Superiority of heterodyning over homodyning: An assessment with quadrature moments

Y. S. Teo,[1] C. R. Müller,[2] H. Jeong,[3] Z. Hradil,[4] J. Řeháček,[4] and L. L. Sánchez-Soto[2,5]

[1]*BK21 Frontier Physics Research Division, Seoul National University, 08826 Seoul, South Korea*
[2]*Max-Planck-Institut für die Physik des Lichts, Staudtstraße 2, 91058 Erlangen, Germany*
[3]*Center for Macroscopic Quantum Control, Seoul National University, 08826 Seoul, South Korea*
[4]*Department of Optics, Palacký University, 17. listopadu 12, 77146 Olomouc, Czech Republic*
[5]*Departamento de Óptica, Facultad de Física, Universidad Complutense, 28040 Madrid, Spain*

We examine the moment-reconstruction performance of both the homodyne and heterodyne (double-homodyne) measurement schemes for arbitrary quantum states and introduce moment estimators that optimize the respective schemes for any given data. In the large-data limit, these estimators are as efficient as the maximum-likelihood estimators. We then illustrate the superiority of the heterodyne measurement for the reconstruction of the first and second moments by analyzing Gaussian states and many other significant nonclassical states. Finally, we present an extension of our theories to two-mode sources, which can be straightforwardly generalized to all other multimode sources.

I. INTRODUCTION

The next-generation quantum technologies introduce novel and innovative routes to the understanding and implementation of measurements, communication, and computation. In this respect, the manipulation of a quantum light source using continuous-variable (CV) measurements offers many advantages [1–5]. There exist two standard CV measurement schemes. The more commonly employed homodyne detection [6–8], which performs an approximate measurement of rotated photonic quadratures [9], probes the marginal distribution of the Wigner function of the unknown quantum state [10]. The other less widely adopted double-homodyne detection, or the heterodyne detection, involves the joint measurement of complementary observables [11–18] that directly samples the phase space according to the Husimi function [19] and is connected to the conventional heterodyne scheme [20–29]. This measurement also has fundamental influence on quantum-mechanical uncertainties, which were also investigated in a different context [30].

These measurement schemes, which experimentally probe quasiprobability distributions, can also be equivalently understood as practical means to directly characterize the source in terms of the ordered moments of the quadrature operators in phase space. Gaussian states [2], for example, which are important in analyzing CV quantum information processing [31–34], are conveniently described by this representation since all their operator moments are functions of only the first and second moments. Therefore, estimating the first and second moments is enough to fully reconstruct the Gaussian state or verify if the reconstructed state is accurately Gaussian [35]. Higher moments come into play for general quantum states. On its own right, the topic of operator moments of quantum states draws interest in the context of generalized uncertainty relations [36,37], nonclassicality detection [38,39], entanglement detection [40,41], and cryptography [42,43].

In Refs. [44,45], we theoretically and experimentally compared the two measurement schemes, using a polarization-squeezing setup [46–51] for the latter. We analyzed the physical implications of having the unavoidable Arthurs-Kelly–type noise that is inherent in the joint measurement heterodyne scheme on moment reconstruction. We found that despite this additional noise, for a single-mode central-Gaussian source the heterodyne scheme still results in second-moment estimators that are more accurate than the homodyne scheme for a wide range of the squeezing strength and temperature parameter.

In this article, we extend the theory of these two CV measurement schemes to general quantum states and show that the tomographic advantage in using the heterodyne scheme carries over to other interesting and important non-Gaussian states. This message is conveyed in five main sections. Section II gives an overview of the fundamental elements in first- and second-moment tomography, as well as the concept of reconstruction accuracy. These elements are then used to discuss the general theory of moment reconstruction for the homodyne and heterodyne schemes in Sec. III. In that section, we shall also introduce optimal moment estimators that asymptotically approach the respective Cramér-Rao bounds, which are derived in Appendix A. In Sec. IV, we shall study the CV schemes in first-moment estimation where it shall be shown that heterodyne detection will always outperform homodyne detection unless the state is of minimum uncertainty, in which case the two schemes give equal reconstruction accuracy per sampling event. This result shall be discussed with some interesting classes of non-Gaussian states. Next, we study the results for second-moment estimation in Sec. V with the same classes of non-Gaussian states and illustrate once again the tomographic advantages of using the heterodyne scheme in moment tomography. Finally, Sec. VII concludes the presented results in a summary.

II. COVARIANCE MATRIX AND MOMENT-RECONSTRUCTION ACCURACY

In dealing with single-mode bosonic systems, such as photons, for the pair of position X and momentum P quadrature operators obeying $[X,P] = i$ (with the quantum unit $\hbar \equiv 1$) that form the column $\boldsymbol{R} = (X \; P)^{\mathrm{T}}$, the covariance

matrix can be written as

$$\boldsymbol{G} = \mathrm{Re}\{\langle \boldsymbol{R}\boldsymbol{R}^{\mathrm{T}}\rangle\} - \langle \boldsymbol{R}\rangle\langle \boldsymbol{R}\rangle^{\mathrm{T}} = \boldsymbol{G}_2 - \boldsymbol{G}_1, \qquad (2.1)$$

where we have introduced the first-moment ($\boldsymbol{G}_1 = \langle \boldsymbol{R}\rangle\langle \boldsymbol{R}\rangle^{\mathrm{T}}$) and second-moment ($\boldsymbol{G}_2 = \mathrm{Re}\{\langle \boldsymbol{R}\boldsymbol{R}^{\mathrm{T}}\rangle\}$) matrices. The two independent parameters $\{\langle X\rangle, \langle P\rangle\}$ in $\boldsymbol{G}_1$ and three independent parameters $\{\langle X^2\rangle, \frac{1}{2}\langle\{X,P\}\rangle, \langle P^2\rangle\}$ in $\boldsymbol{G}_2$ constitute the complete set of five parameters that characterize $\boldsymbol{G}$. The well-known class of Gaussian states possesses a Gaussian Wigner function or any other kind of well-behaved quasiprobability distribution. As a consequence, any Gaussian state is fully described by only $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$.

The covariance matrix for any quantum state obeys the inequality $\boldsymbol{G} \geqslant -i\,\boldsymbol{\Omega}/2$ in terms of $\boldsymbol{\Omega} \,\widehat{=}\, \bigl(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\bigr)$ that is related to the two-dimensional symplectic group, which is a recast of the Heisenberg-Robertson-Schrödinger (HRS) uncertainty relation for position and momentum operators. This gives the equivalent stricter inequality $\det\{\boldsymbol{G}\} \geqslant \frac{1}{4}$ in addition to the standard positivity constraint for $\boldsymbol{G}$. The reconstruction of the full covariance matrix $\boldsymbol{G}$ involves the quantum tomography of all the five independent parameters that define the first and second operator moments of the state. Here, the figure of merit of the reconstruction accuracy is the mean-squared error (MSE) $\mathscr{D} = \mathbb{E}[\mathrm{Tr}\{(\widehat{\boldsymbol{G}} - \boldsymbol{G})^2\}]$ between $\boldsymbol{G}$ and its estimator $\widehat{\boldsymbol{G}}$. In terms of $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$,

$$\mathscr{D} = \underbrace{\mathbb{E}[\mathrm{Tr}\{(\widehat{\boldsymbol{G}}_1 - \boldsymbol{G}_1)^2\}]}_{= \,\mathscr{D}_1} + \underbrace{\mathbb{E}[\mathrm{Tr}\{(\widehat{\boldsymbol{G}}_2 - \boldsymbol{G}_2)^2\}]}_{= \,\mathscr{D}_2}$$
$$+ \{\text{cross terms}\}. \qquad (2.2)$$

To illustrate the physics behind moment reconstruction, we shall analyze both the first- and second-moment reconstruction accuracy separately. In practice, these analyses are relevant to the situation where the reconstructions of $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ are carried out with independent data. For this situation, the {cross terms} in Eq. (2.2) vanish so that the total MSE is the sum of the respective MSEs $\mathscr{D}_1$ and $\mathscr{D}_2$ of the reconstructed moments. From hereon, to facilitate discussions, we shall analyze the quantity $\boldsymbol{r} = \langle \boldsymbol{R}\rangle$ in place of $\boldsymbol{G}_1$, where $\mathscr{D}_1 = \mathbb{E}[(\widehat{\boldsymbol{r}} - \boldsymbol{r})^2]$.

In unbiased statistical estimation theory [52], the MSE $\mathscr{D} \geqslant \mathrm{Tr}\{\boldsymbol{F}^{-1}\}$ is bounded from below by the inverse of the Fisher information matrix $\boldsymbol{F}$, or the Cramér-Rao bound (CRB). Consequently, we have the respective first- and second-moment CRBs $\mathscr{D}_1 \geqslant \mathrm{Tr}\{\boldsymbol{F}_1^{-1}\}$ and $\mathscr{D}_2 \geqslant \mathrm{Tr}\{\boldsymbol{F}_2^{-1}\}$. Therefore, the general theory of the Fisher matrices $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ for the two CV schemes is in order.

Two crucial factors that affect the CRB are the type of measurement used for reconstruction, as well as the measurement resources or number of sampling events N detected by this measurement. In this article, we are interested in comparing the reconstruction accuracies of different measurement types. Therefore, the important quantities of interest are the scaled versions of the CRBs, $\mathscr{H}_1 = N\,\mathrm{Tr}\{\boldsymbol{F}_1^{-1}\}$ and $\mathscr{H}_2 = N\,\mathrm{Tr}\{\boldsymbol{F}_2^{-1}\}$, or their respective scaled Fisher matrices $\widetilde{\boldsymbol{F}}_1 = \boldsymbol{F}_1/N$ and $\widetilde{\boldsymbol{F}}_2 = \boldsymbol{F}_2/N$.
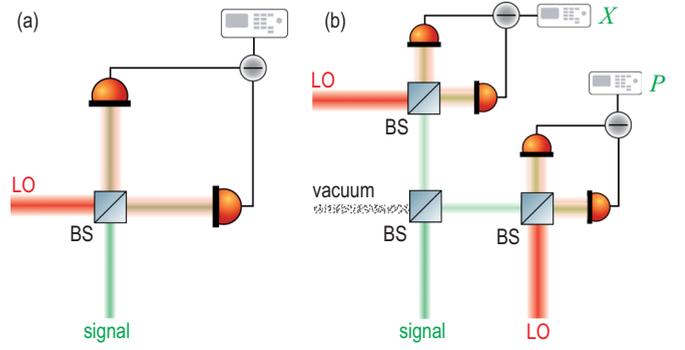
FIG. 1. Schema for the (a) homodyne and (b) heterodyne setups. Here, BS denotes the beam splitter and LO denotes the local oscillator.

III. GENERAL THEORY

A. Homodyne detection

The homodyne detection [6–8] involves a 50:50 beam splitter that introduces an interference between the optical source of an unknown state (signal) and the local oscillator (coherent-state reference source), the latter of which is set to a much larger optical intensity than the mean intensity of the optical source of an unknown quantum state ρ [see Fig. 1(a)]. A subtraction of the output photocurrents gives a distribution of voltage readouts $-\infty < x_\vartheta < \infty$ for the local-oscillator phase $0 \leqslant \vartheta \leqslant \pi$, which essentially corresponds to the eigenvalue probability distribution of the quadrature operator $X_\vartheta = X \cos\vartheta + P \sin\vartheta$. It then follows that statistically, the expectation value $\langle X_\vartheta^m\rangle$ for any integer value m contains all measurable information about the mth operator moments of X and P. Since the data acquired with this scheme are the marginals of the Wigner function, the first ($m = 1$) and second ($m = 2$) moments, or $\boldsymbol{G}$, that are reconstructed with these data may be attributed to this quasiprobability distribution function.

In a typical homodyne experiment, the value of ϑ is set to increase linearly. The data collected would then be binned for all the measured ϑ values. The data bins are mutually independent, so that the Fisher matrices $\boldsymbol{F}_{1,\mathrm{hom}}$ and $\boldsymbol{F}_{2,\mathrm{hom}}$ for the respective first- and second-moment CRBs can each be understood as a summation of Fisher matrices of every local-oscillator phase bin according to the additivity property of the Fisher information. In the limit of large number of sampling events N, the central limit theorem states that the unbiased estimator $\widehat{\langle X_\vartheta^m\rangle}$ of the mth quadrature moment $\langle X_\vartheta^m\rangle$ that is defined as an average sum of independently collected random voltage values for the phase ϑ follows a Gaussian distribution of data mean $\mu = \mu(\vartheta) = \langle X_\vartheta^m\rangle$ and data variance σ^2/N where $\sigma^2 = \sigma(\vartheta)^2 = \langle X_\vartheta^{2m}\rangle - \langle X_\vartheta^m\rangle^2$, so that the Fisher matrix

$$\boldsymbol{F}_{\vartheta,m} = \frac{N}{\sigma^2}\frac{\partial\mu}{\partial\boldsymbol{a}}\frac{\partial\mu}{\partial\boldsymbol{a}} + \frac{1}{2\sigma^4}\frac{\partial\sigma^2}{\partial\boldsymbol{a}}\frac{\partial\sigma^2}{\partial\boldsymbol{a}} \qquad (3.1)$$

for a given phase ϑ in the large-N limit follows the well-known expression for Gaussian distributions, where in our case $\boldsymbol{a}$ is the column of mth moment parameters we are interested in reconstructing. As it is clear that only the first term of (3.1) would survive in this limit, we thus have the *scaled* homodyne

Fisher matrix

$$\widetilde{\boldsymbol{F}}_{m,\mathrm{hom}} = \int_{(\pi)} \frac{d\vartheta}{\pi} \frac{\boldsymbol{F}_{\vartheta,m}}{N} = \int_{(\pi)} \frac{d\vartheta}{\pi} \frac{1}{\sigma(\vartheta)^2} \frac{\partial\mu(\vartheta)}{\partial\boldsymbol{a}} \frac{\partial\mu(\vartheta)}{\partial\boldsymbol{a}} \tag{3.2}$$

with respect to the number of sampling events N for the complete set of homodyne quadrature-eigenstate outcomes.

1. First-moment reconstruction

All information about the first moments, $\boldsymbol{a} = \boldsymbol{r}$, of the covariance matrix is completely encoded in the expectation value $\mu(\vartheta) = \langle X_\vartheta \rangle$. The variance for the data is then given by $\sigma(\vartheta)^2 = \langle X_\vartheta^2 \rangle - \langle X_\vartheta \rangle^2$. The scaled Fisher matrix for the first-moment estimation with homodyne data is therefore given by

$$\widetilde{\boldsymbol{F}}_{1,\mathrm{hom}} = \int_{(\pi)} \frac{d\vartheta}{\pi} \frac{\boldsymbol{m}_\vartheta}{\langle X_\vartheta^2 \rangle - \langle X_\vartheta \rangle^2}, \tag{3.3}$$

where

$$\boldsymbol{m}_\vartheta = \begin{pmatrix} (\cos\vartheta)^2 & \sin\vartheta\cos\vartheta \\ \sin\vartheta\cos\vartheta & (\sin\vartheta)^2 \end{pmatrix}. \tag{3.4}$$

The integral can be evaluated exactly, bringing us to the closed-form expression

$$\mathscr{H}_{1,\mathrm{hom}} = \mathrm{Tr}\{\boldsymbol{G}\} + 2\sqrt{\det\{\boldsymbol{G}\}}. \tag{3.5}$$

With the machinery of quantum tomography (see Appendix A 1), an observer can construct the optimal moment estimator that achieves the scaled CRB. Suppose that the observer collects homodyne data for N sampling events and bins the voltage values into $\{x_{jk}\}$ according to a discrete number n_ϑ of phase bins ϑ_k, where j labels the n_x real voltage values per phase bin $\vartheta = \vartheta_k$ and k labels the phase bins. Then, an unbiased estimator for any particular expectation value $\langle X_k \rangle \equiv \langle X_{\vartheta_k} \rangle$ would be

$$\widehat{\langle X_k \rangle} = \frac{1}{N} \sum_{j=1}^{n_x} n_{jk} x_{jk}, \quad \sum_{k=1}^{n_\vartheta} \sum_{j=1}^{n_x} n_{jk} = \sum_{k=1}^{n_\vartheta} N_k = N. \tag{3.6}$$

Upon denoting $\boldsymbol{u}_k \equiv \boldsymbol{u}_{\vartheta_k} = (\cos\vartheta_k \ \ \sin\vartheta_k)^{\mathrm{T}}$ where we note that $\boldsymbol{m}_k = \boldsymbol{u}_k \boldsymbol{u}_k^{\mathrm{T}}$, the optimal first-moment estimator is given by

$$\widehat{\boldsymbol{r}}_{\mathrm{hom}}^{(\mathrm{opt})} = \boldsymbol{W}_1^{-1} \sum_{k=1}^{n_\vartheta} \boldsymbol{u}_k \frac{N_k \widehat{\langle X_k \rangle}}{\widehat{\langle X_k^2 \rangle} - \widehat{\langle X_k \rangle}^2},$$

$$\boldsymbol{W}_1 = \sum_{k=1}^{n_\vartheta} \boldsymbol{m}_k \frac{N_k}{\widehat{\langle X_k^2 \rangle} - \widehat{\langle X_k \rangle}^2}, \tag{3.7}$$

which is immediately computable given the processed data $\{\widehat{\langle X_k \rangle}\}$ and $\{\widehat{\langle X_k^2 \rangle}\}$ that are defined by

$$\widehat{\langle X_k^m \rangle} = \frac{1}{N} \sum_{j=1}^{n_x} n_{jk} x_{jk}^m \quad (m = 1, 2, \ldots). \tag{3.8}$$

That this estimator achieves the scaled CRB asymptotically is also shown in Appendix A 1. This equivalently implies that the optimal estimator is as efficient as the maximum-likelihood estimator for the multinomially distributed binned data $\{x_{jk}\}$.

2. Second-moment reconstruction

To estimate $\boldsymbol{G}_2$, it is clear that second-moment information is completely encoded in the second quadrature moment $\langle X_\vartheta^2 \rangle$, which is a function of the three independent parameters $a_1 = \langle X^2 \rangle$, $a_2 = \langle \frac{1}{2}\{\Delta X, \Delta P\}\rangle$, and $a_3 = \langle P^2 \rangle$. From Eq. (3.2), the corresponding 3×3 Fisher matrix for these three parameters is

$$\widetilde{\boldsymbol{F}}_{2,\mathrm{hom}} = \int_{(\pi)} \frac{d\vartheta}{\pi} \frac{\boldsymbol{M}_\vartheta}{\langle X_\vartheta^4 \rangle - \langle X_\vartheta^2 \rangle^2}, \tag{3.9}$$

where

$$\boldsymbol{M}_\vartheta \,\widehat{=}\, \begin{pmatrix} (\cos\vartheta)^2 \\ \sqrt{2}\sin\vartheta\cos\vartheta \\ (\sin\vartheta)^2 \end{pmatrix}$$

$$\times ((\cos\vartheta)^2 \quad \sqrt{2}\sin\vartheta\cos\vartheta \quad (\sin\vartheta)^2). \tag{3.10}$$

The analytical answer to $\widetilde{\boldsymbol{F}}_{2,\mathrm{hom}}$ for an arbitrary state, and its subsequent inverse $\mathscr{H}_{2,\mathrm{hom}} = \mathrm{Tr}\{\widetilde{\boldsymbol{F}}_{2,\mathrm{hom}}^{-1}\}$, is difficult to calculate, as the denominator in the integrand generally contains trigonometric functions in a complicated manner. Nevertheless, the integral can be calculated explicitly for many interesting and important quantum sources.

The optimal second-moment estimator (see Appendix A 2) that achieves the corresponding scaled CRB can be cleanly expressed using the vectorization operation $\mathrm{vec}(\boldsymbol{Y})$ that turns a matrix into a column according to

$$\boldsymbol{Y} \,\widehat{=}\, \begin{pmatrix} y_1 & y_2 \\ y_2 & y_3 \end{pmatrix} \quad \mapsto \quad \mathrm{vec}(\boldsymbol{Y}) \,\widehat{=}\, \begin{pmatrix} y_1 \\ \sqrt{2}\,y_2 \\ y_3 \end{pmatrix} \tag{3.11}$$

in any prechosen basis, such that $\mathrm{Tr}\{\boldsymbol{Y}_1 \boldsymbol{Y}_2\} = \mathrm{vec}(\boldsymbol{Y}_1)^{\mathrm{T}} \mathrm{vec}(\boldsymbol{Y}_2)$ for any two 2×2 symmetric matrices $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$. Given the processed data defined in Eq. (3.8), the final operationally ready expressions for this optimal estimator are given as follows:

$$\widehat{\boldsymbol{G}}_{2,\mathrm{hom}}^{(\mathrm{opt})} = \boldsymbol{W}_2^{-1} \sum_{k=1}^{n_\vartheta} \mathrm{vec}(\boldsymbol{m}_k) \frac{N_k \widehat{\langle X_k^2 \rangle}}{\widehat{\langle X_k^4 \rangle} - \widehat{\langle X_k^2 \rangle}^2},$$

$$\boldsymbol{W}_2 = \sum_{k=1}^{n_\vartheta} \boldsymbol{M}_k \frac{N_k}{\widehat{\langle X_k^4 \rangle} - \widehat{\langle X_k^2 \rangle}^2}. \tag{3.12}$$

For accurate tomography, the value of N is typically large enough such that $\widehat{\boldsymbol{G}}_{2,\mathrm{hom}}^{(\mathrm{opt})}$ is a proper covariance matrix and approaches the maximum-likelihood estimator that asymptotically achieves the scaled CRB, which is strictly speaking the correct regime where $\widehat{\boldsymbol{G}}_{2,\mathrm{hom}}^{(\mathrm{opt})}$ is to be used for second-moment tomography. On a separate note, optimal estimators for overcomplete quantum-state tomography of ρ was developed in [53] and later rederived in [54] with the variational principle that is also used to construct the optimal moment estimators in Appendix A.

B. Heterodyne detection

The heterodyne detection scheme essentially uses two homodyne setups to perform a joint measurement of two complementary observables [see Fig. 1(b)], which are in this case chosen to be the standard X and P quadrature pair for convenience. It is well known [11–18] that the product of their joint-measurement standard deviations has a larger lower bound than the usual one-half of a quantum unit given by the original Heisenberg relation owing to the additional quantum noise introduced by the joint measurement.

The outcomes for this scheme are in fact the overcomplete set of coherent states. This means that the resulting data are direct phase-space samples of the Husimi function for the statistical operator ρ. The technical complication of having additional measurement noise can therefore be translated completely into the phase-space language that is relevant in our subsequent analysis. Given an infinite set of the Husimi-function data, we have access to the moments $\overline{x^k p^l}$ (the overline denotes the average with respect to the Husimi function, or simply the Husimi average), with which the corresponding "$\boldsymbol{G}$" operator

$$\boldsymbol{G}_{\text{het}} \,\widehat{=}\, \begin{pmatrix} \overline{x^2} - \overline{x}^2 & \overline{xp} - \overline{x}\,\overline{p} \\ \overline{xp} - \overline{x}\,\overline{p} & \overline{p^2} - \overline{p}^2 \end{pmatrix} \quad (3.13)$$

can be directly constructed. One can then show that for any quantum state,

$$\boldsymbol{G}_{\text{het}} = \boldsymbol{G} + \frac{1}{2}. \quad (3.14)$$

The corresponding Arthurs-Kelly–type measurement uncertainty relation

$$\text{Var}_{\text{Q}}[x]\text{Var}_{\text{Q}}[p] = \left[\langle(\Delta X)^2\rangle + \frac{1}{2}\right]\left[\langle(\Delta P)^2\rangle + \frac{1}{2}\right] \geqslant 1, \quad (3.15)$$

which is saturated by coherent states [$\langle(\Delta X)^2\rangle = \langle(\Delta P)^2\rangle = \frac{1}{2}$], can thereafter be understood as a physical manifestation of the Gauss-Weierstrass transform [related to Eq. (3.14)] between the Wigner and Husimi functions if the joint-measurement data are directly used to calculate variances (here denoted by $\text{Var}_{\text{Q}}[y] = \overline{y^2} - \overline{y}^2$ for a complete Husimi-function data $\{y\}$). We shall show that this additional quantum noise, when combined with optimal tomography strategies, can still lead to better moment-reconstruction accuracies relative to the homodyne scheme.

1. First-moment reconstruction

From Sec. III B, we note that the data collected from the heterodyne scheme are a scatter set of phase-space coordinates $\{(x_j, p_j)\}$ that are distributed according to the Husimi function. As Eq. (3.14) tells us that there is no difference between the state average $\boldsymbol{r}$ and Husimi average of $(x\ p)^{\text{T}}$, being a two-parameter estimation scheme, the first-moment scaled CRB with respect to the heterodyne data can again be found by taking the average of the distance between the estimator

$$\widehat{\boldsymbol{r}}_{\text{het}} \,\widehat{=}\, \frac{1}{N}\sum_{j=1}^{N}\begin{pmatrix} x_j \\ p_j \end{pmatrix} \quad (3.16)$$

and the true column $\boldsymbol{r}^{\text{T}} \,\widehat{=}\, (\overline{x}\ \ \overline{p})$:

$$\mathscr{D}_{1,\text{het}} = \mathbb{E}[(\widehat{\boldsymbol{r}}_{\text{het}} - \boldsymbol{r})^2] = \frac{1}{N}(\text{Var}_{\text{Q}}[x] + \text{Var}_{\text{Q}}[p]), \quad (3.17)$$

so that

$$\mathscr{H}_{1,\text{het}} = \text{Var}_{\text{Q}}[x] + \text{Var}_{\text{Q}}[p] = \text{Tr}\{\boldsymbol{G}\} + 1. \quad (3.18)$$

That $N\mathscr{D}_{1,\text{het}} = \mathscr{H}_{1,\text{het}}$ follows in the limit of large N, where the unbiased estimator $\widehat{\boldsymbol{r}}_{\text{het}}$ is asymptotically optimal since in this limit, the distribution of $\widehat{\boldsymbol{r}}_{\text{het}}$ follows a bivariate Gaussian distribution with vanishing widths, such that $\widehat{\boldsymbol{r}}_{\text{het}}$ becomes the maximum-likelihood estimator that approaches the scaled CRB for this Gaussian distribution.

2. Second-moment reconstruction

Similarly, to arrive at the optimal accuracy for estimating $\boldsymbol{G}_2$ using heterodyne data, we define the natural second-moment estimator

$$\widehat{\boldsymbol{G}}_{2,\text{het}} \,\widehat{=}\, \frac{1}{N}\sum_{j=1}^{N}\begin{pmatrix} x_j^2 & x_j p_j \\ x_j p_j & p_j^2 \end{pmatrix}, \quad (3.19)$$

where $\{(x_j, p_j)\}$ are again the sampled Husimi-function data collected during heterodyne detection. From Eq. (3.14), we get

$$\boldsymbol{G}_{2,\text{het}} = \boldsymbol{G}_2 + \frac{1}{2}. \quad (3.20)$$

The MSE $\mathscr{D}_{2,\text{het}}$ for heterodyne detection concerning second-moment estimation is consequently given by

$$\begin{aligned} \mathscr{D}_{2,\text{het}} &= \mathbb{E}[\text{Tr}\{(\widehat{\boldsymbol{G}}_{2,\text{het}} - \boldsymbol{G}_{2,\text{het}})^2\}] \\ &= \text{Tr}\{\mathbb{E}[\widehat{\boldsymbol{G}}_{2,\text{het}}^2]\} - \text{Tr}\{\boldsymbol{G}_{2,\text{het}}^2\} \\ &= \frac{1}{N}(\text{Var}_{\text{Q}}[x^2] + \text{Var}_{\text{Q}}[p^2] + 2\,\text{Var}_{\text{Q}}[xp]). \end{aligned} \quad (3.21)$$

In the large-N limit, this MSE is essentially the scaled CRB

$$\mathscr{H}_{2,\text{het}} = \text{Var}_{\text{Q}}[x^2] + \text{Var}_{\text{Q}}[p^2] + 2\,\text{Var}_{\text{Q}}[xp] \quad (3.22)$$

since $\widehat{\boldsymbol{G}}_{2,\text{het}}$ again becomes the maximum-likelihood estimator. To see this, we inspect the Fisher matrix $\boldsymbol{F}_{2,\text{het}}$ for the estimator $\widehat{\boldsymbol{G}}_{2,\text{het}}$. If we look at the random column

$$\boldsymbol{x} = \text{vec}(\widehat{\boldsymbol{G}}_{2,\text{het}}) \,\widehat{=}\, \frac{1}{N}\sum_{j=1}^{N}\begin{pmatrix} x_j \\ \sqrt{2}\,x_j p_j \\ p_j \end{pmatrix} \quad (3.23)$$

that represents $\widehat{\boldsymbol{G}}_{2,\text{het}}$, we find that in the limit of large N, the central limit theorem again says that $\boldsymbol{x}$ follows a Gaussian distribution defined by the mean $\boldsymbol{\mu} = \overline{\boldsymbol{x}} \,\widehat{=}\, (\overline{x^2}\ \sqrt{2}\,\overline{xp}\ \ \overline{p^2})^{\text{T}}$ and the covariance matrix (unimportant entries marked by $*$)

$$\boldsymbol{\Sigma} = \overline{\boldsymbol{x}\boldsymbol{x}^{\text{T}}} - \boldsymbol{\mu}\boldsymbol{\mu}^{\text{T}} \,\widehat{=}\, \frac{1}{N}\begin{pmatrix} \text{Var}_{\text{Q}}[x^2] & * & * \\ * & 2\,\text{Var}_{\text{Q}}[xp] & * \\ * & * & \text{Var}_{\text{Q}}[p^2] \end{pmatrix}, \quad (3.24)$$

so that we eventually recover the well-known result $\boldsymbol{\Sigma} = \boldsymbol{F}_{2,\text{het}}^{-1}$ for Gaussian scatter data that saturates the CRB as we

remember that $\mathrm{Tr}\{\boldsymbol{\Sigma}\} = \mathscr{D}_{2,\mathrm{het}}$. Equation (3.22) then follows tout de suite.

IV. FIRST-MOMENT ESTIMATION

A. General optimality of heterodyne tomography

As far as first-moment estimation is concerned, the general results in Eqs. (3.5) and (3.18) imply that $\mathscr{H}_{1,\mathrm{het}} \leqslant \mathscr{H}_{1,\mathrm{hom}}$ for *any* quantum state. This main result hinges on the physical HRS uncertainty relation, which is equivalent to the constraint $\det\{\boldsymbol{G}\} \geqslant \frac{1}{4}$ for the covariance matrix $\boldsymbol{G}$. This constraint means that

$$\mathscr{H}_{1,\mathrm{hom}} = \mathrm{Tr}\{\boldsymbol{G}\} + 2\sqrt{\det\{\boldsymbol{G}\}} \geqslant \mathrm{Tr}\{\boldsymbol{G}\} + 1 = \mathscr{H}_{1,\mathrm{het}}. \quad (4.1)$$

This implies that for *all* quantum states, the reconstruction accuracy of the optimal heterodyne first-moment estimator is always higher or equal to that of the optimal homodyne first-moment estimator in locating the average center of the quantum state in phase space. For minimum-uncertainty states, the accuracies of the two schemes are equal ($\mathscr{H}_{1,\mathrm{hom}} = \mathscr{H}_{1,\mathrm{het}}$). Subsequent well-known and interesting examples merely illustrate this fundamental fact. In terms of the first-moment performance ratio

$$\gamma_1 = \frac{\mathscr{H}_{1,\mathrm{het}}}{\mathscr{H}_{1,\mathrm{hom}}}, \quad (4.2)$$

a subunit magnitude indicates that the heterodyne scheme outperforms the homodyne scheme.

B. Gaussian states

For a Gaussian state where the covariance matrix $\boldsymbol{G}$ characterizes the spread of its Wigner function, the state variance of X_ϑ is simply

$$\langle X_\vartheta^2 \rangle - \langle X_\vartheta \rangle^2 = \boldsymbol{u}_\vartheta^{\mathrm{T}} \boldsymbol{G} \boldsymbol{u}_\vartheta. \quad (4.3)$$

From Eqs. (3.5) and (3.18), the first-moment performance ratio

$$\gamma_1 = \frac{\mathrm{Tr}\{\boldsymbol{G}\} + 1}{\mathrm{Tr}\{\boldsymbol{G}\} + 2\sqrt{\det\{\boldsymbol{G}\}}} \leqslant 1 \quad (4.4)$$

clearly cannot exceed one since any physical state satisfying the HRS uncertainty relation must take $\det\{\boldsymbol{G}\} \geqslant \frac{1}{4}$. The maximum value of $\gamma_1 = 1$ is attained for minimum-uncertainty states.

C. Fock states

A Fock state of the ket $|n\rangle$ is always centered at the origin of the phase space ($\boldsymbol{r} = \boldsymbol{0}$). The circular symmetry of these states imply the fact that $(\Delta X)^2 = (\Delta P)^2 = n + 1/2 = (\Delta X_\vartheta)^2$, hence

$$\mathscr{H}_{1,\mathrm{hom}} = 2(2n+1) \quad (4.5)$$

since such states have zero first moments. On the other hand, for the heterodyne scheme, we get

$$\mathscr{H}_{1,\mathrm{het}} = 2(n+1) \quad (4.6)$$

by simply using the Husimi characteristic function from Table I in Appendix B. Therefore, we get a

$$\gamma_1 = \frac{n+1}{2n+1} \quad (4.7)$$

that is always subunity unless $n = 0$, a result that is again familiar from Sec. IV B. In the limit of large photon numbers, the first-moment γ_1 approaches $\frac{1}{2}$.

D. Even and odd coherent states

Another popular class of non-Gaussian states with interesting phase-space quantum interference features are the even and odd coherent states characterized by the ket $|\pm;\alpha_0\rangle = (|\alpha_0\rangle \pm |-\alpha_0\rangle)\mathscr{N}_\pm$ of appropriate normalization constants $\mathscr{N}_\pm = 1/\sqrt{2 \pm 2\,e^{-2|\alpha_0|^2}}$, whose first moments $\boldsymbol{r}$ are all equal to zero. Using the definitions $a = \frac{1}{2}[\langle(\Delta X)^2\rangle - \langle(\Delta P)^2\rangle] = \alpha_0^2$ and $b_\pm = \frac{1}{2}[\langle(\Delta X)^2\rangle + \langle(\Delta P)^2\rangle] = \alpha_0^2[\tanh(\alpha_0^2)]^{\pm 1} + \frac{1}{2}$,

$$\mathscr{H}_{1,\mathrm{hom}} = 2\left(b_\pm + \sqrt{b_\pm^2 - a^2}\right). \quad (4.8)$$

For the heterodyne counterpart, one finds that

$$\mathscr{H}_{1,\mathrm{het}} = 2\left(b_\pm + \frac{1}{2}\right), \quad (4.9)$$

which contributes to the performance ratio

$$\gamma_1 = \frac{b_\pm + \frac{1}{2}}{b_\pm + \sqrt{b_\pm^2 - a^2}}. \quad (4.10)$$

For both types of coherent-state superpositions, $\gamma_1 \to 1$ as $\alpha_0 \to \infty$. For even coherent states, the performance ratio $\gamma_1 = 1$ when $\alpha_0 = 0$ as it should. Otherwise, this ratio is always less than one for any positive α_0. There exists a single local minimum of $\gamma_1 \approx 0.7577$ at $\alpha_0 \approx 1.715$. For odd coherent states, $\gamma_1 < 1$ for *all* α_0 values, with the minimum value of $\gamma_1 = \frac{1}{3}$ at $\alpha_0 = 0$. For these states, γ_1 increases monotonically to one as α_0 tends to infinity.

E. Displaced Fock states

Displacement and photon addition are two important physical procedures that are frequently discussed in quantum physics. The different orders in which these processes are carried out on the vacuum state give output states of a different nature. Displacing an m-photon-added vacuum state by a complex amplitude α_0 results in displaced Fock states defined by the ket $D(\alpha_0)|m\rangle$ can be effectively performed using a beam splitter with a high transmissivity and a strong coherent state [55,56].

It can be shown easily that the first-moment scaled CRBs are indeed given by Eqs. (4.5) and (4.6), so that the performance ratio is then completely identical to that of the usual central Fock states in Eq. (4.7). This reflects the physical fact that the accuracy in estimating the displacement cannot explicitly depend on where the center of the displaced Fock states is when full sets of CV measurement outcomes are considered, as the tomographic coverage of the entire phase space is then complete. This accuracy depends only on the variances, which describe the second-order symmetry and is unaffected at all by the displacement.

F. Photon-added coherent states

A swap in the order of photon addition and displacement on the vacuum state gives the photon-added coherent state of m added photons and reference amplitude α_0 is defined by the ket $|m; \alpha_0\rangle = \mathcal{N}_{m,|\alpha_0|^2} A^{\dagger m} |\alpha_0\rangle$ with the bosonic annihilation operator A, where the normalization constant $\mathcal{N}_{m,|\alpha_0|^2} = e^{|\alpha_0|^2/2}/\sqrt{m!\,_1F_1(m+1;1;|\alpha_0|^2)}$ involves the confluent hypergeometric function of the first kind $_1F_1(a;b;y)$. The integer value m denotes the extent to which the mean photon number

$$\langle A^{\dagger} A \rangle = (m+1) \frac{_1F_1(m+2;1;|\alpha_0|^2)}{_1F_1(m+1;1;|\alpha_0|^2)} - 1, \qquad (4.11)$$

which is always larger than $|\alpha_0|^2 + m$ whenever $\alpha_0 \neq 0$, is increased nonlinearly by the operation by $A^{\dagger m}$ on the reference coherent ket $|\alpha_0\rangle$. This particular class of quantum states is but one of many possible kinds of photon-added states, which are of interest to the quantum community for testing some fundamental statements [57–59].

For these photon-added coherent states, the second-order symmetry is now affected by the combined action of the displacement and photon addition, so that $\langle (\Delta X)^2 \rangle$ and $\langle (\Delta P)^2 \rangle$ are functions of m and α_0. These expressions can be straightforwardly computed with the help of the characteristic functions given in Table I in Appendix B. By defining

$$a = -\frac{\alpha_0^2(m+1)}{2\,_1F_1(m+1;1;\alpha_0^2)^2} \Big[2(m+1)\,_1F_1(m+2;2;\alpha_0^2)^2$$
$$- (m+2)\,_1F_1(m+1;1;\alpha_0^2)\,_1F_1(m+3;3;\alpha_0^2)\Big],$$
$$b = m + \frac{1}{2} - \frac{\alpha_0^2 m\,_1F_1(m+1;2;\alpha_0^2)}{_1F_1(m+1;1;\alpha_0^2)^2} \Big[\,_1F_1(m+1;1;\alpha_0^2)$$
$$+ m\,_1F_1(m+1;2;\alpha_0^2)\Big], \qquad (4.12)$$

such that $b > a$, the first-moment scaled CRB for homodyne detection is of the same form as in Eq. (4.8), namely,

$$\mathscr{H}_{1,\text{hom}} = 2(b + \sqrt{b^2 - a^2}). \qquad (4.13)$$

The first-moment scaled CRB for heterodyne detection is given by

$$\mathscr{H}_{1,\text{het}} = 2\left[a + (m+1)\frac{_1F_1(m+2;2;\alpha_0^2)}{_1F_1(m+1;1;\alpha_0^2)}\right]. \qquad (4.14)$$

Clearly, when $\alpha_0 = 0$, the answers in Eqs. (4.5) and (4.6) for an m-number Fock state are reproduced exactly. With $m = 0$, the respective scaled CRBs of a value of 2 for all α_0's are furthermore consistent with Sec. IV B. Otherwise, γ_1 is always subunity, and approaches unity as $\alpha_0 \to \infty$.

V. SECOND-MOMENT ESTIMATION

A. Gaussian states

It seems fitting to commence the discussion of second-moment estimation with the Gaussian state, for it is natural to begin with the generalization of the results that already appeared in Refs. [44,45] to general noncentral Gaussian states ($\boldsymbol{r} \neq \boldsymbol{0}$). We suppose that the Gaussian state of the covariance matrix $\boldsymbol{G}$ is centered at $\boldsymbol{r} = \boldsymbol{r}_0 = (x_0 \ p_0)^{\mathsf{T}}$. From Table I in Appendix B, by defining $\mu_\vartheta = \boldsymbol{u}_\vartheta^{\mathsf{T}} \boldsymbol{r}_0$ and $\sigma_\vartheta^2 = \boldsymbol{u}_\vartheta^{\mathsf{T}} \boldsymbol{G} \boldsymbol{u}_\vartheta$, the variance for the second quadrature moment reads as

$$\langle X_\vartheta^4 \rangle - \langle X_\vartheta^2 \rangle^2 = 2\,\sigma_\vartheta^2 (\sigma_\vartheta^2 + 2\,\mu_\vartheta^2). \qquad (5.1)$$

For *central* Gaussian states ($\langle X \rangle = \langle P \rangle = 0$), we have $\langle X_\vartheta^4 \rangle = 3\langle X_\vartheta^2 \rangle^2$ and the scaled Fisher matrix in Eq. (3.9) turns into the familiar form in [44,45]. For the more general situation, one can repeat the contour-method integration in [44] to calculate the scaled Fisher matrix in Eq. (3.9). The answer is given as

$$\widetilde{\boldsymbol{F}}_{2,\text{hom}} = \frac{-2}{(c+ib)(w_3+iw_2)}\bigg[\frac{M_{z=0}}{z_{1+}z_{1-}z_{2+}z_{2-}}$$
$$+ \frac{M_{z=z_{1-}}}{z_{1-}(z_{1-}-z_{1+})(z_{1-}-z_{2+})(z_{1-}-z_{2-})}$$
$$+ \frac{M_{z=z_{2-}}}{z_{2-}(z_{2-}-z_{1-})(z_{2-}-z_{1+})(z_{2-}-z_{2+})}\bigg] \quad (5.2)$$

together with the definitions

$$a = \frac{1}{2}\text{Tr}\{\boldsymbol{G}\}, \quad b = \frac{1}{2}(\boldsymbol{G}_{11} - \boldsymbol{G}_{22}), \quad c = \boldsymbol{G}_{12},$$
$$w_1 = a + \boldsymbol{r}_0^2, \quad w_2 = b + x_0^2 - p_0^2, \quad w_3 = c + 2x_0 p_0,$$
$$z_{1\pm} = \frac{-a \pm i\sqrt{-a^2 + b^2 + c^2}}{b - ic}, \quad z_{2\pm} = \frac{-w_1 \pm i\sqrt{-w_1^2 + w_2^2 + 2w_3^2}}{w_2 - iw_3},$$
$$M_z \mathbin{\widehat{=}} \frac{1}{16}\begin{pmatrix} (z+1)^4 & -i\sqrt{2}(z-1)(z+1)^3 & -(z^2-1)^2 \\ -i\sqrt{2}(z-1)(z+1)^3 & -2(z^2-1)^2 & i\sqrt{2}(z+1)(z-1)^3 \\ -(z^2-1)^2 & i\sqrt{2}(z+1)(z-1)^3 & (z-1)^4 \end{pmatrix}. \qquad (5.3)$$

When $\boldsymbol{r}_0 = 0$, we have $w_1 = a$, $w_2 = b$, and $w_3 = c$ and the scaled Fisher matrix $\widetilde{\boldsymbol{F}}_{2,\text{hom}}$ reduces to that for the central Gaussian state in [44]. For the general setting, the full expression of $\mathscr{H}_{2,\text{hom}}$ is omitted here in this case due to

its complexity. On the other hand, the scaled CRB with the heterodyne scheme for these noncentral Gaussian states can be calculated directly from Eq. (3.22) using the characteristic

function in Table I and is given by

$$\mathcal{H}_{2,\text{het}} = 2 \left(\text{Tr}\{\boldsymbol{G}_{\text{het}}\}^2 - \det\{\boldsymbol{G}_{\text{het}}\} \right.$$
$$\left. + \boldsymbol{r}_0{}^{\text{T}} \boldsymbol{G}_{\text{het}} \boldsymbol{r}_0 + \text{Tr}\{\boldsymbol{G}_{\text{het}}\} \boldsymbol{r}_0^2 \right), \qquad (5.4)$$

where one immediately verifies the counterpart expression in [44] for the central Gaussian states upon setting $\boldsymbol{r}_0 = \boldsymbol{0}$.

At this stage, we reassure ourselves the physics of the problem of second-moment tomography by understanding, first, that in the case where tomography is performed on the *full* covariance matrix $\boldsymbol{G}$, then the scaled CRB, which is the minimum of the MSE, should not depend on the orientation of the two-dimensional uncertainty region (here being an ellipse for any Gaussian state) described by the eigenvectors of this matrix but only its eigenvalues owing to the form of the MSE. Additionally, the accuracy should also be independent of $\boldsymbol{r}_0$. When only the second-moment matrix $\boldsymbol{G}_2$ is reconstructed, the scaled CRB should also not depend on its eigenvectors but only its eigenvalues. The physics remains the same. However, there is a difference between estimating the full matrix $\boldsymbol{G}$ and estimating just $\boldsymbol{G}_2$. Since $\boldsymbol{G}_2$ is in general an increasing function of the first moments, this means that as the displacement of the center from the phase-space origin for the quantum state increases, the geometric mean of eigenvalues (GME) of $\boldsymbol{G}_2$ correspondingly becomes larger so that the second-order "temperature" of the state, a terminology borrowed from Gaussian states, as described by the GME is now higher and this results in a stronger $\boldsymbol{G}_2$-"thermal" property much like the thermal Gaussian states. So we would expect, based on the findings in [44], that states with large displacements give poor second-moment tomographic accuracies for *both* CV schemes, and yet provide a subunit

$$\gamma_2 = \frac{\mathcal{H}_{2,\text{het}}}{\mathcal{H}_{2,\text{hom}}} \qquad (5.5)$$

performance ratio. It is also physically intuitive that the accuracies for both schemes should also be independent of the angle of displacement, but depend only on the magnitude of the displacement. For non-Gaussian states, the *fourth* moments arising from the structure of the MSE, which are no longer functions of the first and second moments as is the case for Gaussian states, also contribute to the scaled CRB, and therefore γ_2, as described in the general theory in Sec. III.

This physics, however, *seems* to be violated by the noncentral-Gaussian-state expressions in (5.2) and (5.4), namely, that $\mathcal{H}_{2,\text{het}}$ depends on the explicit displacement vector $\boldsymbol{r}_0$ and covariance matrix $\boldsymbol{G}$, for instance. This mishap has nothing to do with any kind of physical violation, but has only to do with the way we specify Gaussian states. By choosing to parametrize a multivariate Gaussian distribution using the natural independent parameters $\boldsymbol{r}_0$ and $\boldsymbol{G}$ (the full matrix), we inadvertently change the eigenvalues of $\boldsymbol{G}_2$ by changing $\boldsymbol{r}_0$ and fixing $\boldsymbol{G}$. This becomes obvious when one finds that the two positive eigenvalues $\lambda_\pm$ of $\boldsymbol{G}_2$ are given by

$$\lambda_\pm = |\alpha_0|^2 + \tfrac{1}{2} \text{Tr}\{\boldsymbol{G}_{\text{het}}\} \pm |\alpha_0^2 + \boldsymbol{w}^{\text{T}} \boldsymbol{G}_{\text{het}} \boldsymbol{w}|^2, \qquad (5.6)$$

where $\boldsymbol{w} = \frac{1}{\sqrt{2}}(1 \quad i)^{\text{T}}$ and $\alpha_0 = (x_0 + ip_0)/\sqrt{2}$. The consequence of this natural definition results in such an apparent observation. The noncentral Gaussian states so defined form the singular example in this article where this happens, and
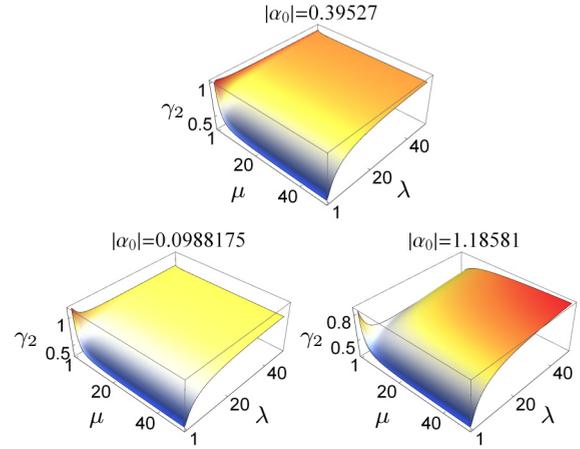
FIG. 2. Plots of γ_2 surfaces for $\phi = 0$ and different displacement magnitudes along the x axis in phase space. The center plot refers to the critical displacement magnitude of $\sqrt{5/32} \approx 0.395$, beyond which $\gamma_2 < 1$ for all μ and λ. The surface tip at $\mu = \lambda = 1$ for the coherent states is invariant under a displacement rotation. It is clear from these plots that increasing the temperature reduces the value of γ_2, while increasing the squeezing strength counters this reduction.

the two other noncentral non-Gaussian states which we shall soon visit do not have this technical issue.

To investigate the second-moment performance ratio $\gamma_2 = \mathcal{H}_{2,\text{het}}/\mathcal{H}_{2,\text{hom}}$, we may reparametrize the eigenvalues of $\boldsymbol{G}$ with the squeezing strength $1 \leqslant \lambda < \infty$ and the temperature parameter $1 \leqslant \mu < \infty$ that measures how thermal the state is relative to minimum-uncertainty states, which are commonly adopted in describing Gaussian states. Then $\boldsymbol{G}$ has the spectral decomposition

$$\boldsymbol{G} \,\hat{=}\, \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \frac{\mu}{2\lambda} & 0 \\ 0 & \frac{\mu\lambda}{2} \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix},$$
$$(5.7)$$

where ϕ orientates the eigenvectors of $\boldsymbol{G}$. In this parametrization, we can clearly see that a large displacement magnitude contributes to a large temperature, so that a small value of γ_2 can be anticipated for these highly displaced or $\boldsymbol{G}_2$-thermal Gaussian states based on the conclusions in [44,45]. The behavior of γ_2 is very similar to that for the central Gaussian states and is plotted in Fig. 2 for various values of $|\alpha_0|$. The lowest achievable γ_2 values go with the highly thermal Gaussian states ($\lambda = 1$, $\mu \gg |\boldsymbol{r}_0|$), whose covariance matrix $\boldsymbol{G} = \mu\boldsymbol{1}/2$ is simply a multiple of the identity. Their second quadrature moment has a variance $\langle X_\vartheta^4 \rangle - \langle X_\vartheta^2 \rangle^2 = \mu^2/2$, according to Eq. (5.1), that is of course independent of the phase ϑ due to the rotational symmetry. The performance ratio then takes the minimum value of $\frac{3}{10}$.

The maximum of γ_2 occurs with the coherent states ($\mu = \lambda = 1$) and takes a value of $\frac{6}{5}$ for $\boldsymbol{r}_0 = \boldsymbol{0}$. For larger magnitudes of α_0, the value of γ_2 drops below unity beyond the magnitude of $|\alpha_0| = \sqrt{5/32}$, which can be obtained through optimization. One may verify that at this critical magnitude, $\mathcal{H}_{2,\text{hom}} = \mathcal{H}_{2,\text{het}} = \frac{63}{8}$. So, given a displacement magnitude larger than $\sqrt{5/32}$, the heterodyne scheme always outperforms the homodyne scheme in second-moment estimation. In the
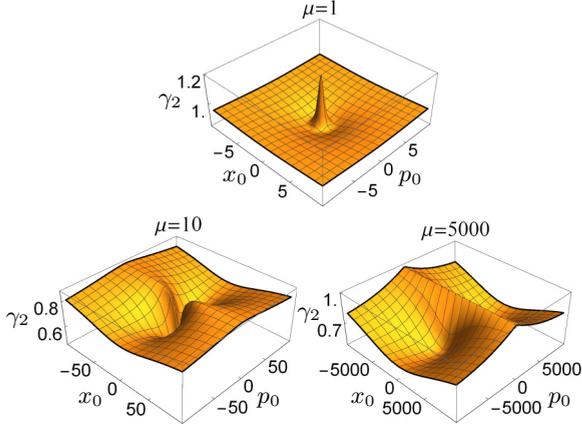
FIG. 3. Plots of γ_2 surfaces against the displacement $\mathbf{r}_0$ for $\phi = 0$ and different values of $\mu = \lambda$. In the limit $\mu \to \infty$, $\gamma_2 \leqslant 1$ approaches unity at $p_0 = 0$. The signature peak of $\gamma_2 = \frac{6}{5} = 1.2$ at $x_0 = p_0 = 0$ for $\mu = 1$ is consistent with the finding in Refs. [44,45] for central Gaussian states.

limit of large μ and λ, where we may take this limit such that $\mu = \lambda$ without loss of generality, if one considers the spectral decomposition in Eq. (5.7), then γ_2 for $\phi = 0$ plotted in Fig. 3 shows the values for different μ as an indication that $\gamma_2 \leqslant 1$ in this limit. Different ϕ values simply rotate these plots in the x_0-p_0 plane.

B. Fock states

Owing to the rotational symmetry of the Fock states $[\boldsymbol{G}_2 = (n + 1/2)\mathbf{1}]$, the second and fourth quadrature moments

$$\langle X_\vartheta^4 \rangle - \langle X_\vartheta^2 \rangle^2 = \frac{1}{2}\langle X_\vartheta^2 \rangle^2 + \frac{3}{8} \qquad (5.8)$$

are independent of the phase ϑ, so that the Fisher matrix

$$\boldsymbol{F}_{2,\text{hom}} = \frac{1}{4(n^2 + n + 1)}\begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix}. \qquad (5.9)$$

It then follows that the scaled CRB is given by

$$\mathscr{H}_{2,\text{hom}} = 5\,(n^2 + n + 1). \qquad (5.10)$$

On the other hand, the Husimi characteristic function for the Fock states in Appendix B produces the answer

$$\mathscr{H}_{2,\text{het}} = 2\,(n + 1)(n + 3). \qquad (5.11)$$

The performance ratio

$$\gamma_2 = \frac{2\,(n + 1)(n + 3)}{5\,(n^2 + n + 1)} \qquad (5.12)$$

is less than one for $n \geqslant 2$, in which regime the Fock states are sufficiently $\boldsymbol{G}_2$ "thermal". For $n = 0$, we evidently obtain the familiar answer $\gamma_2 = \frac{6}{5}$ for the vacuum state, whereas for $n = 1$, $\gamma_2 = \frac{16}{15}$. In the limit of large n, $\gamma_2 \to \frac{2}{5}$ (see Fig. 4).

C. Even and odd coherent states

Since the eigenvalues

$$\lambda_\pm^{(\pm)} = \frac{1}{2} + |\alpha_0|^2\{[\tanh(|\alpha_0|^2)]^{(\pm 1)} \pm 1\} \qquad (5.13)$$
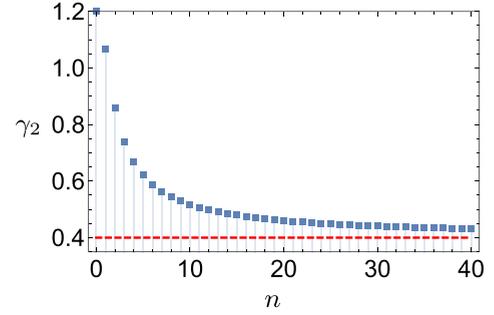
FIG. 4. Plot of γ_2 (solid blue squares) against n for Fock states. As n increases, γ_2 decreases monotonically and eventually saturates at a subunit constant of $\frac{2}{5}$ (dashed red line).

of $\boldsymbol{G}_2$ are simple functions of $|\alpha_0|^2$ for the even and odd $(\pm)$ coherent states, we may take $\alpha_0 \geqslant 0$ without loss of generality. The quadrature moments can be easily derived with the help of Appendix B, which give the following second-moment variance:

$$\langle X_\vartheta^4 \rangle - \langle X_\vartheta^2 \rangle^2 = \frac{1}{2} + 2\alpha_0^2\{\cos(2\vartheta) + [\tanh(\alpha_0^2)]^{\pm 1}\}$$
$$\pm \frac{4\alpha_0^4}{(e^{\alpha_0^2} \pm e^{-\alpha_0^2})^2}. \qquad (5.14)$$

By relying on the asymptotic behaviors $\coth y \approx 1/y$ and $\text{cosech}\, y \approx 1/y$ of the hyperbolic trigonometric functions for small arguments, we revert to the limiting second-moment variances for $n = 0$ and 1, which is consistent with the fact that the even states approach the vacuum state and the odd states approach the single-photon Fock state. The Fisher matrix $\boldsymbol{F}_{2,\text{hom}}$ thus takes the simple form

$$\boldsymbol{F}_{2,\text{hom}} = \int_{(\pi)} \frac{d\vartheta}{\pi}\,\frac{\boldsymbol{M}_\vartheta}{m_\pm + l\cos(2\vartheta)} \quad (l = 2\alpha_0^2 < m_\pm),$$
$$m_\pm = \frac{1}{2} + 2\alpha_0^2[\tanh(\alpha_0^2)]^{\pm 1} \pm \frac{4\alpha_0^4}{(e^{\alpha_0^2} \pm e^{-\alpha_0^2})^2}, \qquad (5.15)$$

hence, one obtains

$$\mathscr{H}_{2,\text{hom}} = 6m_\pm + 4\sqrt{m_\pm^2 - l^2} \qquad (5.16)$$

after carrying out the integration, matrix inversion, and matrix trace. On the other hand, the Husimi-average moments of the heterodyne data contribute to the result

$$\mathscr{H}_{2,\text{het}} = 6 + 12\alpha_0^2[\tanh(\alpha_0^2)]^{\pm 1} \pm \frac{8\alpha_0^4}{(e^{\alpha_0^2} \pm e^{-\alpha_0^2})^2} \qquad (5.17)$$

for the heterodyne scaled CRB.

We once again remind the reader that the scaled CRBs stated in Eqs. (5.16) and (5.17) are independent of the phase of the even and odd coherent states, as this phase amounts to a rotation in phase space that is immaterial in determining the moment-estimation accuracy. For arbitrary complex values of α_0, the expressions are still valid after the change $\alpha_0^2 \to |\alpha_0|^2$.

The ratio γ_2 is greater than one for small values of α_0, with the special limiting cases ($\alpha_0 = 0$) being those of the respective Fock states, and less than one for large values of α_0.
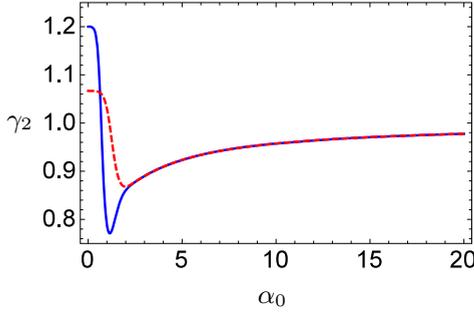
FIG. 5. Plots of γ_2 for the even (solid blue curve) and odd (dashed red curve) coherent states against the parameter α_0 that characterizes the even and odd coherent states. For the even coherent states, the unit-γ_2 crossover occurs at $\alpha_0 \approx 0.693$, whereas for the odd coherent states, this happens at $\alpha_0 \approx 1.128$. Furthermore, for each type of states, γ_2 possesses a stationary global minimum. For the even states, the minimum value of $\gamma_{2,\min} = 0.770\,96$ is attained at $\alpha_0 = 1.148 \approx 1$. For the odd states, this optimum value is $\gamma_{2,\min} = 0.867\,96$ and is achieved with $\alpha_0 = 1.980 \approx 2$.

The crossover values for which these states become sufficiently $\boldsymbol{G}_2$ "thermal" such that $\gamma_2 = 1$ differ for both the even and odd states (see Fig. 5). For sufficiently large α_0, γ_2 approaches unity from below. This can be clearly seen by taking the limit $\alpha_0 \to \infty$. In this limit, we have $m_\pm \to 2\alpha_0^2 = l$ so that $\mathscr{H}_{2,\mathrm{hom}} \to 12\alpha_0^2 \approx \mathscr{H}_{2,\mathrm{het}}$. For these class of states, γ_2 has a stationary minimum that is again different for the two types of states, and this is elucidated in Fig. 5. At $\alpha_0 \approx 0.631$, the γ_2 values for the even and odd states are equal, even though their $\boldsymbol{G}_2$ matrices are very different. The reason is that the combined contributions of all the second and fourth moments give an overall multiplicative factor of about 2.0694 to both $\mathscr{H}_{2,\mathrm{het}}$ and $\mathscr{H}_{2,\mathrm{hom}}$ for the odd state relative to the even state.

D. Displaced Fock states

As opposed to the previous three classes of states, the displaced Fock states (as well as the photon-added coherent states that follow) possess a nonzero quadrature first moment. As the only two parameters $\alpha_0 = (x_0 + i p_0)/\sqrt{2}$ and m that characterize these displaced Fock states do not, in any way, restrict the covariance matrix $\boldsymbol{G}$, it is easy to show that the $\boldsymbol{G}_2$ geometry, and hence its reconstruction accuracy, depend only on the displacement magnitude $|\alpha_0|^2$ and not its phase. This is done by directly inspecting the eigenvalues of $\boldsymbol{G}_2$, namely,

$$\lambda_1 = m + \tfrac{1}{2},$$
$$\lambda_2 = m + 2|\alpha_0|^2 + \tfrac{1}{2}, \qquad (5.18)$$

one of which is an increasing function of $|\alpha_0|^2$. As a result, we only need to consider the case where $\alpha_0 = x_0/\sqrt{2}$ is positive. As α_0 increases, the GME increases, which means that the quantum state becomes more $\boldsymbol{G}_2$ "thermal." We shall soon see that an increase in $|\alpha_0|^2$ results in a smaller performance ratio γ_2 in favor of the heterodyne scheme.

To calculate the homodyne scaled CRB, we first note that the relevant even-order quadrature moments (see Appendix B)

supply the second-moment quadrature variance

$$\langle X_\vartheta^4 \rangle - \langle X_\vartheta^2 \rangle^2 = m_0 + l\cos(2\vartheta),$$
$$m_0 = \tfrac{1}{2}\big[m^2 + m + \alpha_0^2(8m + 4)\big], \quad (5.19)$$
$$l = 2\alpha_0^2(2m + 1) < m_0,$$

which bears striking resemblance in form with that for the even and odd coherent states, so that the scaled CRB also takes the same closed form as Eq. (5.16) inasmuch as

$$\mathscr{H}_{2,\mathrm{hom}} = 6m_0 + 4\sqrt{m_0^2 - l^2}. \qquad (5.20)$$

For the heterodyne scheme, we subsequently get

$$\mathscr{H}_{2,\mathrm{het}} = 2(m + 1)\big(m + 6\alpha_0^2\big) \qquad (5.21)$$

by again referring to Table I.

The interplay between the discrete (m) and continuous (α_0) parameters gives rise to familiar cases that have already been analyzed previously for the Gaussian and Fock states. For $m = 0$, we of course have the coherent state of amplitude α_0 where the maximum $\gamma_2(\alpha_0 = 0) = \frac{6}{5}$ and the crossover point $\gamma_2(\alpha_0 = \sqrt{5/32}) = 1$ beyond which $\gamma_2 < 1$ are reproduced by Eqs. (5.20) and (5.21). For $m = 1$, we have the $m = 1$ Fock state for $\alpha_0 = 0$ so that the unsurprising number $\gamma_2(\alpha_0) = \frac{16}{15}$ comes up from the same scaled CRB expressions. The crossover point for $\gamma_2 = 1$ is located at $\alpha_0 = \frac{1}{2}\sqrt{19/3 - 2\sqrt{87}/3} \approx 0.1696$. The performance ratio becomes subunit for *all* displacements α_0 for $m \geqslant 2$, just like the Fock states. In the limit of large displacements $\alpha_0^2 \gg m$, we have $m_0 \to l$ and

$$\gamma_2\big(\alpha_0^2 \gg m\big) = \frac{m+1}{2m+1}, \qquad (5.22)$$

which approaches $\frac{1}{2}$ in the regime $\alpha_0^2 \gg m \gg 1$.

For this two-parameter quantum state, it is interesting to look at the minimum value of γ_2 over all possible displacement magnitudes α_0 for each m [see Fig. 6(a)]. To calculate the minimum stationary points $\alpha_0 = \widetilde{\alpha}_0$, we differentiate γ_2 with respect to α_0 and set the derivative to zero. While the analytical form for the optimal $\gamma_2 = \gamma_{2,\min}$ as a complicated function of m exists, the approximated forms

$$\gamma_{2,\min} \approx \begin{cases} 0.8504 - 0.5893\,m & \text{(small-}m\text{ regime)}, \\ 0.3693 + \dfrac{0.6565}{m} & \text{(large-}m\text{ regime)} \end{cases} \qquad (5.23)$$

are enough to understand the optimal-γ_2 curve in terms of a power law already for moderately large m. Interestingly, the saturation point for γ_2 is slightly lower than $\frac{2}{5}$, which is the γ_2 for the Fock state of an infinitely large-m value. This hints that the optimal center for the displaced Fock state of a large m for which $\gamma_2 = \gamma_{2,\min}$ is significantly far away from the phase-space origin. This is indeed consistent with the behavior of the minimum point $\widetilde{\alpha}_0$, which also has a complicated closed-form expression [plotted in Fig. 6(a)], so that we only present the more useful approximated forms

$$\widetilde{\alpha}_0 \approx \begin{cases} 1.2929 + 2.2060\,m - 3.2976\,m^2 & \text{(small-}m\text{ regime)}, \\ 0.3993\sqrt{m} + \dfrac{2.8174}{\sqrt{m}} & \text{(large-}m\text{ regime)} \end{cases}$$
$$(5.24)$$
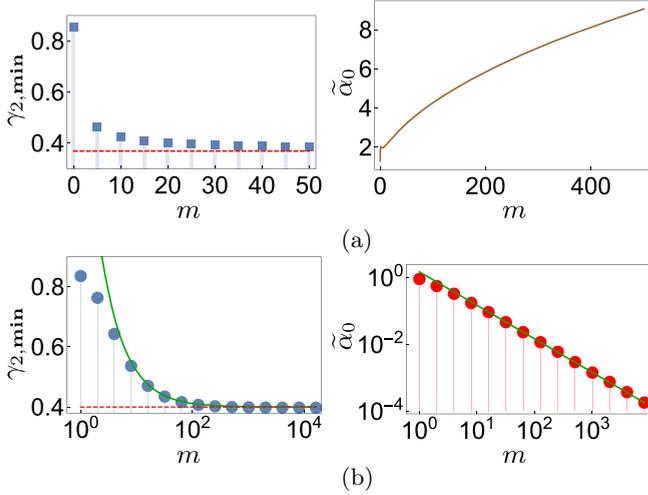
FIG. 6. Plots of (a) the optimum (minimum) γ_2 over all α_0 with m (left) and the minimum point $\alpha_0 = \widetilde{\alpha}_0$ (right) for the displaced Fock states, as well as those of (b) the photon-added coherent states. For the displaced Fock states in (a), $\gamma_{2,\mathrm{min}}$ tends to the limiting value of $6/5 - \sqrt{69}/10 \approx 0.3693$ (dashed red line), and the brown curve representing the exact expression for $\widetilde{\alpha}_0$ shows the quadratic behavior for small m and the approximate square-root behavior for large m. On the other hand, for the photon-added states in (b), the numerically found $\gamma_{2,\mathrm{min}}$ values (solid blue circles) are plotted with the theoretical asymptotic power-law curve (solid dark green curve) to illustrate the accuracy of the latter for $m \gtrsim 10$, both of which approach the limiting value of $\frac{2}{5}$ (dashed red line). The $\gamma_{2,\mathrm{min}}$ value for $m = 0$ (not plotted) has the analytical value of $3(6 - \sqrt{21})/5 \approx 0.85$ that occurs at $\widetilde{\alpha}_0 = \sqrt{13 + 3\sqrt{21}}/4 \approx 1.29$. The approximate model [see Eq. (5.30)] for $\widetilde{\alpha}_0$ (green line) is compared with the numerical minima (solid red circles) as a showcase of its remarkable fit.

that highlight the main gradient features. To summarize, the minimum value of γ_2 essentially behaves as a power law in m, and the corresponding stationary minimum $\widetilde{\alpha}_0$ is quadratic for small m and goes as a square-root curve for large m.

E. Photon-added coherent states

As in the case of the displaced Fock states, the eigenvalues of $\boldsymbol{G}_2$ for the photon-added coherent states,

$$\lambda_1 = (m+1)\frac{{}_1F_1(m+2;2;|\alpha_0|^2)}{{}_1F_1(m+1;1;|\alpha_0|^2)} - \frac{1}{2},$$

$$\lambda_2 = 2m + 2|\alpha_0|^2 + \frac{1}{2}$$
$$+ m(2|\alpha_0|^2 - 1)\frac{{}_1F_1(m+1;2;|\alpha_0|^2)}{{}_1F_1(m+1;1;|\alpha_0|^2)}, \quad (5.25)$$

are also functions of $|\alpha_0|^2$, which correctly coincides with the physics of the second-moment estimation problem. This also means that discussing in terms of the range $\alpha_0 \geqslant 0$ covers the tomography analysis sufficiently. Moreover, the eigenvalues are increasing functions of the displacement magnitude, so that the GME becomes larger with α_0, thereby rendering the photon-added states more $\boldsymbol{G}_2$ "thermal." This again gives a smaller performance ratio γ_2, or a better tomographic

performance for the heterodyne scheme compared to the homodyne scheme.

Once more with the help of Table I in Appendix B, the quadrature moments can be written in principle, but they are represented by bulky expressions that are hardly worth any analytical value and the Fisher-matrix integral in Eq. (3.9) has no known closed-form expression. However, we may still briefly discuss the important limiting cases. For $\alpha_0 \ll \sqrt{m}$, to second order in α_0, it can be shown that

$$\mathscr{H}_{2,\mathrm{hom}} \approx 5(m^2 + m + 1) + 10\alpha_0^2(m+1)(m+2), \quad (5.26)$$

where the asymptotic connection with Fock states is clear. On the other hand, in the regime of large $\alpha_0 \gg \sqrt{m}$, we find that

$$\mathscr{H}_{2,\mathrm{hom}} = 3 + 12\alpha_0^2 + 2\sqrt{1 + 8\alpha_0^2} \approx 12\alpha_0^2, \quad (5.27)$$

which is the second-moment homodyne-scaled CRB for coherent states. This is also the homodyne-scaled CRB for large-intensity even and odd coherent states. The reason is that for large amplitudes, all these states behave like a coherent state of amplitude α_0 as far as second-moment estimation is concerned since all their $\boldsymbol{G}_2$ eigenvalues are indistinguishable in this limit.

Upon revisiting Eq. (3.22), the heterodyne-scaled CRB can be shown to have the closed form

$$\mathscr{H}_{2,\mathrm{het}} = 2\left\{3 + 4m + 2\alpha_0^2(m+3) - m\frac{{}_1F_1(m+1;2;\alpha_0^2)}{\left[{}_1F_1(m+1;1;\alpha_0^2)\right]^2}\right.$$
$$\times \left[2(\alpha_0^4 - 3\alpha_0^2 - m)\,{}_1F_1(m+1;1;\alpha_0^2)\right.$$
$$\left.+ m(2\alpha_0^4 - 2\alpha_0^2 + 1)\,{}_1F_1(m+1;2;\alpha_0^2)\right]\right\} \quad (5.28)$$

for $\alpha_0 > 0$. The behavior to leading order in α_0^2 for $\alpha_0 \ll \sqrt{m}$,

$$\mathscr{H}_{2,\mathrm{het}} \approx (2 + 4\alpha_0^2)(m+1)(m+3), \quad (5.29)$$

is evidently consistent with the known result for Fock states. For $\alpha_0 \gg \sqrt{m}$, we once again have $\mathscr{H}_{2,\mathrm{het}} \approx 12\alpha_0^2$. Note also that, as expected, the first equality in Eq. (5.27) is equal to 5, and Eq. (5.27) gives a value of 6 for the vacuum state ($m = 0$).

For $m \geqslant 2$, the ratio $\gamma_2 < 1$ for all α_0. This natural extension to the result for the Fock states means that for highly nonlinear photon-"adding" operations, the performance of heterodyne detection is always better than that of homodyne detection in terms of second-moment covariance-dyadic estimation. For $m = 0$, the analysis reverts to that for the coherent state, where the crossover occurs at $\alpha_0 = \sqrt{5/32}$ after solving for $\mathscr{H}_{2,\mathrm{hom}} = \mathscr{H}_{2,\mathrm{het}} = 2(3 + 6\alpha_0^2)$ so that $\gamma_2(\alpha_0 > \sqrt{5/32}) < 1$, which is again consistent with Sec. V A. For $m = 1$, the crossover point $\alpha_0 \approx 0.2$ may be obtained as the numerical solution. As α_0 approaches infinity, previous arguments imply that $\gamma_2 \to 1$ for any m.

In view of the behavior of γ_2, another interesting limit is the high-nonlinearity limit ($m \to \infty$). In this case, we notice that the value $\alpha_0 = \widetilde{\alpha}_0$ for which γ_2 is minimum approaches zero. A good model to estimate this minimum point in this limit is given by

$$\widetilde{\alpha}_0 \approx \frac{3}{2m}, \quad (5.30)$$

which can be approximated from curve fitting. Therefore, in the large-m limit, the optimum performance ratio γ_2 is that of an intense Fock state of a large photon number and so we expect the minimum value of γ_2 to approach $\frac{2}{5}$ as discussed in Sec. V B. In other words, for sufficiently large m, the minimum of γ_2 follows the noncentral power law

$$\gamma_{2,\mathrm{min}} = \gamma_2|_{\alpha_0=\widetilde{\alpha}_0} \approx \frac{2}{5} + \frac{6}{5m}. \tag{5.31}$$

Figure 6(b) succinctly highlights these observations.

VI. TWO-MODE SOURCES: A TREATISE TO MULTIMODE MOMENT ESTIMATION

The general theory of CV measurements naturally applies to sources of any number of independent optical modes. As a brief yet concrete demonstration, we shall discuss matters regarding two-mode sources. Their first-moment column

$$\boldsymbol{r} = \begin{pmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \end{pmatrix}, \quad \boldsymbol{r}_l \widehat{=} \begin{pmatrix} \langle X_l \rangle \\ \langle P_l \rangle \end{pmatrix} \tag{6.1}$$

collects all single-mode first-moment expectation values while their second-moment matrix takes the general form

$$\boldsymbol{G}_2 = \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{12}^{\mathrm{T}} & \boldsymbol{A}_2 \end{pmatrix} \geqslant -\frac{i}{2}\begin{pmatrix} \boldsymbol{\Omega} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Omega} \end{pmatrix},$$

$$\boldsymbol{A}_l \widehat{=} \begin{pmatrix} \langle X_l^2 \rangle & \frac{1}{2}\langle \{X_l, P_l\} \rangle \\ \frac{1}{2}\langle \{X_l, P_l\} \rangle & \langle P_l^2 \rangle \end{pmatrix}, \tag{6.2}$$

$$\boldsymbol{A}_{12} \widehat{=} \begin{pmatrix} \langle X_1 X_2 \rangle & \langle X_1 P_2 \rangle \\ \langle X_2 P_1 \rangle & \langle P_1 P_2 \rangle \end{pmatrix},$$

where the positive single-mode $\boldsymbol{A}_l$'s separately obey the HRS inequality and the, generally, nonsymmetric $\boldsymbol{A}_{12}$ accounts for nontrivial two-mode correlations for the two-mode quantum state ρ.

The estimation of $\boldsymbol{r}$ and $\boldsymbol{G}_2$ now involves the estimation of both the single-mode ($\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{A}_1, \boldsymbol{A}_2$) and two-mode ($\boldsymbol{A}_{12}$) components. The block structures of $\boldsymbol{r}$ and $\boldsymbol{G}_2$ suggest the natural situation where these five components are estimated with independent data, which we shall henceforth consider. In other words, the full data are invested equally into the estimation of all five components. Consequently, the corresponding scaled CRBs

$$\mathscr{H}_1 = \mathscr{H}_1^{(1)} + \mathscr{H}_1^{(2)}, \tag{6.3}$$

$$\mathscr{H}_2 = \mathscr{H}_2^{(1)} + \mathscr{H}_2^{(2)} + \mathscr{H}_2^{(12)} \tag{6.4}$$

reflect the independent additive contributions of the single-mode and two-mode components.

For two-mode sources, one can concurrently perform two separate CV measurements, one on each optical mode. The resulting joint measurement has a tensor-product structure, and it collects informationally complete data that directly probe the full quantum state. Analysis of this type provides closed-form expressions, which can be straightforwardly generalized to arbitrary number of modes, albeit with a rapidly increasing complexity. This method is by no means the only one for characterizing multimode sources. For more complex sources, it might be more economical to carry out tomography on linear superpositions of all the available modes using a single-mode measurement setup. Cross-correlation information may be acquired by studying different kinds of such superpositions. To this end, the general theory for single-mode sources presented in Sec. III is all we need to understand the accuracies in estimating moments for these superpositions. The robustness in controlling these superpositions as well as their high customizability are important prerequisites for practical multimode schemes [60].

A. Homodyne detection

If the two modes of ρ are physically fed into two separate homodyne setups, then the data collected must come from the rank-one measurement outcomes defined by the kets $|x_{\vartheta_1}\rangle|x'_{\vartheta_2}\rangle$, and the data scatter is characterized by the joint probabilities $p(x_{\vartheta_1}, \vartheta_1; x'_{\vartheta_2}, \vartheta_2)$ that are the Born probabilities of these outcomes. The two-mode expectation values are then given by

$$\langle X_{1,\vartheta_1}^m X_{2,\vartheta_2}^n \rangle = \int dx_{\vartheta_1} \int dx'_{\vartheta_2} \, x_{\vartheta_1}^m \, x_{\vartheta_2}'^n \, p\big(x_{\vartheta_1}, \vartheta_1; x'_{\vartheta_2}, \vartheta_2\big). \tag{6.5}$$

The single-mode $\langle X_{1,\vartheta_1}^m \rangle = \langle X_{\vartheta_1}^m \otimes 1 \rangle$ expectation values can hence be derived easily by first setting $x_{\vartheta_2}'^n = 1$ and, next, marginalizing the joint probabilities over the second mode. The procedure is similar for obtaining $\langle X_{2,\vartheta_2}^n \rangle = \langle 1 \otimes X_{\vartheta_2}^n \rangle$, so that we have

$$\langle X_{1,\vartheta_1}^m \rangle = \int dx_{\vartheta_1} \, x_{\vartheta_1}^m \, p^{(1)}\big(x_{\vartheta_1}, \vartheta_1\big),$$
$$\langle X_{2,\vartheta_2}^n \rangle = \int dx'_{\vartheta_2} \, x_{\vartheta_2}'^n \, p^{(2)}\big(x'_{\vartheta_2}, \vartheta_2\big), \tag{6.6}$$

where $p^{(1)}(x_{\vartheta_1}, \vartheta_1) = \int dx'_{\vartheta_2} \, p(x_{\vartheta_1}, \vartheta_1; x'_{\vartheta_2}, \vartheta_2)$, for instance.

The first and second moments can be fully characterized by the full two-mode homodyne data. For the first moments, $\langle X_{1,\vartheta_1} \rangle$ and $\langle X_{2,\vartheta_2} \rangle$ are parametrized by the pairs $(\langle X_1 \rangle, \langle P_1 \rangle)$ and $(\langle X_2 \rangle, \langle P_2 \rangle)$ needed to characterize $\boldsymbol{r}$. For the second moments, the joint expectation value $\langle X_{1,\vartheta_1} X_{2,\vartheta_2} \rangle$ contains all the four parameters of $\boldsymbol{A}_{12}$, and their marginalized versions $\langle X_{1,\vartheta_1}^2 \rangle$ and $\langle X_{2,\vartheta_2}^2 \rangle$ together contain all the six parameters of $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, all of which characterize the 10 independent parameters in $\boldsymbol{G}_2$.

With careful accounting of measurement resources, arguments similar to those that are responsible for Sec. III give the scaled Fisher matrix

$$\widetilde{\boldsymbol{F}}_{1,\mathrm{hom}}^{(l)} = \int_{(\pi)} \frac{d\vartheta_l}{\pi} \frac{\boldsymbol{m}_{\vartheta_l}}{\langle X_{l,\vartheta_l}^2 \rangle - \langle X_{l,\vartheta_l} \rangle^2} \tag{6.7}$$

first-moment single-mode components, and that for the second-moment single-mode components

$$\widetilde{\boldsymbol{F}}_{2,\mathrm{hom}}^{(l)} = \int_{(\pi)} \frac{d\vartheta_l}{\pi} \frac{\boldsymbol{M}_{\vartheta_l}}{\langle X_{l,\vartheta_l}^4 \rangle - \langle X_{l,\vartheta_l}^2 \rangle^2} \tag{6.8}$$

for $l = 1,2$. To figure out $\widetilde{\boldsymbol{F}}_{2,\mathrm{hom}}^{(12)}$, we observe that the MSE $\mathbb{E}[\mathrm{Tr}\{(\widehat{\boldsymbol{A}}_{12} - \boldsymbol{A}_{12})^{\mathrm{T}}(\widehat{\boldsymbol{A}}_{12} - \boldsymbol{A}_{12})\}] = \mathbb{E}[\widehat{\boldsymbol{a}}_{12}^{\mathrm{T}}\widehat{\boldsymbol{a}}_{12}] - \boldsymbol{a}_{12}^{\mathrm{T}}\boldsymbol{a}_{12}$ for the two-mode component $\boldsymbol{A}_{12}$ is identical to the MSE for the two-mode column $\boldsymbol{a}_{12} = (\langle X_1 X_2 \rangle \ \langle X_1 P_2 \rangle \ \langle X_2 P_1 \rangle \ \langle P_1 P_2 \rangle)^{\mathrm{T}}$.

The 4×4 Fisher matrix for estimating $\boldsymbol{A}_{12}$ now reads as

$$\widetilde{\boldsymbol{F}}_{2,\text{hom}}^{(12)} = \int_{(\pi)} \frac{d\vartheta_1}{\pi} \int_{(\pi)} \frac{d\vartheta_2}{\pi} \frac{\boldsymbol{m}_{\vartheta_1} \otimes \boldsymbol{m}_{\vartheta_2}}{\langle X_{1,\vartheta_1}^2 X_{2,\vartheta_2}^2 \rangle - \langle X_{1,\vartheta_1} X_{2,\vartheta_2} \rangle^2}.$$

(6.9)

B. Heterodyne detection

On the other hand, when the two modes are respectively fed into two separate heterodyne setups, the resulting measurement consists of all rank-one outcomes defined by product coherent states. In effect, this separable heterodyne detection directly samples the two-mode Husimi function. The sample mean of the phase-space data obtained, which tends to the two-mode Husimi function in the large-N limit, can again be shown to asymptotically converge to a joint Gaussian distribution, such that

$$\widehat{\boldsymbol{r}} = \frac{1}{N} \sum_{j=1}^{N} \begin{pmatrix} \widehat{\boldsymbol{r}}_{1,j} \\ \widehat{\boldsymbol{r}}_{2,j} \end{pmatrix}, \quad \widehat{\boldsymbol{r}}_{l,j} \cong \begin{pmatrix} x_{l,j} \\ p_{l,j} \end{pmatrix}$$

(6.10)

and

$$\widehat{\boldsymbol{G}}_{2,\text{het}} = \frac{1}{N} \sum_{j=1}^{N} \begin{pmatrix} \widehat{\boldsymbol{A}}_{1,j} & \widehat{\boldsymbol{A}}_{12,j} \\ \widehat{\boldsymbol{A}}_{12,j}^{\mathrm{T}} & \widehat{\boldsymbol{A}}_{2,j} \end{pmatrix},$$

$$\widehat{\boldsymbol{A}}_{l,j} \cong \begin{pmatrix} x_{l,j}^2 & x_{l,j} p_{l,j} \\ x_{l,j} p_{l,j} & p_{l,j}^2 \end{pmatrix},$$

$$\widehat{\boldsymbol{A}}_{12,j} \cong \begin{pmatrix} x_{1,j} x_{2,j} & x_{1,j} p_{2,j} \\ x_{2,j} p_{1,j} & p_{1,j} p_{2,j} \end{pmatrix}$$

(6.11)

become asymptotically efficient estimators for $\boldsymbol{r}$ and $\boldsymbol{G}_2$.

The relations between single- and two-mode heterodyne data share a similar behavior to those for the two-mode homodyne data. The Husimi average $\overline{x_1^m p_1^n}$, for example, can be obtained from the joint average $\overline{x_1^m p_1^n x_2^s p_2^t}$ by simply setting the phase-space values to 1 and marginalizing the two-mode Husimi function over the second mode, and *vice versa*. With these facts, we can immediately calculate the expressions for $\mathcal{H}_{1,\text{het}}$ and $\mathcal{H}_{2,\text{het}}$ by direct evaluations of the corresponding MSEs, as we did before in Sec. III B for single-mode sources. One then finds that

$$\mathcal{H}_{1,\text{het}} = \text{Var}_{\text{Q}}[x_1] + \text{Var}_{\text{Q}}[p_1] + \text{Var}_{\text{Q}}[x_2] + \text{Var}_{\text{Q}}[p_2],$$

$$\mathcal{H}_{2,\text{het}} = \text{Var}_{\text{Q}}[x_1^2] + \text{Var}_{\text{Q}}[p_1^2] + 2\text{Var}_{\text{Q}}[x_1 p_1]$$
$$+ \text{Var}_{\text{Q}}[x_2^2] + \text{Var}_{\text{Q}}[p_2^2] + 2\text{Var}_{\text{Q}}[x_2 p_2]$$
$$+ 2\text{Var}_{\text{Q}}[x_1 x_2] + 2\text{Var}_{\text{Q}}[x_1 p_2]$$
$$+ 2\text{Var}_{\text{Q}}[x_2 p_1] + 2\text{Var}_{\text{Q}}[p_1 p_2].$$

(6.12)

C. Optimality of heterodyne tomography in first-moment estimations

As in Eq. (3.5) for the single-mode case, we recognize here that $\mathcal{H}_{1,\text{hom}}$ takes the explicit expression

$$\mathcal{H}_{1,\text{hom}} = \text{Tr}\{\boldsymbol{G}^{(1)} + \boldsymbol{G}^{(2)}\} + 2(\sqrt{\det\{\boldsymbol{G}^{(1)}\}} + \sqrt{\det\{\boldsymbol{G}^{(2)}\}})$$

(6.13)

for any two-mode state ρ, where

$$\boldsymbol{G}^{(l)} \cong \begin{pmatrix} \langle (\Delta X_l)^2 \rangle & \frac{1}{2}\langle \{\Delta X_l, \Delta P_l\} \rangle \\ \frac{1}{2}\langle \{\Delta X_l, \Delta P_l\} \rangle & \langle (\Delta P_l)^2 \rangle \end{pmatrix}.$$

(6.14)

Since $\boldsymbol{G}^{(l)}$ also obeys the HRS uncertainty relation, one has $\det\{\boldsymbol{G}^{(l)}\} \geqslant \frac{1}{4}$, which again implies the universal inequality $\mathcal{H}_{1,\text{hom}} \geqslant \mathcal{H}_{1,\text{het}}$, or $\gamma_1 \leqslant 1$. Equality holds when both the marginalized $\boldsymbol{G}^{(1)}$ and $\boldsymbol{G}^{(2)}$ are, respectively, the covariance matrices of single-mode minimum-uncertainty states.

D. Two-mode Fock states

The first example to which we apply our two-mode theory is the class of two-mode Fock states $|n_1\rangle|n_2\rangle$. Since they are product states, the two-mode expectation values are product of single-mode expectation values. As such, the expressions for $\mathcal{H}_{1,\text{hom}}$ and $\mathcal{H}_{1,\text{het}}$ are easy to get, and they are

$$\mathcal{H}_{1,\text{hom}} = 4(n_1 + n_2 + 1),$$
$$\mathcal{H}_{1,\text{het}} = 2(n_1 + n_2 + 2).$$

(6.15)

When $n_1 = n_2 = 0$, ρ is a two-mode vacuum state, so that $\boldsymbol{G}^{(1)}$ and $\boldsymbol{G}^{(2)}$ are identically the covariance matrix of the single-mode vacuum state, where $\gamma_1 = 1$ as it should be. For any other values of n_1 and n_2, γ_1 is clearly subunity. Furthermore, since the scaled CRBs are clearly symmetric in n_1 and n_2, upon examining the limit $n_1 = n_2 \to \infty$, we find that the performance ratio approaches the constant $\frac{1}{2}$.

Likewise, the product nature of these states and their rotational symmetry allow us to comfortably obtain

$$\mathcal{H}_{2,\text{hom}} = 5(n_1^2 + n_2^2) + 21(n_1 + n_2) + 18,$$
$$\mathcal{H}_{2,\text{het}} = 2(n_1^2 + n_2^2) + 16(n_1 + n_2) + 20.$$

(6.16)

It turns out that although for the two-mode vacuum state, we have $\gamma_2 = 10/9 > 1$, the heterodyne scheme still beats the homodyne scheme for all other values of n_1 and n_2. The performance ratio γ_2 approaches the constant $\frac{2}{7}$ in the limit $n_1 = n_2 \to \infty$.

E. Two-mode squeezed vacuum state

The two-mode squeezed vacuum state, for a squeezing parameter $\zeta \geqslant 0$, is defined by the ket

$$|\text{SQV}\rangle = \frac{1}{\cosh(\zeta)} \sum_{n=0}^{\infty} |nn\rangle [\tanh(\zeta)]^n.$$

(6.17)

This two-mode entangled state is an important basic ingredient for many entanglement-manipulation and generation strategies for CV settings as well as quantum interferometry studies [61–64].

Even though this state is no longer a product state, the scaled CRBs still have closed-form expressions. In particular, for the first-moment scaled CRBs, we have

$$\mathcal{H}_{1,\text{hom}} = 4\cosh(2\zeta),$$
$$\mathcal{H}_{1,\text{het}} = 2\cosh(2\zeta) + 2.$$

(6.18)

We check that for $\zeta = 0$, we recover the scaled CRBs for the two-mode vacuum state, where again $\gamma_1 = 1$. For $\zeta > 0$, $\gamma < 1$ and tends to $\frac{1}{2}$ for large-ζ values.

This time, we witness that although all first moments for this entangled state are zero, the joint first moments need not. In particular, $\langle X_{1,\vartheta_1} X_{2,\vartheta_2} \rangle = \sinh(\zeta) \cosh(\zeta) \cos(\vartheta_1 + \vartheta_2)$ and $\overline{x_1 x_2} = \overline{p_1 p_2} = \sinh(\zeta) \cosh(\zeta)$. For the homodyne scheme, the scaled Fisher matrices $\widetilde{\boldsymbol{F}}_{2,\text{hom}}^{(1)}$, $\widetilde{\boldsymbol{F}}_{2,\text{hom}}^{(2)}$, and $\widetilde{\boldsymbol{F}}_{2,\text{hom}}^{(12)}$ can be calculated with the respective state variance functions

$$\left\langle X_{l,\vartheta_l}^4 \right\rangle - \left\langle X_{l,\vartheta_l}^2 \right\rangle^2 = \tfrac{1}{2}[\cosh(2\zeta)]^2,$$

$$\left\langle X_{1,\vartheta_1}^2 X_{2,\vartheta_2}^2 \right\rangle - \left\langle X_{1,\vartheta_1} X_{2,\vartheta_2} \right\rangle^2 = a \cos(2\vartheta_1 + 2\vartheta_2) + b,$$
(6.19)

with the coefficients $a = \tfrac{1}{2}[\sinh(2\zeta)]^2$ and $b = \tfrac{1}{4}[1 + 3\cosh(4\zeta)]$. The results are as follows:

$$\mathscr{H}_{2,\text{hom}} = \frac{11}{2} + \frac{13}{2}\cosh(4\zeta) + 6\cosh(2\zeta)\sqrt{\cosh(4\zeta)},$$

$$\mathscr{H}_{2,\text{het}} = 4[\cosh(\zeta)]^2[2 + 3\cosh(2\zeta)]. \qquad (6.20)$$

When $\zeta = 0$, the scaled CRBs indeed give the correct values for the two-mode vacuum state. From these formulas, we see that γ_2 is a monotonically decreasing function, starting from $\gamma_2|_{\zeta=0} = \frac{10}{9}$ all the way to the lowest value $\gamma_{2,\min} = 6/(13 + 6\sqrt{2}) \approx 0.279\,26$.

VII. CONCLUSION

We compare the moment-reconstruction performances of the homodyne and heterodyne measurement schemes using optimal moment estimators that minimize the mean-squared error. We first showed that in first-moment tomography, the heterodyne scheme is always tomographically superior to, or at least as good as, the homodyne scheme for *all* quantum states in terms of the mean-squared error of the moment estimators. The underlying physical reason is solely the Heisenberg-Robertson-Schrödinger uncertainty relation for complementary observables. For second-moment tomography, we showed that the heterodyne scheme can often outperform the homodyne scheme for Gaussian states and many other interesting and important classes of non-Gaussian states. All these states indicate a trend that a larger geometric mean of second-moment eigenvalues (second-moment "temperature") improves the moment-reconstruction accuracy with the heterodyne scheme relative to the homodyne scheme. This trend, however, is not monotonic in the second-moment "temperature" because there is also influence from the fourth moments originating from the form of the mean-squared error, the combined contributions of both give interesting features to the reconstruction accuracy, as illustrated by the examples in this article. The general theory introduced in Sec. III can be applied to higher-moment estimations that are important in general operator-moment applications and source-calibration protocols, and these shall be reported in the future.

ACKNOWLEDGMENTS

APPENDIX A: OPTIMAL ESTIMATORS FOR HOMODYNE TOMOGRAPHY

1. First-moment estimation

In this discussion, the reconstruction accuracy of the estimator $\widehat{\boldsymbol{r}}_{\text{hom}}$ for $\boldsymbol{r}$ shall be taken to be the usual MSE distance measure

$$\mathscr{D}_{1,\text{hom}} = \mathbb{E}[(\widehat{\boldsymbol{r}}_{\text{hom}} - \boldsymbol{r})^2] \qquad (A1)$$

that is typically defined for columns. One easy way to obtain an estimator $\widehat{\boldsymbol{r}}_{\text{hom}}$ is to make use of $\langle X_\vartheta \rangle = \langle X \rangle \cos\vartheta + \langle P \rangle \sin\vartheta$ to ascertain that

$$\boldsymbol{L}_\vartheta \boldsymbol{r} = \boldsymbol{r}_\vartheta \qquad (A2)$$

for an $n_\vartheta \times 2$ matrix $\boldsymbol{L}_\vartheta$ (n_ϑ being the number of bins for the phases ϑ) and a column $\boldsymbol{r}_\vartheta$ of n_ϑ true averages $\langle X_\vartheta \rangle$. The highly overcomplete nature of the measurement thus permits us to define, for any experimentally obtained estimates of average values $\boldsymbol{r}_\vartheta \equiv (\widehat{\langle X_1 \rangle} \ \widehat{\langle X_2 \rangle} \ \ldots \ \widehat{\langle X_{n_\vartheta} \rangle})^{\text{T}}$ ($\mathbb{E}[\widehat{\langle X_k \rangle}] = \langle X_{\vartheta_k} \rangle$),

$$\widehat{\boldsymbol{r}}_{\text{hom}}^{(\text{lin})} = \boldsymbol{L}_\vartheta^- \langle \boldsymbol{R}_\vartheta \rangle \qquad (A3)$$

as the linear estimator of interest using the pseudoinverse $\boldsymbol{L}_\vartheta^-$ of $\boldsymbol{L}_\vartheta$. This estimator, however, is suboptimal in the sense that it does *not* minimize the MSE $\mathscr{D}_{1,\text{hom}}$.

To obtain the optimal estimator for $\boldsymbol{r}$ [often known as the best linear unbiased estimator] that minimizes the MSE, we resort to the linear optimization of

$$\widehat{\boldsymbol{r}}_{\text{hom}} = \sum_{k=1}^{n_\vartheta} \boldsymbol{v}_k \widehat{\langle X_k \rangle} \qquad (A4)$$

over all possible *reconstruction columns* $\boldsymbol{v}_k$ for the estimates $\widehat{\langle X_k \rangle}$. Data consistency according to $\widehat{\langle X_k \rangle} = \boldsymbol{u}_k^{\text{T}} \widehat{\boldsymbol{r}}_{\text{HOM}}$ requires these reconstruction columns, or *dual columns*, to satisfy the property

$$\sum_{k=1}^{n_\vartheta} \boldsymbol{v}_k \boldsymbol{u}_k^{\text{T}} = \boldsymbol{1} = \sum_{k=1}^{n_\vartheta} \boldsymbol{u}_k \boldsymbol{v}_k^{\text{T}} \qquad (A5)$$

with the measurement columns $\boldsymbol{u}_k = \boldsymbol{u}_{\vartheta_k} = (\cos\vartheta_k \ \sin\vartheta_k)^{\text{T}}$. Logically, we must have

$$\mathbb{E}[\widehat{\boldsymbol{r}}_{\text{hom}}] = \sum_{k=1}^{n_\vartheta} \boldsymbol{v}_k \langle X_k \rangle = \boldsymbol{r}. \qquad (A6)$$

The Lagrange function for the optimization is therefore

$$\mathscr{L}_{\text{hom}} = \mathscr{D}_{\text{hom}} - \text{Tr}\left\{ \boldsymbol{\Lambda} \left(\sum_{k=1}^{n_\vartheta} \boldsymbol{u}_k \boldsymbol{v}_k^{\text{T}} - \boldsymbol{1} \right) \right\}, \qquad (A7)$$

where $\boldsymbol{\Lambda}$ is the Lagrange matrix for the dual-column constraint in (A5). In terms of the dual columns,

$$
\begin{aligned}
\mathscr{D}_{1,\mathrm{hom}} &= \sum_{k=1}^{n_\vartheta}\sum_{k'=1}^{n_\vartheta}\boldsymbol{v}_k^{\mathrm{T}}\boldsymbol{v}_{k'}\big(\mathbb{E}[\widehat{\langle X_k\rangle}\,\widehat{\langle X_{k'}\rangle}] - \langle X_k\rangle\langle X_{k'}\rangle\big)\\
&= \sum_{k=1}^{n_\vartheta}\boldsymbol{v}_k^{\mathrm{T}}\boldsymbol{v}_k\,\mathbb{E}[\widehat{\langle X_k\rangle}^2] + \sum_{k\neq k'}\boldsymbol{v}_k^{\mathrm{T}}\boldsymbol{v}_{k'}\,\mathbb{E}[\widehat{\langle X_k\rangle}]\,\mathbb{E}[\widehat{\langle X_{k'}\rangle}]\\
&\quad - \sum_{k=1}^{n_\vartheta}\sum_{k'=1}^{n_\vartheta}\boldsymbol{v}_k^{\mathrm{T}}\boldsymbol{v}_{k'}\langle X_k\rangle\langle X_{\vartheta_{k'}}\rangle.
\end{aligned}
\tag{A8}
$$

Since the unbiased estimate

$$
\widehat{\langle X_k\rangle} = \frac{1}{N_k}\sum_{j=1}^{n_x} n_{jk}x_{jk}
\tag{A9}
$$

is an average sum of all the measured n_x voltage readings x_{jk} per phase that are distributed according to the multinomial distribution of random multinomial weights $\sum_j n_{jk} = N_k$, the second moment is given by

$$
\begin{aligned}
\mathbb{E}[\widehat{\langle X_k\rangle}^2] &= \frac{1}{N_k^2}\sum_{j=1}^{n_x}\sum_{j'=1}^{n_x}\mathbb{E}[n_{jk}n_{j'k}]x_{jk}x_{j'k}\\
&= \frac{1}{N_k}\sum_{j=1}^{n_x} p_{jk}x_{jk}^2 + \frac{N_k-1}{N_k}\sum_{j=1}^{n_x}\sum_{j'=1}^{n_x} p_{jk}p_{j'k}x_{jk}x_{j'k}\\
&= \frac{1}{N_k}\langle X_k^2\rangle + \frac{N_k-1}{N_k}\langle X_k\rangle^2.
\end{aligned}
\tag{A10}
$$

The final equality is valid for sufficiently large data (bins) for all phases, as $p_{jk} \to dx_\vartheta\, p(x_\vartheta,\vartheta)$ and

$$
\begin{aligned}
\sum_{j=1}^{n_x} p_{jk}x_{jk}^2 &\to \int dx_\vartheta\, p(x_\vartheta,\vartheta)\,x_\vartheta^2\\
&= \int dx_\vartheta\,\langle|x_\vartheta\rangle\langle x_\vartheta|\rangle x_\vartheta^2 = \langle X_\vartheta^2\rangle.
\end{aligned}
\tag{A11}
$$

So, we finally get

$$
\mathscr{D}_{1,\mathrm{hom}} = \sum_{k=1}^{n_\vartheta}\frac{\boldsymbol{v}_k^{\mathrm{T}}\boldsymbol{v}_k}{N_k}\big(\langle X_k^2\rangle - \langle X_k\rangle^2\big).
\tag{A12}
$$

A simple variation of $\mathscr{L}_{\mathrm{hom}}$ therefore gives

$$
\begin{aligned}
\delta\mathscr{L}_{\mathrm{hom}} &= \sum_{k=1}^{n_\vartheta}\frac{\delta\boldsymbol{v}_k^{\mathrm{T}}\boldsymbol{v}_k + \boldsymbol{v}_k^{\mathrm{T}}\delta\boldsymbol{v}_k}{N_k}\big(\langle X_k^2\rangle - \langle X_k\rangle^2\big)\\
&\quad - \frac{1}{2}\mathrm{Tr}\left\{\boldsymbol{\Lambda}\sum_{k=1}^{n_\vartheta}\big(\boldsymbol{u}_k\delta\boldsymbol{v}_k^{\mathrm{T}} + \delta\boldsymbol{v}_k\boldsymbol{u}_k^{\mathrm{T}}\big)\right\} \equiv 0
\end{aligned}
\tag{A13}
$$

or

$$
\begin{aligned}
\frac{1}{2}\boldsymbol{\Lambda} &= \boldsymbol{F}\big(\{\langle X_k\rangle,\langle X_k^2\rangle\}\big) \equiv \sum_{k=1}^{n_\vartheta}\boldsymbol{u}_k\boldsymbol{u}_k^{\mathrm{T}}\frac{N_k}{\langle X_k^2\rangle - \langle X_k\rangle^2},\\
\boldsymbol{v}_k &= \frac{N_k}{\langle X_k^2\rangle - \langle X_k\rangle^2}\boldsymbol{F}\big(\{\langle X_k\rangle,\langle X_k^2\rangle\}\big)^{-1}\boldsymbol{u}_k.
\end{aligned}
\tag{A14}
$$

The matrix $\boldsymbol{F}(\{\langle X_k\rangle,\langle X_k^2\rangle\})$ is known as the frame matrix.

The optimal estimator therefore depends on the true moments which are certainly unavailable in the first place, for no tomography is otherwise necessary at all. Nonetheless, one can substitute the estimated moments for them to obtain an asymptotically efficient optimal estimator that approximates the optimal estimator. An unbiased estimate for the second moment is given by

$$
\widehat{\langle X_k^2\rangle} = \frac{1}{N_k}\sum_{j=1}^{n_x} n_{jk}x_{jk}^2,
\tag{A15}
$$

so that the asymptotically optimal estimator is given by

$$
\begin{aligned}
\widehat{\boldsymbol{r}}_{\mathrm{hom}}^{(\mathrm{opt})} &= \boldsymbol{W}_1^{-1}\sum_{k=1}^{n_\vartheta}\boldsymbol{u}_k\frac{N_k\widehat{\langle X_k\rangle}}{\widehat{\langle X_k^2\rangle} - \widehat{\langle X_k\rangle}^2},\\
\boldsymbol{W}_1 &= \sum_{k=1}^{n_\vartheta}\boldsymbol{m}_k\frac{N_k}{\widehat{\langle X_k^2\rangle} - \widehat{\langle X_k\rangle}^2}.
\end{aligned}
\tag{A16}
$$

It is easy to see that when the estimated moments approach the true moments, this optimal estimator attains the scaled CRB. Directly from Eq. (A12), we immediately know that its corresponding MSE is given by

$$
\mathscr{D}_{1,\mathrm{hom}}^{(\mathrm{opt})} = \mathrm{Tr}\big\{\boldsymbol{F}\big(\{\langle X_k\rangle,\langle X_k^2\rangle\}\big)^{-1}\big\}
\tag{A17}
$$

and all we need to realize is that for sufficiently large N and uniformly distributed quadrature outcomes, $N_k/N \to d\vartheta/\pi$ and the frame matrix

$$
\frac{1}{N}\boldsymbol{F}\big(\{\langle X_k\rangle,\langle X_k^2\rangle\}\big) \to \int_{(\pi)}\frac{d\vartheta}{\pi}\frac{\boldsymbol{m}_\vartheta}{\langle X_\vartheta^2\rangle - \langle X_\vartheta\rangle^2} = \widetilde{\boldsymbol{F}}_{1,\mathrm{hom}}
\tag{A18}
$$

is nothing more than the Fisher matrix introduced in Eq. (3.3). This also means that the optimal estimator and the asymptotically optimal estimator are both asymptotically as efficient as the maximum-likelihood estimator.

This construction comes with a basic and important lesson. The simple linear estimator $\widehat{\boldsymbol{r}}_{\mathrm{hom}}^{(\mathrm{lin})}$ in Eq. (A3), which is suboptimal, depends only on the first moments. To improve the reconstruction accuracy, more aspects of the data that are attributed to the figure of merit chosen to measure this accuracy would have to be incorporated systematically. In the case of the MSE, these are linear combinations of both the first and second moments, or at least their estimates. Put differently, we should always use the reconstruction estimator that optimize the figure of merit we choose to rank the goodness of the reconstruction.

2. Second-moment estimation

By the same token, we can construct the optimal estimator that approximates the optimal estimator for second-moment estimation by minimizing the MSE

$$
\mathscr{D}_{2,\mathrm{hom}} = \mathbb{E}[\mathrm{Tr}\{(\widehat{\boldsymbol{G}}_{2,\mathrm{hom}} - \boldsymbol{G}_{2,\mathrm{hom}})^2\}]
\tag{A19}
$$

over the estimator that is of the linear form

$$
\widehat{\boldsymbol{G}}_{2,\mathrm{hom}} = \sum_{k=1}^{n_\vartheta}\boldsymbol{\Theta}_k\widehat{\langle X_k^2\rangle}
\tag{A20}
$$

with respect to the second-moment estimates. This form is a natural extension to the column estimator $\widehat{\boldsymbol{r}}_{\mathrm{hom}}$ *via* a generalization of the dual columns $\boldsymbol{v}_k$ to *dual* matrices $\boldsymbol{\Theta}_k$.

Completely analogous to the discussion in Appendix A 1, consistency with $\langle \widetilde{X_k^2} \rangle = \boldsymbol{u}_k^{\mathrm{T}} \widehat{\boldsymbol{G}}_{2,\mathrm{hom}} \boldsymbol{u}_k$ implies that

$$\widehat{\boldsymbol{G}}_{2,\mathrm{hom}} = \sum_{k=1}^{n_\vartheta} \boldsymbol{\Theta}_k \boldsymbol{u}_k^{\mathrm{T}} \widehat{\boldsymbol{G}}_{2,\mathrm{hom}} \boldsymbol{u}_k. \tag{A21}$$

The above relation can be simplified by introducing the *vectorization* notation vec($\boldsymbol{Y}$) that turns a matrix $\boldsymbol{Y}$ into a column. Since all two-dimensional matrices considered here are real and symmetric, they are essentially characterized by three real parameters. Hence, in our context, given that

$$\boldsymbol{Y} \,\widehat{=}\, \begin{pmatrix} y_1 & y_2 \\ y_2 & y_3 \end{pmatrix}, \tag{A22}$$

the vectorized quantity is defined as

$$\mathrm{vec}(\boldsymbol{Y}) \,\widehat{=}\, \begin{pmatrix} y_1 \\ \sqrt{2}\,y_2 \\ y_3 \end{pmatrix}. \tag{A23}$$

This operation is a variant of the usual column-stacking vectorization operation to apply on 2×2 real symmetric matrices for our case to make contact with the property $\mathrm{Tr}\{\boldsymbol{Y}_1 \boldsymbol{Y}_2\} = \mathrm{vec}(\boldsymbol{Y}_1)^{\mathrm{T}} \mathrm{vec}(\boldsymbol{Y}_2)$ between any pair of such matrices $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$. In this notation, Eq. (A21) becomes

$$\mathrm{vec}(\widehat{\boldsymbol{G}}_{2,\mathrm{hom}}) = \sum_{k=1}^{n_\vartheta} \mathrm{vec}(\boldsymbol{\Theta}_k) \mathrm{vec}(\boldsymbol{m}_k)^{\mathrm{T}} \mathrm{vec}(\widehat{\boldsymbol{G}}_{2,\mathrm{hom}}), \tag{A24}$$

which is equivalent to the vectorized constraint

$$\sum_{k=1}^{n_\vartheta} \mathrm{vec}(\boldsymbol{\Theta}_k) \mathrm{vec}(\boldsymbol{m}_k)^{\mathrm{T}} = \mathbf{1} = \sum_{k=1}^{n_\vartheta} \mathrm{vec}(\boldsymbol{m}_k) \mathrm{vec}(\boldsymbol{\Theta}_k)^{\mathrm{T}}. \tag{A25}$$

As usual, to derive the expression for the optimal estimator, we first calculate $\mathscr{D}_{2,\mathrm{hom}}$ in terms of the dual matrices. For this we shall need the average of the square of the estimate $\langle \widetilde{X_k^2} \rangle$ defined in Eq. (A15):

$$\mathbb{E}\left[\langle \widetilde{X_k^2} \rangle^2\right] = \frac{1}{N_k}\langle X_k^4 \rangle + \frac{N_k - 1}{N_k}\langle X_k^2 \rangle^2, \tag{A26}$$

from which gives the expression

$$\mathscr{D}_{2,\mathrm{hom}} = \sum_{k=1}^{n_\vartheta} \frac{\mathrm{vec}(\boldsymbol{\Theta})_k^{\mathrm{T}} \mathrm{vec}(\boldsymbol{\Theta})_k}{N_k}\left(\langle X_k^4 \rangle - \langle X_k^2 \rangle^2\right) \tag{A27}$$

for the MSE. Then, by carrying out the variation of the appropriate Lagrange function similar to the calculations in Appendix A 1 and remembering the additional association $\boldsymbol{M}_k = \mathrm{vec}(\boldsymbol{m}_k)\mathrm{vec}(\boldsymbol{m}_k)^{\mathrm{T}}$, we find that the matrices for the optimal estimator are

$$\boldsymbol{F}\left(\{\langle X_k^2 \rangle, \langle X_k^4 \rangle\}\right) \equiv \sum_{k=1}^{n_\vartheta} \boldsymbol{M}_k \frac{N_k}{\langle X_k^4 \rangle - \langle X_k^2 \rangle^2},$$

$$\mathrm{vec}(\boldsymbol{\Theta}_k) = \frac{N_k}{\langle X_k^4 \rangle - \langle X_k^2 \rangle^2} \boldsymbol{F}\left(\{\langle X_k^2 \rangle, \langle X_k^4 \rangle\}\right)^{-1} \mathrm{vec}(\boldsymbol{m}_k). \tag{A28}$$

Finally, the asymptotically optimal estimator is given by

$$\widehat{\boldsymbol{G}}_{2,\mathrm{hom}}^{(\mathrm{opt})} = \boldsymbol{W}_2^{-1} \sum_{k=1}^{n_\vartheta} \mathrm{vec}(\boldsymbol{m}_k) \frac{N_k \langle \widehat{X_k^2} \rangle}{\langle \widehat{X_k^4} \rangle - \langle \widehat{X_k^2} \rangle^2},$$

$$\boldsymbol{W}_2 = \sum_{k=1}^{n_\vartheta} \boldsymbol{M}_k \frac{N_k}{\langle \widehat{X_k^4} \rangle - \langle \widehat{X_k^2} \rangle^2}. \tag{A29}$$

That this estimator asymptotically attains the scaled CRB for second-moment estimation is again clear.

APPENDIX B: LIST OF CHARACTERISTIC FUNCTIONS

In calculating the moments for both the homodyne and heterodyne schemes, it is extremely useful to start with the relevant characteristic functions for both schemes. To facilitate the discussions in the main article, we have supplied a list of quadrature characteristic functions ($\langle e^{ikX_\vartheta} \rangle$) for the homodyne scheme and a list of Husimi characteristic functions $\left[\overline{e^{g^*\alpha + g\alpha^*}}, g = (u + iv)/\sqrt{2}\right]$ for the heterodyne scheme respectively in Table I in this appendix. Then, the two kinds of moments can then be readily computed by the prescriptions

$$\langle X_\vartheta^m \rangle = \left(-i\frac{\partial}{\partial k}\right)^m \langle e^{ikX_\vartheta} \rangle\Big|_{k=0},$$

$$\overline{x^k p^l} = \left(\frac{\partial}{\partial u}\right)^k \left(\frac{\partial}{\partial v}\right)^l \overline{e^{g^*\alpha + g\alpha^*}}\Big|_{u,v=0}. \tag{B1}$$

TABLE I. A list of characteristic functions for all the quantum states discussed. The symbols in this table are defined as $\alpha_0 e^{-i\vartheta} = (x_\vartheta + ip_\vartheta)/\sqrt{2}$ where $x_0 = x_{\vartheta=0}$ and $p_0 = p_{\vartheta=0}$, $\boldsymbol{g} \,\widehat{=}\, (-g \; g^*)^{\mathrm{T}}$, $\boldsymbol{M} = \boldsymbol{H}^\dagger \boldsymbol{G}_{\mathrm{het}}^{-1} \boldsymbol{H}$, and $\boldsymbol{H} \,\widehat{=}\, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$.

Class of quantum states	Quadrature characteristic function	Husimi characteristic function
Gaussian	$\exp\left(-\frac{1}{2}\left(\boldsymbol{u}_\vartheta^{\mathrm{T}} \boldsymbol{G} \boldsymbol{u}_\vartheta\right)^2 k^2 + i\,\boldsymbol{u}_\vartheta^{\mathrm{T}} \boldsymbol{r}_0\, k\right)$	$e^{g^*\alpha_0 + g\alpha_0^*} \exp\left(\frac{\det\{\boldsymbol{G}_{\mathrm{het}}\}}{2}\boldsymbol{g}^\dagger \boldsymbol{M} \boldsymbol{g}\right)$
Fock	$e^{-\frac{k^2}{4}} L_n\left(\frac{k^2}{2}\right)$	$_1F_1(n+1; 1; \|g\|^2)$
Even and odd coherent	$e^{-\frac{k^2}{4}} \dfrac{\cos(kx_\vartheta) \pm e^{-2\|\alpha_0\|^2}\cosh(kp_\vartheta)}{1 \pm e^{-2\|\alpha_0\|^2}}$	$\dfrac{e^{-\|\alpha_0\|^2}}{2 \pm 2\,e^{-2\|\alpha_0\|^2}}\left[\begin{array}{c} e^{\|g+\alpha_0\|^2} + e^{\|g-\alpha_0\|^2} \\ \pm e^{(g^*-\alpha_0^*)(g+\alpha_0)} \pm \mathrm{c.c.} \end{array}\right]$
Displaced Fock	$e^{-\frac{k^2}{4} + ikx_\vartheta} L_m\left(\frac{k^2}{2}\right)$	$e^{g^*\alpha_0 + g\alpha_0^*}{}_1F_1(m+1; 1; \|g\|^2)$
Photon-added coherent	$e^{\frac{k^2}{4}} \dfrac{{}_1F_1\left(m+1; 1; \left(\alpha_0 + \frac{ik}{\sqrt{2}}e^{i\vartheta}\right)\left(\alpha_0^* + \frac{ik}{\sqrt{2}}e^{-i\vartheta}\right)\right)}{{}_1F_1(m+1; 1; \|\alpha_0\|^2)}$	$\dfrac{{}_1F_1(m+1; 1; \|g+\alpha_0\|^2)}{{}_1F_1(m+1; 1; \|\alpha_0\|^2)}$

which simply involves multiple differentiations with respect to the free variables and later setting these variables to zero. The identity $_1F_1(n+1; 1; -x) = e^{-x}L_n(x)$ relating the confluent hypergeometric functions and Laguerre polynomials is particularly useful for consistency verification between two characteristic functions of different quantum states.

[1] S. L. Braunstein and P. van Loock, Quantum information with continuous variables, Rev. Mod. Phys. **77**, 513 (2005).

[2] A. Ferraro, S. Olivares, and M. G. A. Paris, *Gaussian States in Continuous Variable Quantum Information* (Bibliopolis, Napoli, 2005).

[3] *Quantum Information with Continuous Variables of Atoms and Light*, edited by N. Cerf, G. Leuchs, and E. S. Polzik (Imperial College Press, London, 2007).

[4] U. L. Andersen, G. Leuchs, and C. Silberhorn, Continuous-variable quantum information processing, Laser Photon. Rev. **4**, 337 (2010).

[5] G. Adesso, S. Ragy, and A. R. Lee, Continuous variable quantum information: Gaussian states and beyond, Open Syst. Inf. Dyn. **21**, 1440001 (2014).

[6] H. P. Yuen and V. W. S. Chan, Noise in homodyne and heterodyne detection, Opt. Lett. **8**, 177 (1983).

[7] G. L. Abbas, V. W. S. Chan, and T. K. Yee, Local-oscillator excess-noise suppression for homodyne and heterodyne detection, Opt. Lett. **8**, 419 (1983).

[8] B. L. Schumaker, Noise in homodyne detection, Opt. Lett. **9**, 189 (1984).

[9] K. Banaszek and K. Wódkiewicz, Operational theory of homodyne detection, Phys. Rev. A **55**, 3117 (1997).

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

[11] E. Arthurs and J. L. Kelly, On the simultaneous measurement of a pair of conjugate observables, Bell Syst. Technol. J. **44**, 725 (1965).

[12] H. P. Yuen, Generalized quantum measurements and approximate simultaneous measurements of noncommuting observables, Phys. Lett. A **91**, 101 (1982).

[13] E. Arthurs and M. S. Goodman, Quantum Correlations: A Generalized Heisenberg Uncertainty Relation, Phys. Rev. Lett. **60**, 2447 (1988).

[14] H. Martens and W. M. de Muynck, The inaccuracy principle, Found. Phys. **20**, 357 (1990).

[15] H. Martens and W. M. de Muynck, Towards a new uncertainty principle: Quantum measurement noise, Phys. Lett. A **157**, 441 (1991).

[16] M. G. Raymer, Uncertainty principle for joint measurement of noncommuting variables, Am. J. Phys. **62**, 986 (1994).

[17] A. Trifonov, G. Björk, and J. Söderholm, Simultaneous Minimum-Uncertainty Measurement of Discrete-Valued Complementary Observables, Phys. Rev. Lett. **86**, 4423 (2001).

[18] R. F. Werner, The uncertainty relation for joint measurement of position and momentum, Quantum Inf. Comput. **4**, 546 (2004).

[19] S. Stenholm, Simultaneous measurement of conjugate variables, Ann. Phys. (NY) **218**, 233 (1992).

[20] A Javan, E. A. Ballik, and W. L. Bond, Frequency characteristics of a continuous-wave He–Ne optical maser, J. Opt. Soc. Am. **52**, 96 (1962).

[21] W. S. Read and R. G. Turner, Tracking heterodyne detection, Appl. Opt. **4**, 1570 (1965).

[22] H. R. Carleton and W. T. Maloney, A balanced optical heterodyne detector, Appl. Opt. **7**, 1241 (1968).

[23] H. Gerhardt, H. Welling, and A. Güttner, Measurements of the laser linewidth due to quantum phase and quantum amplitude noise above and below threshold, I, Z. Phys. **253**, 113 (1972).

[24] H. Yuen and J. H. Shapiro, Optical communication with two-photon coherent states–part III: Quantum measurements realizable with photoemissive detectors, IEEE Trans. Inf. Theory **26**, 78 (1980).

[25] J. H. Shapiro and S. Wagner, Phase and amplitude uncertainties in heterodyne detection, IEEE J. Quantum Electron. **QE20**, 803 (1984).

[26] J. Shapiro, Quantum noise and excess noise in optical homodyne and heterodyne receivers, IEEE J. Quantum Electron. **QE21**, 237 (1985).

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

[28] M. J. Collett, R. Loudon, and C. W. Gardiner, Quantum theory of optical homodyne and heterodyne detection, J. Mod. Opt. **34**, 881 (1987).

[29] Y. Lai and H. A. Haus, Characteristic functions and quantum measurements of optical observables, Quantum Opt. **1**, 99 (1989).

[30] T. Kiesel, Statistical analysis of sampling methods in quantum tomography, Phys. Rev. A **85**, 052114 (2012).

[31] S. Lorenz, N. Korolkova, and G. Leuchs, Continuous variable quantum key distribution using polarization encoding and post selection, Appl. Phys. B **79**, 273 (2004).

[32] A. M. Lance, T. Symul, V. Sharma, C. Weedbrook, T. C. Ralph, and P. K. Lam, No-Switching Quantum Key Distribution Using Broadband Modulated Coherent Light, Phys. Rev. Lett. **95**, 180503 (2005).

[33] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, The security of practical quantum key distribution, Rev. Mod. Phys. **81**, 1301 (2009).

[34] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, Rev. Mod. Phys. **84**, 621 (2012).

[35] J. Řeháček, S. Olivares, D. Mogilevtsev, Z. Hradil, M. G. A. Paris, S. Fornaro, V. D'Auria, A. Porzio, and S. Solimeno, Effective method to estimate multidimensional Gaussian states, Phys. Rev. A **79**, 032111 (2009).

[36] J. C. Angulo, Uncertainty relationships in many-body systems, J. Phys. A: Math. Gen. **26**, 6493 (1993).

[37] J. C. Angulo, Information entropy and uncertainty in *d*-dimensional many-body systems, Phys. Rev. A **50**, 311 (1994).

[38] R. Simon, M. Selvadoray, Arvind, and N. Mukunda, Necessary and sufficient classicality conditions on photon number distributions, arXiv:quant-ph/9709030v1.

[39] Arvind, N. Mukunda, and R. Simon, Characterisations of classical and non-classical states of quantised radiation, J. Phys. A: Math. Gen. **31**, 565 (1998).

[40] R. Namiki, Photonic families of non-Gaussian entangled states and entanglement criteria for continuous-variable systems, Phys. Rev. A **85**, 062307 (2012).

[41] J. S. Ivan, N. Mukunda, and R. Simon, Generation and distillation of non-gaussian entanglement from nonclassical photon statistics, Quantum Inf. Process. **11**, 873 (2012).

[42] A. Leverrier, Symmetrization technique for continuous-variable quantum key distribution, Phys. Rev. A **85**, 022339 (2012).

[43] O. Thearle, S. M. Assad, and T. Symul, Estimation of output-channel noise for continuous-variable quantum key distribution, Phys. Rev. A **93**, 042343 (2016).

[44] J. Řeháček, Y. S. Teo, Z. Hradil, and S. Wallentowitz, Surmounting intrinsic quantum-measurement uncertainties in gaussian-state tomography with quadrature squeezing, Sci. Rep. **5**, 12289 (2015).

[45] C. R. Müller, C. Peuntinger, T. Dirmeier, I. Khan, U. Vogl, Ch. Marquardt, G. Leuchs, L. L. Sánchez-Soto, Y. S. Teo, Z. Hradil, and J. Řeháček, Evading Vacuum Noise: Wigner Projections or Husimi Samples?, Phys. Rev. Lett. **117**, 070801 (2016).

[46] J Heersink, V. Josse, G. Leuchs, and U. L. Andersen, Efficient polarization squeezing in optical fibers, Opt. Lett. **30**, 1192 (2005).

[47] P. Grangier, R. E. Slusher, B. Yurke, and A. LaPorta, Squeezed-Light–Enhanced Polarization Interferometer, Phys. Rev. Lett. **59**, 2153 (1987).

[48] V. Josse, A. Dantan, A. Bramati, and E. Giacobino, Entanglement and squeezing in a two-mode system: theory and experiment, J. Opt. B **6**, S532 (2004).

[49] Ch. Marquardt, J. Heersink, R. Dong, M. V. Chekhova, A. B. Klimov, L. L. Sánchez-Soto, U. L. Andersen, and G. Leuchs, Quantum Reconstruction of an Intense Polarization Squeezed Optical State, Phys. Rev. Lett. **99**, 220401 (2007).

[50] C. R. Müller, B. Stoklasa, C. Peuntinger, C. Gabriel, J. Řeháček, Z. Hradil, A. B. Klimov, G. Leuchs, Ch. Marquardt, and L. L. Sánchez-Soto, Quantum polarization tomography of bright squeezed light, New J. Phys. **14**, 085002 (2012).

[51] C. Peuntinger, B. Heim, C. R. Müller, C. Gabriel, Ch. Marquardt, and G. Leuchs, Distribution of Squeezed States Through an Atmospheric Channel, Phys. Rev. Lett. **113**, 060502 (2014).

[52] D. R. Cox, *Principles of Statistical Inference* (Cambridge University Press, Cambridge, 2006).

[53] H. Zhu, Quantum state estimation with informationally over-complete measurements, Phys. Rev. A **90**, 012115 (2014).

[54] Y. S. Teo, *Introduction to Quantum-State Estimation* (World Scientific, Singapore, 2015).

[55] M. G. A. Paris, Displacement operator by beam splitter, Phys. Lett. A **217**, 78 (1996).

[56] K. Banaszek and K. Wódkiewicz, Testing Quantum Nonlocality in Phase Space, Phys. Rev. Lett. **82**, 2009 (1999).

[57] V. Parigi, A. Zavatta, M. S. Kim, and M. Bellini, Probing quantum commutation rules by addition and subtraction of single photons to/from a light field, Science **317**, 1890 (2007).

[58] M. S. Kim, H. Jeong, A. Zavatta, V. Parigi, and M. Bellini, Scheme for Proving the Bosonic Commutation Relation using Single-Photon Interference, Phys. Rev. Lett. **101**, 260401 (2008).

[59] A. Zavatta, V. Parigi, M. S. Kim, H. Jeong, and M. Bellini, Experimental Demonstration of the Bosonic Commutation Relation via Superpositions of Quantum Operations on Thermal Light Fields, Phys. Rev. Lett. **103**, 140406 (2009).

[60] D. Mogilevtsev, J. Reháček, and Z. Hradil, Relative tomography of an unknown quantum state, Phys. Rev. A **79**, 020101(R) (2009).

[61] P. van Loock and S. L. Braunstein, Unconditional teleportation of continuous-variable entanglement, Phys. Rev. A **61**, 010302(R) (1999).

[62] M. Ban, Quantum dense coding via a two-mode squeezed-vacuum state, J. Opt. B: Quantum Semiclass. Opt. **1**, L9 (1999).

[63] K. Huang, H. Le Jeannic, J. Ruaudel, V. B. Verma, M. D. Shaw, F. Marsili, S. W. Nam, E. Wu, H. Zeng, Y.-C. Jeong, R. Filip, O. Morin, and J. Laurat, Optical Synthesis of Large-Amplitude Squeezed Coherent-State Superpositions with Minimal Resources, Phys. Rev. Lett. **115**, 023602 (2015).

[64] I. Ruo-Berchera, I. P. Degiovanni, S. Olivares, N. Samantaray, P. Traina, and M. Genovese, One- and two-mode squeezed light in correlated interferometry, Phys. Rev. A **92**, 053821 (2015).