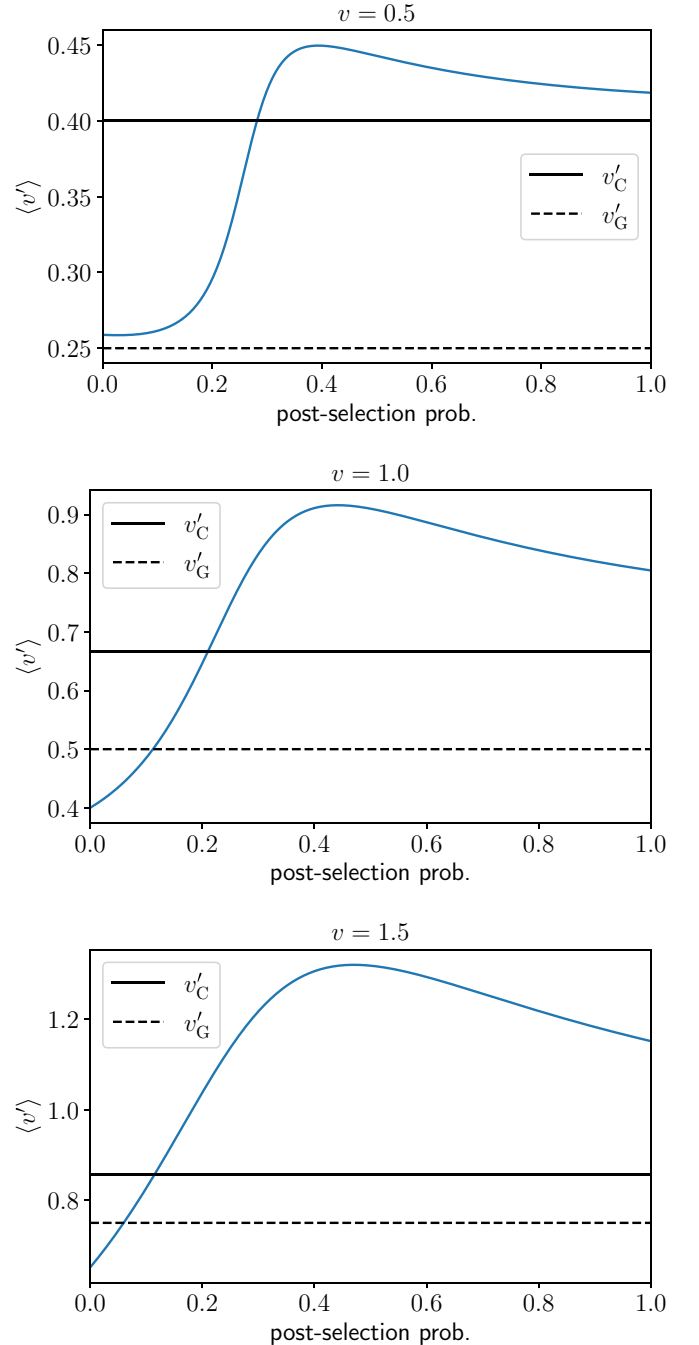
From Eqs. (B1) and (B2), Eq. (18) is obtained.

FIG. 8. The relation between the postselection probability [Eq. (C2)] and $\langle v'\rangle$ [Eq. (C1)] for $v = 0.5, 1.0, 1.5$, when we take $\hat{\rho} = \hat{\rho}' = |1\rangle\langle 1|$ in the setup in Sec. IV and postselect a range $y_x^2 + y_p^2 \leqslant r^2$ for $r > 0$.

APPENDIX C: ANALYSIS OF THE CASE WITH A FINITE POSTSELECTION PROBABILITY

We consider the setup in Sec. IV and take $\hat{\rho} = \hat{\rho}' = |1\rangle\langle 1|$. Rather than postselecting a single point $(y_x, y_p) = (0, 0)$, we postselect a range $y_x^2 + y_p^2 \leqslant r^2$ for $r > 0$. By increasing $r \to \infty$, the postselection probability converges to 1. We define the estimation error $\langle v' \rangle$ as the average of v' within the selected range of y,

$$\langle v' \rangle = \frac{\int_{y_x^2 + y_p^2 \leqslant r^2} v'(y_x, y_p) p(y_x, y_p) dy_x dy_p}{p(y_x^2 + y_p^2 \leqslant r^2)}, \qquad \text{(C1)}$$

where $p(y_x^2 + y_p^2 \leqslant r^2)$ is the postselection probability:

$$p(y_x^2 + y_p^2 \leqslant r^2) = \int_{y_x^2 + y_p^2 \leqslant r^2} p(y_x, y_p) dy_x dy_p. \qquad \text{(C2)}$$

In Fig. 8, the relation between the postselection probability and $\langle v' \rangle$ is plotted, for $v = 0.5, 1.0, 1.5$ (see also Fig. 3). It can be seen that the classical and Gaussian bounds are still beaten in some range of v, with a finite postselection probability.

$$\hat{B}^\dagger \hat{x}_1 \hat{B} = \frac{1}{\sqrt{2}}(\hat{x}_1 + \hat{x}_2), \quad \hat{B}^\dagger \hat{p}_1 \hat{B} = \frac{1}{\sqrt{2}}(\hat{p}_1 + \hat{p}_2),$$

$$\hat{B}^\dagger \hat{x}_2 \hat{B} = \frac{1}{\sqrt{2}}(\hat{x}_1 - \hat{x}_2), \quad \hat{B}^\dagger \hat{p}_2 \hat{B} = \frac{1}{\sqrt{2}}(\hat{p}_1 - \hat{p}_2). \qquad \text{(D3)}$$

Therefore, v' can be expressed as

$$v' = \frac{\int [(\xi - \tilde{\xi}(y))^2 + (\eta - \tilde{\eta}(y))^2] \exp(-\frac{\xi^2 + \eta^2}{v}) W_{\hat{\sigma}}(\xi/\sqrt{2}, \eta/\sqrt{2}) d\xi d\eta}{\int \exp(-\frac{\xi^2 + \eta^2}{v}) W_{\hat{\sigma}}(\xi/\sqrt{2}, \eta/\sqrt{2}) d\xi d\eta}$$

$$= \frac{\int 2[(x - x_0)^2 + (p - p_0)^2] \exp(-\frac{2}{v}(x^2 + p^2)) W_{\hat{\sigma}}(x, p) dx dp}{\int \exp(-\frac{2}{v}(x^2 + p^2)) W_{\hat{\sigma}}(x, p) dx dp}, \qquad \text{(D4)}$$

where we put $x_0 := \tilde{\xi}(y)/\sqrt{2}$, $p_0 := \tilde{\eta}(y)/\sqrt{2}$.

When $v = 2$, one can use the fact that

$$\frac{1}{\pi} \exp(-(x^2 + p^2)) = W_{|0\rangle\langle 0|}(x, p) \qquad \text{(D5)}$$

and

$$\frac{2}{\pi}[(x - x_0)^2 + (p - p_0)^2] \exp(-(x^2 + p^2))$$
$$= W_{|0\rangle\langle 0| + |\varphi\rangle\langle\varphi|}(x, p), \qquad \text{(D6)}$$

where $|\varphi\rangle = \alpha |0\rangle + |1\rangle$ and $\alpha = -\sqrt{2}(x_0 + ip_0)$. Thus, we obtain

$$v' = \frac{\text{Tr}[(|0\rangle\langle 0| + |\varphi\rangle\langle\varphi|)\hat{\sigma}]}{\text{Tr}[|0\rangle\langle 0|\hat{\sigma}]} \qquad \text{(D7)}$$
$$\geqslant 1.$$

The equality holds when $\langle\varphi|\hat{\sigma}|\varphi\rangle = 0$ for some α. The input and the POVM element being the same and pure, i.e., $\hat{\rho} \propto \hat{E}_y \propto |\psi\rangle\langle\psi|$ for some $|\psi\rangle$, is a sufficient condition for

APPENDIX D: PROOF THAT $v' \geqslant 1$ WHEN $v = 2$ FOR ARBITRARY STATES AND MEASUREMENTS

Because the convolution of two Wigner functions can be obtained by considering a half-beam-splitter interaction and tracing out one of the modes, the conditional probability, Eq. (5), can be expressed as

$$p(y|\xi, \eta) = 2\pi \int W_{\hat{\rho}}(x - \xi/2, p - \eta/2)$$
$$\times W_{\hat{E}_y}(x + \xi/2, p + \eta/2) dx dp$$
$$= \pi \int W_{\hat{\rho}}\left(\frac{x - \xi/\sqrt{2}}{\sqrt{2}}, \frac{p - \eta/\sqrt{2}}{\sqrt{2}}\right)$$
$$W_{\hat{E}_y}\left(\frac{x + \xi/\sqrt{2}}{\sqrt{2}}, \frac{p + \eta/\sqrt{2}}{\sqrt{2}}\right) dx dp$$
$$= \pi W_{\hat{\sigma}}(\xi/\sqrt{2}, \eta/\sqrt{2}), \qquad \text{(D1)}$$

where $\hat{\sigma}$ is a positive operator defined as

$$\hat{\sigma} := \text{Tr}_1[\hat{B}^\dagger(\hat{\rho} \otimes \hat{E}_y)\hat{B}]. \qquad \text{(D2)}$$

Here Tr_1 denotes the partial trace of the first system, and the half-beam-splitter operator $\hat{B}$ acts as

this, because

$$\langle 1|\hat{\sigma}|1\rangle \propto \int \left|\int \phi_1(x_2)\psi\left(\frac{x_1 + x_2}{\sqrt{2}}\right)\psi\left(\frac{x_1 - x_2}{\sqrt{2}}\right) dx_2\right|^2 dx_1$$
$$= 0, \qquad \text{(D8)}$$

where $\phi_1(x)$ is the wave function of $|1\rangle$. This holds because $\phi_1(x_2)$ is odd and $\psi(\frac{x_1 + x_2}{\sqrt{2}})\psi(\frac{x_1 - x_2}{\sqrt{2}})$ is even with respect to x_2.

APPENDIX E: ON THE DEPENDENCE OF EQ. (40) ON THE POSTSELECTION PROBABILITY $p(y)$

In Eq. (40), the upper bound of $1/v' - 1/v$ is inversely proportional to the normalized postselection probability density $\frac{p(y)}{\sqrt{\text{Tr}[\hat{E}_y^2]}}$. We consider cases where $p(y)$ is small. Because $p(y)$ is the overlap between the prior distribution $p(\xi, \eta)$ [Eq. (2)] and the likelihood function $p(\xi, \eta|y)$ [Eq. (3)], this can be done by taking $y = y_0 + \epsilon$ for small ϵ, where y_0 is a point such that $p(0, 0|y_0) = 0$, and taking $v \to 0$. We show that the linear dependence of the upper bound of $1/v' - 1/v$ on the

factor $\frac{\sqrt{\mathrm{Tr}[\hat{E}_y^2]}}{p(y)}$ in Eq. (40) is tight in the sense that we can construct an example such that

$$\lim_{\epsilon \to 0} \left[\lim_{v \to 0} \frac{p(y)}{\sqrt{\mathrm{Tr}[\hat{E}_y^2]}} \right] = 0, \tag{E1}$$

$$\lim_{\epsilon \to 0} \left[\lim_{v \to 0} (1/v' - 1/v) \frac{p(y)}{\sqrt{\mathrm{Tr}[\hat{E}_y^2]}} \right] > 0 \tag{E2}$$

for a small parameter ϵ, with a fixed value of the maximum photon number n.

In the heterodyne setting in Sec. IV, we take $\hat{\rho} = \hat{\rho}' = |1\rangle\langle 1|$. Then, for $y = (q, 0)$, the estimation error v' has the following expression:

$$v' = \frac{P_n(v, q)}{P_d(v, q)}, \tag{E3}$$

$$\begin{aligned}
P_n(v, q) = \; & 2v(16q^8 + 32q^6v^2 - 128q^6 + 32q^4v^4 + 48q^4v^3 \\
& - 96q^4v^2 + 64q^4v + 384q^4 + 16q^2v^6 + 80q^2v^5 \\
& + 160q^2v^4 + 256q^2v^3 + 256q^2v^2 - 256q^2v \\
& - 512q^2 + 3v^8 + 20v^7 + 56v^6 + 112v^5 \\
& + 192v^4 + 192v^3 + 128v^2 + 256v + 256),
\end{aligned} \tag{E4}$$

$$\begin{aligned}
P_d(v, q) = \; & (v + 2)(4q^4 + 4q^2v^2 - 16q^2 + v^4 + 4v^3 + 8v^2 \\
& + 16v + 16)^2.
\end{aligned} \tag{E5}$$

On the other hand, the postselection probability density $p(q, 0)$ is expressed as

$$\begin{aligned}
& p(q, 0) \\
& = \frac{(4q^4 + 4q^2v^2 - 16q^2 + v^4 + 4v^3 + 8v^2 + 16v + 16)e^{-\frac{q^2}{v+2}}}{\pi(v^5 + 10v^4 + 40v^3 + 80v^2 + 80v + 32)}.
\end{aligned} \tag{E6}$$

If we take the limit of $v \to 0$, we have

$$\lim_{v \to 0} \left(\frac{1}{v'} - \frac{1}{v} \right) = \frac{q^4 - 4q^2 + 12}{2(q^2 - 2)^2} \tag{E7}$$

and

$$\lim_{v \to 0} p(q, 0) = \frac{(q^2 - 2)^2 e^{-\frac{q^2}{2}}}{8\pi}. \tag{E8}$$

Then, if we put $q^2 = 2 + \epsilon$ and assume $|\epsilon| \ll 1$, we have $p(q, 0) \sim \epsilon^2$ and $(\frac{1}{v'} - \frac{1}{v}) \sim \epsilon^{-2}$. Using $\mathrm{Tr}\,\hat{E}_y^2 = 1/(2\pi)^2$, we have

$$\lim_{\epsilon \to 0} \left[\lim_{v \to 0} \frac{p(q, 0)}{\sqrt{\mathrm{Tr}[\hat{E}_{q,0}^2]}} \right] = 0, \tag{E9}$$

$$\lim_{\epsilon \to 0} \left[\lim_{v \to 0} (1/v' - 1/v) \frac{p(q, 0)}{\sqrt{\mathrm{Tr}[\hat{E}_{q,0}^2]}} \right] = \frac{1}{e}, \tag{E10}$$

which is the desired property.

[1] W. Asavanant, Y. Shiozawa, S. Yokoyama, B. Charoensombutamon, H. Emura, R. N. Alexander, S. Takeda, J. Yoshikawa, N. C. Menicucci, H. Yonezawa, and A. Furusawa, Generation of time-domain-multiplexed two-dimensional cluster state, Science **366**, 373 (2019).

[2] M. V. Larsen, X. Guo, C. R. Breum, J. S. Neergaard-Nielsen, and U. L. Andersen, Deterministic generation of a two-dimensional cluster state, Science **366**, 369 (2019).

[3] W. Asavanant, B. Charoensombutamon, S. Yokoyama, T. Ebihara, T. Nakamura, R. N. Alexander, M. Endo, J. Yoshikawa, N. C. Menicucci, H. Yonezawa, and A. Furusawa, Time-Domain-Multiplexed Measurement-Based Quantum Operations with 25-MHz Clock Frequency, Phys. Rev. Appl. **16**, 034005 (2021).

[4] M. V. Larsen, X. Guo, C. R. Breum, J. S. Neergaard-Nielsen, and U. L. Andersen, Deterministic multi-mode gates on a scalable photonic quantum computing platform, Nat. Phys. **17**, 1018 (2021).

[5] D. Gottesman, An introduction to quantum error correction and fault-tolerant quantum computation, in *Quantum Information Science and Its Contributions to Mathematics. Proceedings of Symposia in Applied Mathematics*, Vol. 68 (American Mathematical Society, Providence, RI, 2010), pp. 13–58.

[6] J. Harrington and J. Preskill, Achievable rates for the Gaussian quantum channel, Phys. Rev. A **64**, 062301 (2001).

[7] J. Eisert, S. Scheel, and M. B. Plenio, Distilling Gaussian States with Gaussian Operations is Impossible, Phys. Rev. Lett. **89**, 137903 (2002).

[8] J. Niset, J. Fiurášek, and N. J. Cerf, No-Go Theorem for Gaussian Quantum Error Correction, Phys. Rev. Lett. **102**, 120501 (2009).

[9] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, Phys. Rev. A **64**, 012310 (2001).

[10] C. Flühmann, T. L. Nguyen, M. Marinelli, V. Negnevitsky, K. Mehta, and J. Home, Encoding a qubit in a trapped-ion mechanical oscillator, Nature (London) **566**, 513 (2019).

[11] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf *et al.*, Quantum error correction of a qubit encoded in grid states of an oscillator, Nature (London) **584**, 368 (2020).

[12] B. M. Terhal, J. Conrad, and C. Vuillot, Towards scalable bosonic quantum error correction, Quantum Sci. Technol. **5**, 043001 (2020).

[13] S. Personick, Application of quantum estimation theory to analog communication over quantum channels, IEEE Trans. Info. Theory **17**, 240 (1971).

[14] H. Yuen and M. Lax, Multiple-parameter quantum estimation and measurement of nonselfadjoint observables, IEEE Trans. Info. Theory **19**, 740 (1973).

[15] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, Vol. 1 (Springer Science & Business Media, New York, 2011).

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

[17] C. W. Helstrom, *Quantum Detection and Estimation Theory*, Vol. 84 (Academic Press, New York, 1976).

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

[19] M. Ghosh, Cramér-Rao bounds for posterior variances, Stat. Prob. Lett. **17**, 173 (1993).

[20] Y. Li, L. Pezzè, M. Gessner, Z. Ren, W. Li, and A. Smerzi, Frequentist and Bayesian quantum phase estimation, Entropy **20**, 628 (2018).

[21] K. Duivenvoorden, B. M. Terhal, and D. Weigand, Single-mode displacement sensor, Phys. Rev. A **95**, 012305 (2017).

[22] S. M. Assad, J. Li, Y. Liu, N. Zhao, W. Zhao, P. K. Lam, Z. Y. Ou, and X. Li, Accessible precisions for estimating two conjugate parameters using Gaussian probes, Phys. Rev. Research **2**, 023182 (2020).

[23] M. Bradshaw, P. K. Lam, and S. M. Assad, Ultimate precision of joint quadrature parameter estimation with a Gaussian probe, Phys. Rev. A **97**, 012106 (2018).

[24] M. G. Genoni, M. G. A. Paris, G. Adesso, H. Nha, P. L. Knight, and M. S. Kim, Optimal estimation of joint parameters in phase space, Phys. Rev. A **87**, 012107 (2013).

[25] C. Oh, K. Park, R. Filip, H. Jeong, and P. Marek, Optical estimation of unitary Gaussian processes without phase reference using Fock states, New J. Phys. **22**, 123039 (2020).

[26] K. C. McCormick, J. Keller, S. C. Burd, D. J. Wineland, A. C. Wilson, and D. Leibfried, Quantum-enhanced sensing of a single-ion mechanical oscillator, Nature (London) **572**, 86 (2019).

[27] U. Leonhardt, *Essential Quantum Optics: From Quantum Measurements to Black Holes* (Cambridge University Press, Cambridge, UK, 2010).

[28] R. D. Gill and B. Y. Levit, Applications of the Van Trees inequality: A Bayesian Cramér-Rao bound, Bernoulli **1**, 59 (1995).

[29] M. Tsang, Physics-inspired forms of the Bayesian Cramér-Rao bound, Phys. Rev. A **102**, 062217 (2020).

[30] R. Jackiw, Minimum uncertainty product, number-phase uncertainty product, and coherent states, J. Math. Phys. **9**, 339 (1968).

[31] N. Walker and J. E. Carroll, Multiport homodyne detection near the quantum noise limit, Opt. Quantum Electron. **18**, 355 (1986).

[32] S. Takeda, T. Mizuta, M. Fuwa, J. Yoshikawa, H. Yonezawa, and A. Furusawa, Generation and eight-port homodyne characterization of time-bin qubits for continuous-variable quantum information processing, Phys. Rev. A **87**, 043803 (2013).

[33] K. Wódkiewicz, Operational Approach to Phase-Space Measurements in Quantum Mechanics, Phys. Rev. Lett. **52**, 1064 (1984).

[34] K. Miyata, H. Ogawa, P. Marek, R. Filip, H. Yonezawa, J. Yoshikawa, and A. Furusawa, Implementation of a quantum cubic gate by an adaptive non-Gaussian measurement, Phys. Rev. A **93**, 022301 (2016).

[35] A. Ourjoumtsev, R. Tualle-Brouri, and P. Grangier, Quantum Homodyne Tomography of a Two-Photon Fock State, Phys. Rev. Lett. **96**, 213601 (2006).

[36] M. Yukawa, K. Miyata, T. Mizuta, H. Yonezawa, P. Marek, R. Filip, and A. Furusawa, Generating superposition of up to three photons for continuous variable quantum information processing, Opt. Express **21**, 5529 (2013).

[37] L. Gatteschi, Asymptotics and bounds for the zeros of Laguerre polynomials: Survey, J. Comput. Appl. Math. **144**, 7 (2002) [selected papers of the International Symposium on Applied Mathematics, August 2000, Dalian, China].

[38] E. Martínez-Vargas, C. Pineda, F. Leyvraz, and P. Barberis-Blostein, Quantum estimation of unknown parameters, Phys. Rev. A **95**, 012136 (2017).

[39] H. Cramér, *Mathematical Methods of Statistics*, Vol. 43 (Princeton University Press, Princeton, NJ, 1999).

[40] H. Nagaoka, A new approach to Cramér-Rao bounds for quantum state estimation, in *Asymptotic Theory Of Quantum Statistical Inference: Selected Papers* (World Scientific, Singapore, 2005), pp. 100–112.

[41] J. Rubio and J. Dunningham, Bayesian multiparameter quantum metrology with limited data, Phys. Rev. A **101**, 032114 (2020).

[42] L. Conlon, J. Suzuki, P. K. Lam, and S. M. Assad, Efficient computation of the Nagaoka–Hayashi bound for multi-parameter estimation with separable measurements, npj Quantum Inf. **7**, 110 (2021).

[43] J. S. Sidhu, Y. Ouyang, E. T. Campbell, and P. Kok, Tight bounds on the simultaneous estimation of incompatible parameters, Phys. Rev. X **11**, 011028 (2021).

[44] H. L. Van Trees, *Detection, Estimation, and Linear Modulation Theory, Part I* (Wiley, New York, 1968).