

Statistical parameter estimation of multimode multiphoton-subtracted thermal states of light

G. V. Avosopiants^{1,*}, B. I. Bantysh², K. G. Katamadze^{1,2}, N. A. Bogdanova², Yu. I. Bogdanov², and S. P. Kulik¹

¹*Quantum Technology Centre, Faculty of Physics, M. V. Lomonosov Moscow State University, 119991 Moscow, Russia*

²*Valiev Institute of Physics and Technology, Russian Academy of Sciences, 117218 Moscow, Russia*



(Received 19 February 2021; revised 2 June 2021; accepted 24 June 2021; published 12 July 2021)

Thermal states of light are widely used in quantum optics for testing of various quantum phenomena. Particularly, they can be utilized for characterization of photon creation and photon annihilation operations. During the last decade the problem of photon subtraction from multimode quantum states has become of much significance. Therefore in this paper we present a technique for statistical parameter estimation of multimode multiphoton-subtracted thermal states of light, which can be used for the multimode photon annihilation test.

DOI: [10.1103/PhysRevA.104.013710](https://doi.org/10.1103/PhysRevA.104.013710)

I. INTRODUCTION

Photon creation and annihilation operators are base elements of quantum optics. Despite the fact that they are non-Hermitian and nonunitary, they can be directly (but probabilistically) implemented [1–4]. Thereby we get a perfect toolbox allowing tests of basic commutation rules [3], Schrödinger’s cat and other non-Gaussian quantum state preparations [1,2,5], probabilistic linear noiseless amplification [6], strong Kerr nonlinearity implementation [7], etc.

Thermal states of light are easy to prepare, and their statistics is modified significantly by both photon creation and photon annihilation. Therefore photon-subtracted thermal states (PSTs) become very attractive for the demonstration of effects in quantum optics and quantum thermodynamics such as the quantum vampire effect [8,9], photonic Maxwell’s demon [10], quantum thermal engine [11], etc. Moreover, it was shown that PSTs can be utilized in some metrological applications [12,13].

Recently, the action of non-Gaussian operations (particularly photon creation and annihilation) on the multimode states of light has become very interesting in the context of cluster-state quantum computing [14,15]. Despite the fact that there are some mode-selective photon subtraction techniques [15,16], generally the annihilation operator is implemented using a low-reflective beam splitter and a photon detector in the reflected channel [1–3]. In this case we cannot control in which optical mode the photon is subtracted.

In this paper, we consider the photon subtraction from a multimode quantum state and study only a part of the output modes (Fig. 1). The case of single-mode detection is of particular interest in, e.g., homodyne detection. We study an example of a multimode thermal state at the input and the general case of multiple-photon detection in the reflected channel that corresponds to multiple-photon subtraction.

Previously, it has been theoretically shown [17] and experimentally verified [18–20] that the photon number distribution of a K -photon-subtracted M -mode thermal state can be described by a negative binomial, or a compound Poisson distribution [21] $P_{cP}(N|\mu_0, a)$ with two parameters, the group parameter $a = K + M$ and the initial per-mode mean photon number μ_0 [17,20,22–24]:

$$P_{cP}(N|\mu_0, a) = \frac{\Gamma(a + N)}{\Gamma(a)} \frac{\mu_0^N}{N!} \left(\frac{1}{1 + \mu_0} \right)^{N+a}. \quad (1)$$

The mean photon number of this distribution is $\mu = \mu_0 a$.

Thus it is impossible to determine separately the number of modes M and the number of subtracted photons K by examining the total photocount statistics in all modes. One can only get the sum of these parameters. This creates problems in the case when we need to determine the parameters of multimode PSTs (MPSTs).

However, the situation changes somewhat if we consider only $m < M$ modes of the state described by the distribution (1). Then the resulting photocount distribution in such a subsystem is the convolution of the compound Poisson distribution (1) and the Pólya distribution $P_{\text{Polya}}(k|m, M, K)$:

$$P(N|\mu_0, m, M, K) = \sum_{k=0}^K P_{\text{Polya}}(k|m, M, K) \times P_{cP}(N|\mu_0, a = k + M), \quad (2)$$

where $P_{\text{Polya}}(k|m, M, K) = \frac{C_{m+k-1}^k C_{M-m+K-k-1}^{K-k}}{C_{M+K-1}^k}$ [21,25–27].

The convolution in (2) is easy to calculate using the generating-function approach. The generating function of the distribution (2) is as follows [28]:

$$G(z|\mu_0, m, M, K) = [G_{\text{BE}}(z|\mu_0)]^m \times {}_2F_1(-K, m, M, 1 - G_{\text{BE}}(z|\mu_0)). \quad (3)$$

*avosopyantsgrant@gmail.com

Here, $G_{\text{BE}}(z|\mu_0) = [1 + \mu_0(1 - z)]^{-1}$ is the generating function of the thermal state (Bose-Einstein distribution), ${}_2F_1$ is the Gaussian hypergeometric function (for more details, see our work [26]). The corresponding photon number statistics is

$$P(N|\mu_0, m, M, K) = \frac{\mu_0^N}{(1 + \mu_0)^{N+m}} \frac{1}{\Gamma(m)} \frac{\Gamma(N+m)}{\Gamma(N+1)} \frac{\Gamma(M)}{\Gamma(M-m)} \frac{\Gamma(M+K-m)}{\Gamma(M+K)} \times {}_2F_1\left(-K, N+m, -K-M+m+1, \frac{1}{1+\mu_0}\right). \quad (4)$$

In previous work [26], we have shown that model (4) is adequate to the experimental data, provided that all parameters are fixed except for the per-mode mean photon number μ_0 that was calculated from the experimental data. In this paper, we present a method for statistical estimation of multiphoton-subtracted multimode thermal states from both photon and quadrature statistics measurements. In the latter case, since the homodyne detection selects exactly a single mode, corresponding to the local oscillator, $m = 1$.

The paper has the following structure. Section II describes the procedures for preparation and measurement of various states of light of the form (4). Section III describes the procedure for the statistical estimation of the state parameters based on measurements of the photocount statistics. The statistical estimation in Sec. IV is based on the quadrature measurements. We conclude that it is possible to use the model (4) for statistical estimation of the parameters of multimode PSTSs, if prior information is provided.

II. EXPERIMENT

A sketch of our experimental setup is presented in Fig. 2. The optical scheme represents the combination of schemes described in Refs. [20,26]. The HeNe continuous-wave (cw) laser beam is split by a fiber beam splitter (FBS) into two channels. The light from the first output is focused on a rotating ground-glass disk (RGGD), and a part of the scattered light is coupled into a single-mode fiber (SMF) for the single-mode thermal state preparation [29,30]. A small part of the fiber output beam is redirected by a 90:10 beam splitter (BS) to a single-photon detector D_k based on a silicon avalanche photodiode (APD), in order to implement conditional photon annihilation [1,2]. Next the radiation is split by a symmetric BS into two parts. In the first one there is another APD detector D_n for photocount distribution measurement. The rest of the beam is subjected to the homodyne detection (HD). The

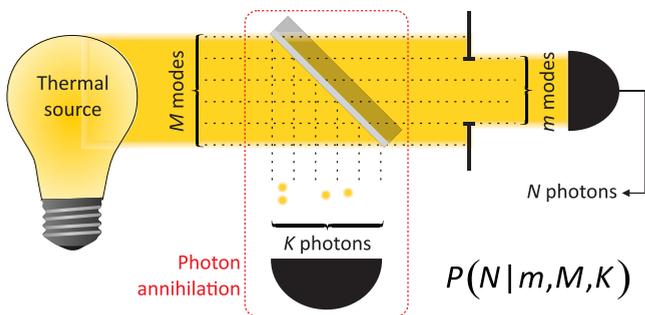


FIG. 1. The registration scheme for photon number statistics of the multiphoton-subtracted multimode thermal state [26].

laser beam from the second output of the FBS serves as a homodyne local oscillator. Since the quadrature distribution of thermal states, as well as MPSTSs, does not depend on the homodyne phase, the phase was not fixed. Thus photocount pulses from D_n and D_k and quadrature values from HD are collected synchronously. This allows us to study the photon statistics registered by the detector D_n and the quadrature statistics obtained by the HD under the condition of a given number of subtracted photons, collected by the detector D_k .

It is important to note that splitting the radiation in half inevitably leads to losses. However, the PSTSs described by the compound Poisson distribution $P_{cP}(N|\mu_0, a)$ under the influence of losses converts to the PSTSs $P_{cP}(N|\mu'_0, a)$ with a lower mean photon number μ'_0 , but with the same a parameter [20,22].

The data-processing algorithm is presented in Fig. 3. First, all the time traces are divided into time bins with the width τ corresponding to the time mode duration [Fig. 3(a)]. The value of τ should satisfy the inequality $T_{\text{coh}} \gg \tau \gg \tau_d$, where T_{coh} is the thermal state coherence time defined by the RGGD velocity and τ_d is the single-photon detector dead time. This inequality defines the possibility of registering several photocounts from a single optical mode (see Ref. [20] for details). In our experiment, $T_{\text{coh}} = 40 \mu\text{s}$, $\tau_d = 220 \text{ ns}$, and $\tau = 10 \mu\text{s}$, so the inequality is satisfied and we were able to register up to 45 photons in each time bin. For each bin the photocount numbers k and n from the detectors D_k and D_n , respectively,

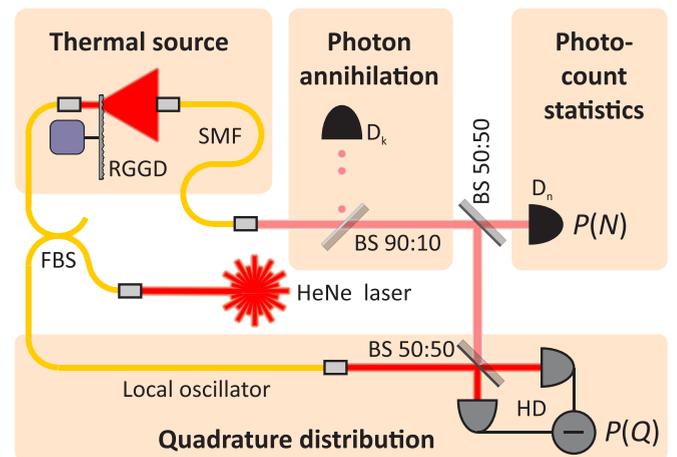


FIG. 2. The experimental setup. BS, beam splitter; D_k and D_n , single-photon APD-based detectors used for photon annihilation and photocount statistics measurements, respectively; FBS, fiber-based beam splitter; HD, homodyne detector, used for quadrature distribution registration; RGGD, rotating ground-glass disk; SMF, single-mode fiber.

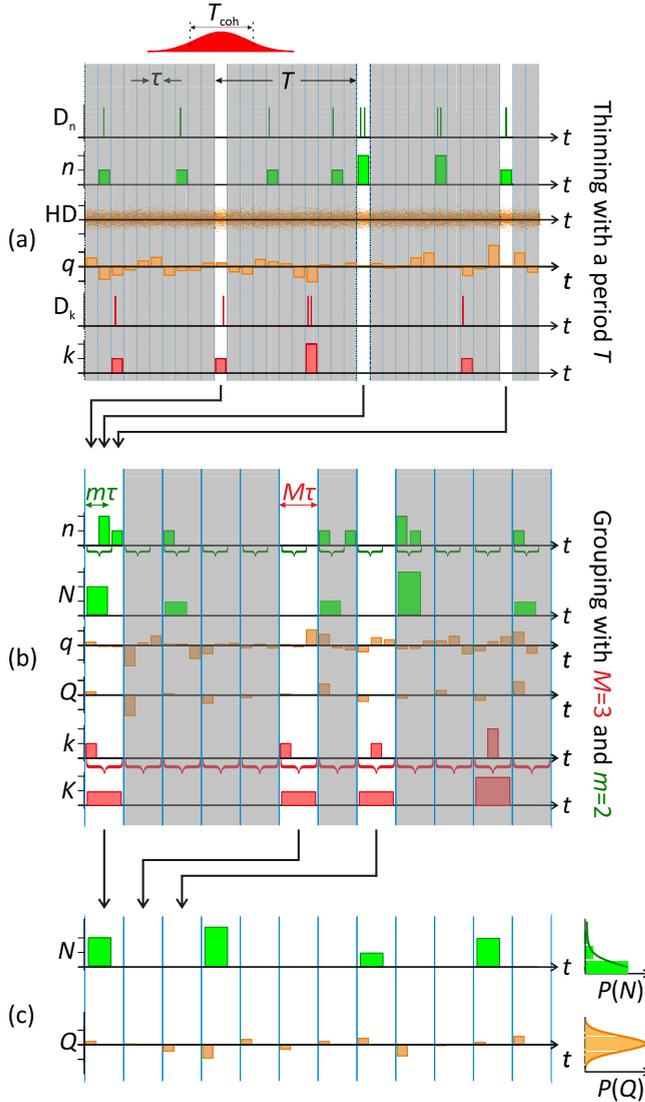


FIG. 3. Signal processing. (a) Initial data set is divided into time bins τ , and then the data are thinned with a period T in order to avoid interbin correlations. (b) Thinned data are grouped by M , and groups are separated according to the total number of subtracted photons in the group (K). The total photon number N is calculated as a sum of the first $m \leq M$ bins. For a group quadrature value Q the first bin value q is selected. (c) The data sets $\{N_1, N_2, \dots\}$ and $\{Q_1, Q_2, \dots\}$ corresponding to the same value of K are collected and subjected to the statistical estimation procedures.

and the quadrature values q from HD are calculated. Next, in order to avoid any interbin correlations, we selected the bins periodically separated by $T = 12T_{\text{coh}}$. Such a large interval T is necessary since the thermal field has a Gaussian correlation function and, even at times significantly longer than the coherence time T_{coh} , correlations are partially preserved, which distorts the photon number and quadrature statistics. Thus only 2% of the collected data are used.

In contrast to the situation considered in Fig. 1, where various spatial modes of the thermal field were considered, in our experiment the field is spatially single mode, so one can select a multimode state by collecting M time modes. Therefore

all the uncorrelated time bins are grouped by M [Fig. 3(b)]. For each group we obtain the total number of subtracted photons K . In order to realize the situation described in Fig. 1, where just a part of the thermal modes is finally collected, we calculate the total photon number N as a sum of the first m bins in a group. Since homodyne detection can select only a single mode, we take only the first bin quadrature value q in a group as a group quadrature value Q . This value corresponds to the single-mode quadrature value under the condition that K photons have been subtracted from the corresponding group of M time modes.

To extract the K -photon-subtracted state, we select the groups with the total number of annihilated photons equal to K [Fig. 3(c)]. Thus, for each value of $M = 1 \div 5$, $m = 1 \div M$, and $K = 0 \div 5$ we derive a set of photocount values $\mathcal{D} = \{N_1, N_2, \dots\}$ and quadrature values $\mathcal{D}_Q = \{Q_1, Q_2, \dots\}$. These data sets are subsequently processed to reconstruct the state parameters using the distribution models (4) and (13), respectively.

Thus we are able to extract the data for an arbitrary state of light with photocount distribution (4) in a wide range of well-controlled parameters m , M , and K . However, the setup does not allow us to control the per-mode mean photon number μ_0 , so we estimate its theoretical value from the data using Eq. [26]:

$$\mu_0 = \frac{\mu}{m(1 + \frac{K}{M})}, \quad (5)$$

where μ is the estimated mean photon number in all registered modes.

Note that $P(N)$ does not exactly correspond to the experimental photocount distribution because of the presence of the dark counts, described by the Poisson distribution $P_{\text{DC}}(N)$ with the mean value $\mu_{\text{DC}} = m \times 0.0015$. Despite the fact that the average number of noise photocounts is much less than the average number of photons per mode (about $\mu_0 = 0.27$ in the examples below), we take it into account to increase the reconstruction accuracy. The resulting photocount distribution is the convolution of (4) and $P_{\text{DC}}(N)$.

III. PARAMETER ESTIMATION BASED ON PHOTOCOUNT STATISTICS

In this section we examine in detail an example based on the experiment with theoretical parameter values $m_t = 2$, $M_t = 3$, $K_t = 3$. The total number of observed events was $n = 58\,623$. The calculated theoretical value of the per-mode mean photon number was $\mu_{0,t} = 0.264$. We denote the number of N -photocount events in the sample \mathcal{D} as $D(N)$. The corresponding histogram and the probability distribution based on theoretical values are shown in Fig. 4(a).

First, we develop a parameter estimation procedure using simulated data. Then we apply it to process the real experimental data.

A. Multicollinearity

Consider the fiducial distribution of parameters $P_F(\mu_0, m, M, K|\mathcal{D})$. One can interpret this distribution as the degree of confidence that a certain set of parameters

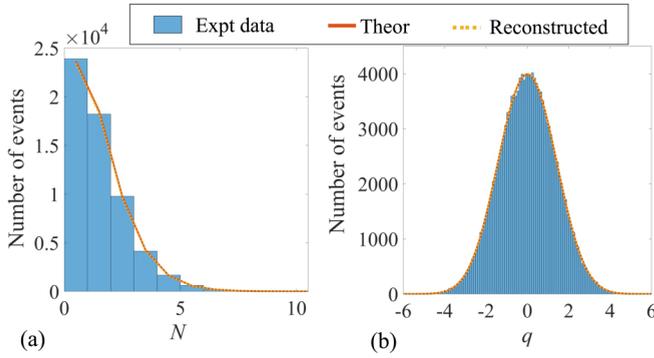


FIG. 4. Experimental data (histograms) compared with probability distributions with theoretical (solid curves) and reconstructed (dashed curves) parameter values for photocount (a) and quadrature (b) statistics.

$\{\mu_0, m, M, K\}$ conditions the data set \mathcal{D} . The distribution is equal to the likelihood function L up to the normalization constant C [31–33]:

$$L(\mu_0, m, M, K|\mathcal{D}) = \prod_{N=0,1,\dots} [P(N|\mu_0, m, M, K)]^{D(N)}. \quad (6)$$

For the sample size $n = 58\,623$, the width of the marginal fiducial distributions over any parameter is quite large. The width of the marginal distribution over K is especially large estimating hundreds of units. This is due to strong correlations between the parameters, or multicollinearity of the initial distribution (4). Figure 5(a) illustrates this effect well.

Note that the likelihood function takes very small values for a high sample size, so we consider its logarithm:

$$\begin{aligned} \ln L(\mu_0, m, M, K|\mathcal{D}) &= \sum_{N=0,1,\dots} D(N) \ln P(N|\mu_0, m, M, K). \end{aligned} \quad (7)$$

Since adding a constant to a given function only affects the proportionality constant C , it is efficient to calculate the fiducial distribution relative to the shifted logarithmic likelihood:

$$\begin{aligned} P_F(\mu_0, m, M, K|\mathcal{D}) &= C' \\ &\cdot \exp[\ln L(\mu_0, m, M, K|\mathcal{D}) \\ &- \max_{\mu_0, m, M, K} \ln L(\mu_0, m, M, K|\mathcal{D})]. \end{aligned} \quad (8)$$

As a result of this shift, the exponent values range from 0 to 1. The constant C' is then calculated by the direct integration.

To numerically characterize multicollinearity, one can calculate the Fisher information matrix $I_{u,v} = n\mathbb{E}_N[(\partial_u P)(\partial_v P)]$, where $\partial_u P$ is the partial derivative of the distribution with respect to the parameter u , and $u, v = m, M, \mu_0$. Here, we assume the parameter K to be fixed. According to the Cramér-Rao bound, the covariance matrix for the estimates of the distribution parameters is bounded by the reciprocal of the Fisher information I^{-1} [33]. Thus the condition number of the Fisher information matrix (the ratio between its maximum and minimum eigenvalues) reflects the robustness of statistical estimates with respect to statistical fluctuations. For

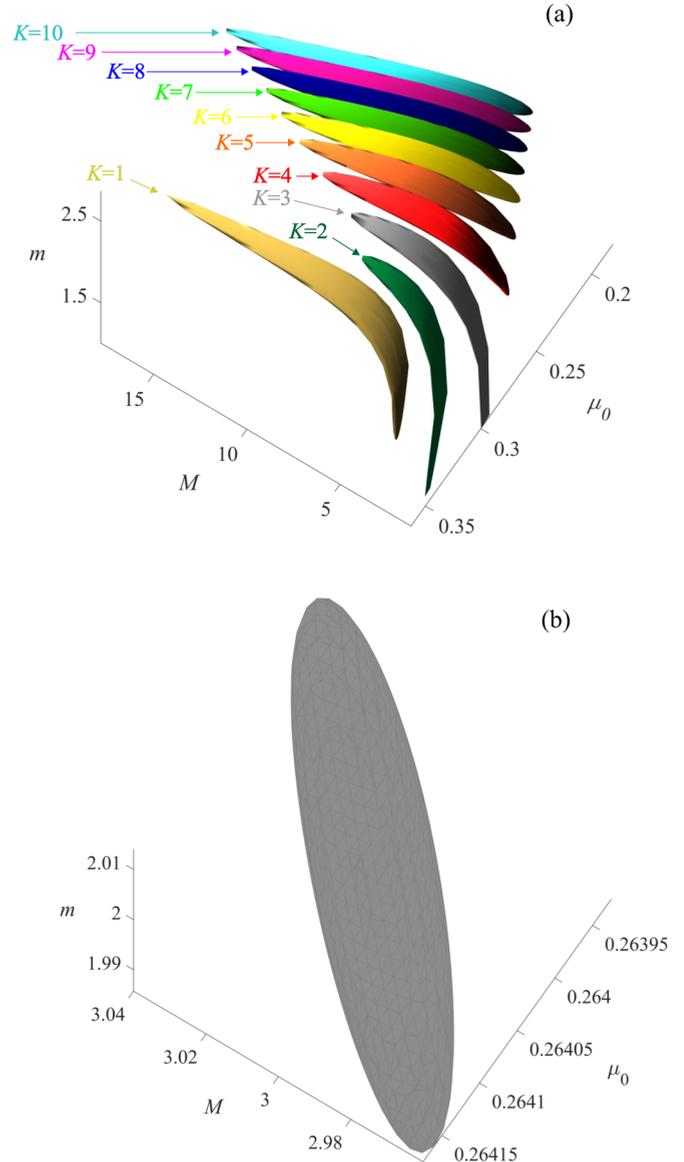


FIG. 5. Isosurfaces of the fiducial distribution $P_F(\mu_0, m, M, K|\mathcal{D})$ at the half-maximum level for the fixed values of K . The data \mathcal{D} were obtained using Monte Carlo simulation with sample size n and distribution parameters $\mu_{0,t} = 0.264$, $m_t = 2$, $M_t = 3$, and $K_t = 3$. (a) $K = 1 \div 10$, $n = 58\,623$. (b) $K = 3$, $n = 420 \times 10^6$.

all practically important parameter values considered in this paper, the information matrix turns out to be ill conditioned. In particular, for the above case the condition number is about 7 000 000. This results in a very low accuracy of statistical estimates.

Note that the inverse Fisher information matrix gives the estimate variances only for a fixed value of K . Therefore we characterize the parameter estimation accuracy by the maximum relative error $\Delta = \max_u(\sigma_u/u_t)$ ($u = m, M, \mu_0, K$) to take fluctuations of K into account. Here, σ_u is the marginal standard deviation of the fiducial distribution $P_F(\mu_0, m, M, K|\mathcal{D})$, and u_t is the parameter theoretical value.

The multicollinearity significantly complicates the procedure for the reconstruction of state parameters. To obtain

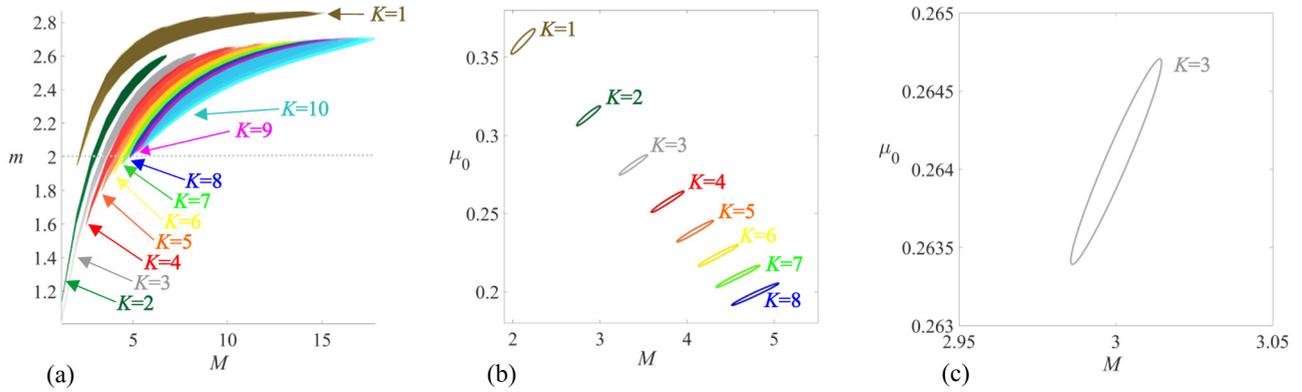


FIG. 6. (a) Projection of the fiducial distribution isosurfaces [Fig. 5(a)] onto the plane $\{M, m\}$. The horizontal line corresponds to the plane with $m = m_t$. (b) and (c) Cross sections of the fiducial distribution isosurfaces at $m = m_t$ for $n = 58\,623$ (b) and $n = 4 \times 10^6$ (c). The latter case provides a single plausible value of K ($K = 3$) and $\Delta = 1\%$ relative error.

a sufficient reconstruction accuracy, a very large amount of data is required (which is difficult to implement for high K values, since they correspond to relatively rare events). For example, numerical experiments show that one needs a sample size of at least $n = 420 \times 10^6$ in order to achieve $\Delta = 1\%$ precision [Fig. 5(b)]. To achieve $\Delta = 10\%$, one needs at least $n = 18 \times 10^6$, which is still a large amount of data.

B. Prior information

Introducing some prior information could, in principle, increase the estimation accuracy. A common choice is to fix the value of some parameter (or a set of parameters). In particular, one can control the number of selected modes m . For example, with homodyne detection, only a single mode of light is selected ($m_t = 1$). In this case, the fiducial distribution takes the form $P_F^m(\mu_0, M, K|\mathcal{D}) = C_m L(\mu_0, m = m_t, M, K|\mathcal{D})$. Graphically, this corresponds to the plot cross section at $m = m_t$ [horizontal line in Fig. 6(a)] that intersects the distribution isosurfaces for $K = 1 \div 8$ only. Thus fixing m reduces the number of plausible values of K from hundreds to of the order of ten. However, corresponding cross sections also show strong parameter correlations [Fig. 6(b)] and multicollinearity.

Note that to achieve $\Delta = 10\%$, one needs a sample size of at least $n = 1.2 \times 10^6$. For $\Delta = 1\%$ [Fig. 6(c)] the sample size $n = 4 \times 10^6$ is required. This is still quite a large amount of data, since the cases of high numbers of subtracted photons are less frequent and require more time to gather experimental data.

C. Bayesian inference

Fixing certain parameters of the initial distribution can significantly improve the reconstruction accuracy. This, however, can introduce systematic errors of reconstruction, if the selected prior values differ significantly from the true values.

Another approach of using prior information is based on Bayes' theorem [28,34]:

$$P_B(\mu_0, m, M, K|\mathcal{D}) = C_B L(\mu_0, m, M, K|\mathcal{D}) \cdot P_P(\mu_0, m, M, K). \quad (9)$$

Here, $P_P(\mu_0, m, M, K)$ is the prior probability distribution of plausible parameter values. The posterior distribution $P_B(\mu_0, m, M, K|\mathcal{D})$ updates the prior information, taking into account the statistical data \mathcal{D} obtained in the experiment.

We rely on the common choice of a multiparameter prior distribution, where all parameters are independent:

$$P_P(\mu_0, m, M, K) = P_P^{\mu_0}(\mu_0) \cdot P_P^m(m) \cdot P_P^M(M) \cdot P_P^K(K). \quad (10)$$

To get single-parameter prior distributions, we consider the conditional distributions: All parameters, except one, are fixed and equal to the expected theoretical values. Furthermore, we will demonstrate that such conditional distributions adequately describe our experimental data.

D. Conditional distribution verification

Consider the conditional distribution for the parameter m . Let us construct [Fig. 7(a)] two fiducial distributions: $P_F^m(m|\mathcal{D})$ and $P_F^m(m|\mathcal{D}_t)$, where $P_F^m(m|\mathcal{D}) = P_F(\mu_{0,t}, m, M_t, K_t)$ and “data” \mathcal{D}_t correspond to the theoretical grouped data $D_t(N) = nP(N|\mu_{0,t}, m_t, M_t, K_t)$. Figure 7(a) shows a strong overlap between these distributions, which suggests that the conditional distribution can be used to describe the data \mathcal{D} . A similar result for parameters M , μ_0 , and K is shown in Figs. 7(b), 7(c), and 7(d), respectively. Note that, as follows from Fig. 7(d), the parameter K is in fact deterministic, since the fiducial probability of $K \neq K_t$ is almost zero. In this regard, below we consider $P_P^K(K) = \delta_{K,K_t}$.

E. Prior distributions

Considering separately the conditional distributions introduced above, it is possible to perform the parameter reconstruction by the maximum likelihood estimation (MLE) technique. Let us take the parameter m as an example. We denote its MLE value as \hat{m}_c . Hereinafter, the subscript c stands for the estimates based on conditional distributions.

According to the general estimation theory, in the limit of a high sample size n , the MLE value is the random variable with normal distribution $f(\hat{m}|m_c, \sigma_{m,c})$ [33]. The expected value m_c corresponds to the asymptotic ($n \rightarrow \infty$) MLE

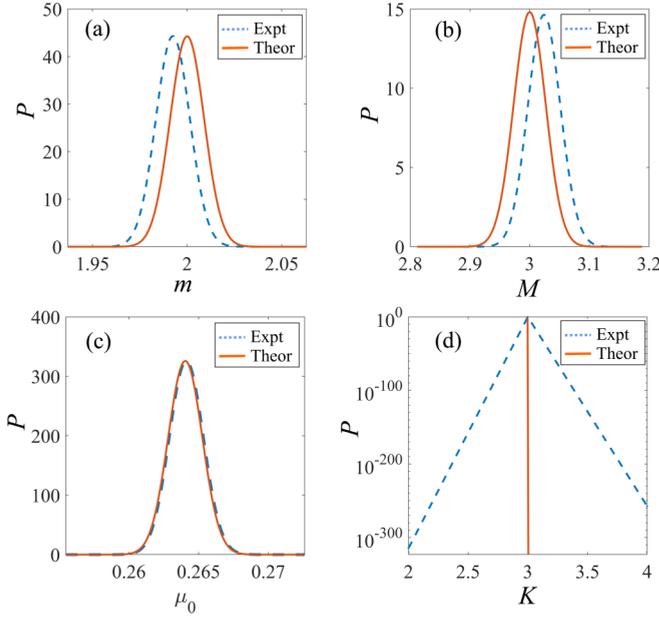


FIG. 7. Sample (experimental) and exact (theoretical) conditional fiducial distributions of the parameters m (a), M (b), μ_0 (c), and K (d) based on experimental (dashed lines) and theoretical (solid lines) data.

estimate, and the variance is related to the single-parameter Fisher information: $\sigma_{m,c}^2 = I_{mm}^{-1}$. Again, we refer to the fiducial inference denoting $P_P^m(m) = f(m|\hat{m}_c, \sigma_{m,c})$ as the prior distribution of the parameter m . The prior distributions $P_P^M(M)$ and $P_P^{\mu_0}(\mu_0)$ are calculated in a similar way.

Following the above technique, we estimated the prior distribution parameters for the real experimental data \mathcal{D} : $\hat{m}_c = 1.993$, $\sigma_{m,c} = 0.009$, $\hat{M}_c = 3.026$, $\sigma_{M,c} = 0.027$, $\hat{\mu}_{0,c} = 0.265$, and $\sigma_{\mu_{0,c}} = 0.001$. One can observe a close relation between the MLE values and theoretical values. This again gives evidence that the theoretical values used to build conditional distributions do not introduce an observable estimator bias.

F. Posterior distribution

Above, we obtained single-parameter prior distributions of all parameters under consideration. Their product forms the multiparameter prior distribution (10). The posterior distribution is the product of the multiparameter distribution (10) and the unconditional fiducial distribution (normalized likelihood function). Prior and posterior distribution are illustrated in Fig. 8.

We calculate the Bayesian posterior distribution expected values to get the point estimates. The obtained values were close to the theoretical ones: $\hat{m}_B = 1.943$, $\hat{M}_B = 3.084$, and $\hat{\mu}_{0,B} = 0.274$. Figure 4(a) shows (dashed curve) the distribution (4) with these parameter values. The resulting curve is in close agreement with the curve corresponding to the theoretical parameters.

Thus our approach implies using predetermined theoretical values of parameters as the starting point for the prior distribution definition. These values are then clarified by means of Bayes' theorem taking statistical data into account.

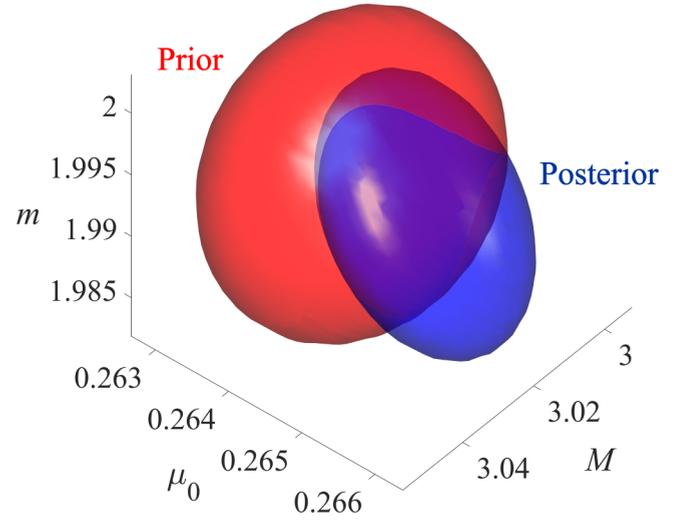


FIG. 8. Half-maximum level isosurfaces of the prior $P_P(\mu_0, m, M, K)$ and posterior $P_F(\mu_0, m, M, K|\mathcal{D})$ distributions of plausible parameter values.

To show that this approach avoids the multicollinearity problem, we construct an information matrix corresponding to the posterior distribution. First of all, note that the Fisher information of the normal distribution $f(\hat{m}|m_c, \sigma_{m,c})$ is $1/\sigma_{m,c}^2 = I_{mm}$ (similarly for the other two parameters). Since the Fisher information matrix of the product of probability distributions is the sum of the corresponding information matrices for each distribution separately, we finally get

$$I_B = I + I_P, \tag{11}$$

where

$$I_P = \begin{pmatrix} \frac{1}{\sigma_{m,c}^2} & 0 & 0 \\ 0 & \frac{1}{\sigma_{M,c}^2} & 0 \\ 0 & 0 & \frac{1}{\sigma_{\mu_{0,c}}^2} \end{pmatrix} = \begin{pmatrix} I_{mm} & 0 & 0 \\ 0 & I_{MM} & 0 \\ 0 & 0 & I_{\mu_0\mu_0} \end{pmatrix}. \tag{12}$$

In fact, the use of the prior information in the form of the product of normal fiducial distributions of MLEs doubles the diagonal of the information matrix of the original unconditional distribution. In our case, the condition number of the resulting matrix is 750. This value is significantly lower than the condition number 7 000 000 for unconditional distribution. Hence the resulting covariance matrix gives low parameter estimator variances. Even for the sample of low size, which was available in our experiment, one could obtain an error rate below $\Delta = 1\%$.

Table I shows the numerical characteristics obtained for the simulated photocount statistics. We derive the sample size required to achieve the error rates $\Delta = 1\%$ and $\Delta = 10\%$.

TABLE I. The sample size values required to achieve error rates below 1 and 10% for different reconstruction methods.

| | No prior information | Fixed m | Bayesian inference |
|-----------------|----------------------|-------------------|--------------------|
| $\Delta = 10\%$ | 18×10^6 | 1.2×10^6 | 8×10^2 |
| $\Delta = 1\%$ | 42×10^7 | 4×10^6 | 5.8×10^4 |

We compare methods that differ in types of prior information: the absence of any prior information, a known fixed value of m , and the approximate knowledge of theoretical parameter values. Table I clearly shows that enormous amounts of data are required to defeat multicollinearity, while the Bayesian method allows one to get an error rate below $\Delta = 1\%$ with the amount of data accumulated in the real experiment.

Using the Bayesian inference, we estimated parameters of 90 different states with $M = 1 \div 5$, $m = 1 \div M$, and $K = 0 \div 5$. The estimation error rate was from 0.0008 to 0.03% for μ_0 , from 0.002 to 1.09% for m , and from 0.015 to 1.89% for M .

IV. PARAMETER ESTIMATION BASED ON QUADRATURE MEASUREMENTS

A homodyne detector selects only a single-mode subsystem of the state ($m = 1$). Let us consider the measurement results of the state with $M_t = 5$, $K_t = 4$. The data \mathcal{D}_Q size was $n = 138\,710$ [Fig. 4(b)]. Using (5) with $\mu = \sigma_q^2 - 1/2$, where σ_q^2 is the sample quadrature variance, we estimate $\mu_{0,t} = 0.752$.

The transition from the photocount distribution to the quadrature distribution of the electromagnetic field is carried out as follows [22,35,36]:

$$\tilde{P}(Q|\mu_0, M, K) = \sum_N P(N|\mu_0, m = 1, M, K) |\varphi_N(Q)|^2, \quad (13)$$

where $\varphi_N(Q)$ are the eigenfunctions of a harmonic oscillator. These functions have the form of the Chebyshev-Hermite basis. The explicit form of these functions is as follows:

$$\varphi_N(Q) = \frac{1}{(2^N N! \sqrt{\pi})^{1/2}} H_N(Q) \exp\left(-\frac{Q^2}{2}\right), \quad (14)$$

where $H_N(Q)$ is the N th Hermite polynomial.

We use the same technique as in the previous section to estimate the parameters of (4) from \mathcal{D}_Q [Fig. 4(b)]. To begin with, we construct a fiducial distribution $\tilde{P}_F(\mu_0, M, K|\mathcal{D}_Q) = C \cdot L(\mu_0, M, K|\mathcal{D}_Q)$, where the likelihood function has the form $L(\mu_0, M, K|\mathcal{D}_Q) = \prod_i \tilde{P}(Q_i|\mu_0, M, K)$.

We again obtain a deterministic value $K = K_t = 4$ and use theoretical values of parameters to get the prior distribution parameters in the same way as was done in Sec. III E: $\hat{\mu}_{0,c} = 0.749$, $\sigma_{\mu_{0,c}} = 0.006$, $\hat{M}_c = 5.064$, and $\sigma_{M,c} = 0.096$. The posterior distribution is shown in Fig. 9. Its expected values are $\hat{M}_B = 5.036$ and $\hat{\mu}_{0,B} = 0.758$.

The resulting quadrature distribution is presented in Fig. 4(b) together with the experimental quadrature histogram and the distribution based on theoretical parameter values. Again, the reconstructed distribution is in close agreement with the expected theoretical one.

Using the Bayesian inference, we measured 30 different states with parameters $M = 1 \div 5$, $m = 1$, and $K = 0 \div 5$. The estimation error rates were from 0.0018 to 0.287% for μ_0 and from 0.044 to 0.287% for M .

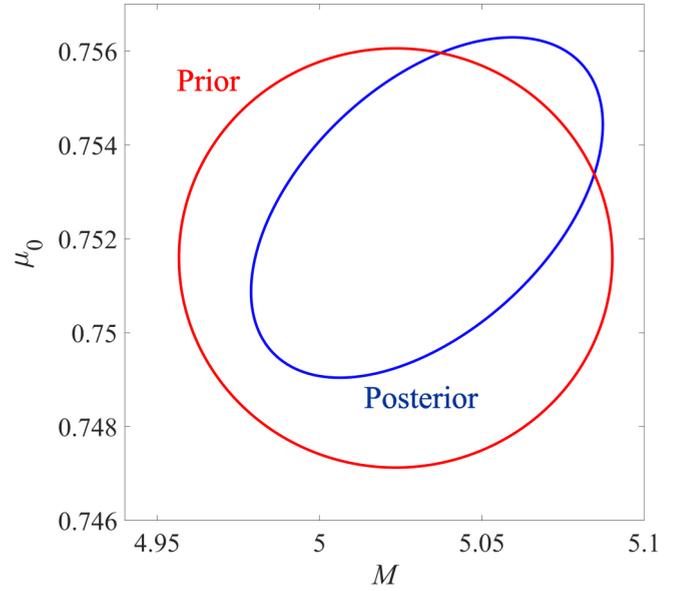


FIG. 9. Half-maximum level isosurfaces of the prior and posterior distributions of plausible parameter values.

V. CONCLUSION

We have considered the problem of statistical parameter estimation of multimode multiphoton-subtracted thermal states of light by analyzing the photon statistics as well as the quadrature distribution. We have studied the subsystems containing only a part of the light modes. For statistical estimation we have used a model of photocount distribution (4), introduced in Ref. [26]. We have shown that the distribution parameters suffer from significant multicollinearity, which complicates the model's use for unconditional parameter estimation. This task can be simplified if we are able to fix one or several parameters at their true values. However, for a more accurate estimate it is better to use variates for the prior knowledge. On the basis of the Bayesian approach, it is possible to accurately estimate the photocount distribution parameters. In particular, we were able to reconstruct all the parameters with an error rate below 1% for a sample of size $n = 5.8 \times 10^4$.

Thus we have developed an approach for statistical parameter estimation of multimode states of thermal light with the subtraction of a given number of photons. It can be used to test photon subtraction in multimode states of light. In a similar way, one can solve the problem of characterization of more complex quantum states of light, which have great potential in the quantum computing field. On the other hand, the developed model can also be used to describe single-mode photon subtraction in the case when there is an error in the selection of exactly a single mode.

ACKNOWLEDGMENTS

This research was performed according to the Development Program of the Interdisciplinary Scientific and Educational School of Lomonosov Moscow State University

“Photonic and quantum technologies: Digital medicine.” The work was supported by the Russian Science Foundation (RSF), Project No. 19-72-10069. G.V.A. expresses special

gratitude to the Russian Foundation for Basic Research (RFBR) for support from the project “Ph.D. students,” No. 19-32-90212.

- [1] A. Ourjoumtsev, *Science* **312**, 83 (2006).
- [2] J. S. Neergaard-Nielsen, B. M. Nielsen, C. Hettich, K. Mølmer, and E. S. Polzik, *Phys. Rev. Lett.* **97**, 083604 (2006).
- [3] V. Parigi, A. Zavatta, M. Kim, and M. Bellini, *Science* **317**, 1890 (2007).
- [4] A. Zavatta, S. Viciani, and M. Bellini, *Science* **306**, 660 (2004).
- [5] J. Wenger, R. Tualle-Brouiri, and P. Grangier, *Phys. Rev. Lett.* **92**, 153601 (2004).
- [6] G. Y. Xiang, T. C. Ralph, A. P. Lund, N. Walk, and G. J. Pryde, *Nat. Photonics* **4**, 316 (2010).
- [7] L. S. Costanzo, A. S. Coelho, N. Biagi, J. Fiurášek, M. Bellini, and A. Zavatta, *Phys. Rev. Lett.* **119**, 013601 (2017).
- [8] K. G. Katamadze, E. V. Kovlakov, G. V. Avosopiants, and S. P. Kulik, *Opt. Lett.* **44**, 3286 (2019).
- [9] Yu. I. Bogdanov, D. V. Fastovets, B. I. Bantysh, A. Y. Chernyavskii, I. A. Semenikhin, N. A. Bogdanova, K. G. Katamadze, Y. A. Kuznetsov, A. A. Kokin, and V. F. Lukichev, *Quantum Electron.* **48**, 1016 (2018).
- [10] M. D. Vidrighin, O. Dahlsten, M. Barbieri, M. S. Kim, V. Vedral, and I. A. Walmsley, *Phys. Rev. Lett.* **116**, 050401 (2016).
- [11] J. Hloušek, M. Ježek, and R. Filip, *Sci. Rep.* **7**, 13046 (2017).
- [12] C. G. Parazzoli, B. A. Capron, B. Koltenbah, D. Gerwe, P. Idell, J. Dowling, C. Gerry, and R. W. Boyd, in *Conference on Lasers and Electro-Optics* (OSA, Washington, DC, 2016), p. FTu3C.4.
- [13] S. M. Hashemi Rafsanjani, M. Mirhosseini, O. S. Magaña-Loaiza, B. T. Gard, R. Birrittella, B. E. Koltenbah, C. G. Parazzoli, B. A. Capron, C. C. Gerry, J. P. Dowling, and R. W. Boyd, *Optica* **4**, 487 (2017).
- [14] U. L. Andersen, J. S. Neergaard-Nielsen, P. van Loock, and A. Furusawa, *Nat. Phys.* **11**, 713 (2015).
- [15] Y.-S. Ra, A. Dufour, M. Walschaers, C. Jacquard, T. Michel, C. Fabre, and N. Treps, *Nat. Phys.* **16**, 144 (2020).
- [16] Y.-S. Ra, C. Jacquard, A. Dufour, C. Fabre, and N. Treps, *Phys. Rev. X* **7**, 031012 (2017).
- [17] G. S. Agarwal, *Phys. Rev. A* **45**, 1787 (1992).
- [18] A. Allevi, A. Andreoni, M. Bondani, M. G. Genoni, and S. Olivares, *Phys. Rev. A* **82**, 013816 (2010).
- [19] Y. Zhai, F. E. Becerra, B. L. Glebov, J. Wen, A. E. Lita, B. Calkins, T. Gerrits, J. Fan, S. W. Nam, and A. Migdall, *Opt. Lett.* **38**, 2171 (2013).
- [20] Yu. I. Bogdanov, K. G. Katamadze, G. V. Avosopiants, L. V. Belinsky, N. A. Bogdanova, A. A. Kalinkin, and S. P. Kulik, *Phys. Rev. A* **96**, 063803 (2017).
- [21] Yu. I. Bogdanov, N. A. Bogdanova, and V. L. Dshkhunyan, *Russ. Microelectron.* **32**, 51 (2003).
- [22] Yu. I. Bogdanov, N. A. Bogdanova, K. G. Katamadze, G. V. Avosopyants, and V. F. Lukichev, *Optoelectron., Instrum. Data Process.* **52**, 475 (2016).
- [23] K. G. Katamadze, G. V. Avosopiants, B. I. Bantysh, Yu. I. Bogdanov, and S. P. Kulik, in *Proceedings of SPIE International Conference on Micro- and Nano-Electronics 2018*, Vol. 11022, edited by V. F. Lukichev and K. V. Rudenko (SPIE, Bellingham, WA, 2019), p. 110222K.
- [24] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*, 1st ed. (Cambridge University Press, Cambridge, 1995), p. 1194.
- [25] Yu. I. Bogdanov, N. A. Bogdanova, K. G. Katamadze, G. V. Avosopiants, and S. P. Kulik, in *Proceedings of SPIE International Conference on Micro- and Nano-Electronics 2018*, Vol. 11022, edited by V. F. Lukichev and K. V. Rudenko (SPIE, Bellingham, WA, 2019), p. 110222L.
- [26] K. G. Katamadze, G. V. Avosopiants, N. A. Bogdanova, Yu. I. Bogdanov, and S. P. Kulik, *Phys. Rev. A* **101**, 013811 (2020).
- [27] L. Landau and E. Lifshitz, *Statistical Physics. Part 1*, 1st ed. (Pergamon, New York, 1959), p. 484.
- [28] W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed. (Wiley, New York, 1968), p. 524.
- [29] W. Martienssen, *Am. J. Phys.* **32**, 919 (1964).
- [30] F. T. Arecchi, *Phys. Rev. Lett.* **15**, 912 (1965).
- [31] R. A. Fisher, *Ann. Eugen.* **6**, 391 (1935).
- [32] D. R. Cox, *Principles of Statistical Inference*, 1st ed. (Cambridge University Press, Cambridge, 2006), p. 236.
- [33] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics, Vol. 2: Inference and Relationship* (Griffin, London, 1961).
- [34] A. Gelman, J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin, *Bayesian Data Analysis*, 3rd ed. (Chapman and Hall, London, 2013), p. 675.
- [35] U. Leonhardt, *Measuring the Quantum State of Light*, 1st ed. (Cambridge University Press, Cambridge, 1997), p. 208.
- [36] Yu. I. Bogdanov, G. V. Avosopyants, L. V. Belinskii, K. G. Katamadze, S. P. Kulik, and V. F. Lukichev, *J. Exp. Theor. Phys.* **123**, 212 (2016).