Experimental demonstration of negative-valued polarization quasiprobability distribution

K. Yu. Spasibko,[1,2,*] M. V. Chekhova,[1,2,3] and F. Ya. Khalili[3,4]
[1]*Max-Planck-Institute for the Science of Light, Staudtstrasse 2, 91058 Erlangen, Germany*
[2]*University of Erlangen-Nürnberg, Staudtstrasse 7/B2, 91058 Erlangen, Germany*
[3]*Department of Physics, M.V. Lomonosov Moscow State University, Leninskie Gory, 119991 Moscow, Russia*
[4]*Russian Quantum Center, 143025 Skolkovo, Russia*
(Received 9 June 2017; published 9 August 2017)

Polarization quasiprobability distribution defined in the Stokes space shares many important properties with the Wigner function for position and momentum. Most notably, they both give correct one-dimensional marginal probability distributions and therefore represent the natural choice for the probability distributions in classical hidden-variable models. In this context, negativity of the Wigner function is considered as proof of nonclassicality for a quantum state. On the contrary, the polarization quasiprobability distribution demonstrates negativity for all quantum states. This feature comes from the discrete nature of Stokes variables; however, it was not observed in previous experiments, because they were performed with photon-number averaging detectors. Here we reconstruct the polarization quasiprobability distribution of a coherent state with photon-number resolving detectors, which allows us to directly observe for the first time its negativity. Furthermore we derive a theoretical polarization quasiprobability distribution for any linearly polarized quantum state.

DOI: 10.1103/PhysRevA.96.023822

I. INTRODUCTION

Noncommuting observables do not exist in classical physics, but arise in quantum mechanics and optics. They lead to difficulties in attempts to describe quantum states in a semiclassical way, because it is impossible to define a joint probability distribution for such observables. As a remedy, *quasi*probability distributions have been proposed, which can take negative values and therefore violate one of the main axioms of probability theory.

The most well-known example of noncommuting observables is the canonical pair of position and momentum, and the most remarkable corresponding joint quasiprobability distribution is the Wigner distribution [1]. Its major distinctive feature is that, in contrast to, e.g., the Glauber-Sudarshan P representation [2,3] or the Husimi-Kano Q representation [4,5], it gives correct marginal distributions for position and momentum [6]. Therefore, it represents the natural choice for the probability distributions in classical hidden-variable models. Because of this property, it is widely accepted that the negativity of a Wigner distribution means the nonclassicality of the quantum state [6–11].

Mathematical objects with properties similar to those of the Wigner function have been defined for many different systems and observables due to its unique features. In particular, it was done for discrete-valued position and momentum [12], for the Hermite-Gaussian and Laguerre-Gaussian modes of an optical beam [13], and for the canonical pair of the angle and the angular momentum of vortex states [14].

The analog of the Wigner distribution for the three noncommuting Stokes observables [see Eqs. (3)]—the polarization quasiprobability distribution (PQPD)—was developed in Refs. [15,16]. Although PQPD represents only part of the density matrix, the so-called polarization sector [17–20], it remains a powerful tool. PQPD fully describes the polarization properties of a quantum state and gives correct one-dimensional marginal probability distributions for all Stokes observables and their linear combinations.

Note also that the Wigner function quantum tomography requires an additional beam phase-locked with the light under study. It could be either the local oscillator beam in the standard implementations, or an additional coherent beam in the method which involves displacement of the state under investigation and then measurement of the photon number parity [21,22], or a coherent beam in another polarization, see [11], Sec. V. At the same time, polarization tomography does not require it. This is a great advantage in experiments, especially in the experiments with broadband light [23].

Because Stokes observables commute with the photon-number operator, PQPD for a quantum state can be described as a sum of PQPDs for photon-number (Fock) states with certain weights. Therefore, in principle, each subspace with a fixed photon number can be represented separately. This property of PQPD was used to find extremal quantum states [19] or to investigate polarization squeezing in each photon-number subspace [20].

An unusual feature of PQPD is that it takes negative values for all quantum states of light, even for the "most classical" coherent ones. The physical origin of this behavior was explored theoretically in Ref. [11]. The negativity appears because the Stokes observables are discrete valued, that is, their single-dimensional marginal distributions are defined only for the integer values of the arguments. On the other hand, the full PQPD as well as its two-dimensional marginal distributions could possess symmetry features impossible for the discrete argument functions. Their negativity resolves this contradiction. At the same time, this feature was never observed in polarization tomography experiments [15,23–26], because all these experiments were performed with photon-number averaging detectors, which smoothed the measured photon-number statistics and washed out the nonclassical features of PQPD.

*Corresponding author: kirill.spasibko@mpl.mpg.de

Thus, photon-number resolving detection is crucial for observing the intrinsic negativity of PQPD. This is a non-Gaussian operation, and it can be used to prepare states with a negative Wigner function [27,28]. At the same time, PQPD is defined in terms of Stokes observables and therefore photon-number resolving detection is an essential part of its measurement. In other words, negativity is not just an (inconvenient) mathematical feature of PQPD but has a deep physical meaning.

In this work, we have performed polarization tomography for a coherent state of light using single-photon detectors. To this end, we have developed a procedure for the high-quality reconstruction of the PQPD using a limited data set. The reconstructed distribution demonstrates well-pronounced negative-valued areas.

Moreover, we have derived the PQPD for any linearly polarized quantum state. Being in general quite complicated, the PQPD expression can be drastically simplified for many important quantum states.

II. STOKES OBSERVABLES AND PQPD

A quantum state of light can be fully described by its density operator $\hat{\rho}$. For states with two polarization modes the PQPD $W(S_1,S_2,S_3)$ is defined as the Fourier transform of the polarization characteristic function $\chi(u_1,u_2,u_3)$,

$$W(S_1,S_2,S_3) = \int_{-\infty}^{\infty} \chi(u_1,u_2,u_3) \times \exp\left(-i\sum_{j=1}^{3} u_j S_j\right) \frac{du_1\,du_2\,du_3}{(2\pi)^3}, \quad (1)$$

where

$$\chi(u_1,u_2,u_3) = \mathrm{Tr}\left[\hat{\rho}\exp\left(i\sum_{j=1}^{3} u_j \hat{S}_j\right)\right], \quad (2)$$

$u_j \in \mathbb{R}$. The Stokes operators $\hat{S}_j$ are defined as

$$\hat{S}_1 = \hat{n}_H - \hat{n}_V, \quad \hat{S}_2 = \hat{a}_V^\dagger \hat{a}_H + \hat{a}_H^\dagger \hat{a}_V, \quad \hat{S}_3 = i(\hat{a}_V^\dagger \hat{a}_H - \hat{a}_H^\dagger \hat{a}_V), \quad (3)$$

where $\hat{a}_H$ and $\hat{a}_V$ are the photon annihilation operators for the horizontal (H) and vertical (V) polarization modes, and $\hat{n}_{H,V} = \hat{a}_{H,V}^\dagger \hat{a}_{H,V}$ are photon-number operators in these modes. All Stokes operators can be represented as the differences of photon-number operators in certain modes, therefore the corresponding Stokes observables, e.g., S_1, can only take integer values $n \in \mathbb{Z}$.

III. PQPD RECONSTRUCTION

A standard setup for polarization tomography (see Fig. 1) consists of quarter- and a half-wave plates ($\lambda/4$ and $\lambda/2$), a polarizing beam splitter, and two detectors (D_1 and D_2). For each pair of settings of the quarter-wave ($\tilde{\beta}$) and half-wave ($\tilde{\alpha}$) plates, such a setup measures a different arbitrary Stokes operator $\hat{S}_{\alpha\beta} = \hat{n}_1 - \hat{n}_2$. The operators $\hat{n}_{1,2}$ correspond to the photon numbers in the mode transmitted or reflected by the polarizing beam splitter and are measured by D_1 or D_2, respectively.

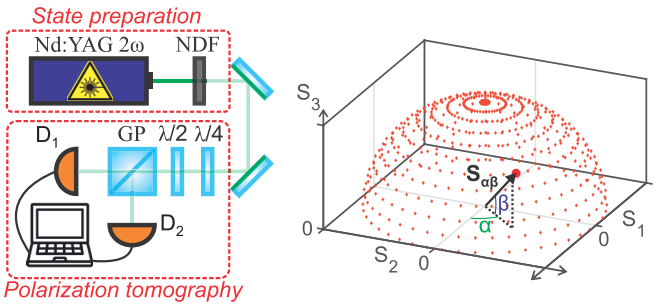

FIG. 1. Left: experimental setup. A weak coherent state is prepared by attenuating the second harmonic of a Nd:YAG laser (Nd:YAG 2ω) with neutral density filters (NDF). A standard setup for polarization tomography consists of a quarter- and a half-wave plate ($\lambda/4$ and $\lambda/2$), a polarizing beam splitter, and two detectors (D_1 and D_2). We use a Glan-Taylor prism (GP) as a polarizing beam splitter and two avalanche photodiodes as detectors. Right: points at which tomographic measurements are performed are shown on the Poincaré sphere.

The angles $\alpha \in [0,2\pi]$ and $\beta \in [-\pi/2,\pi/2]$ that define a point on the Poincaré sphere (see Fig. 1) are determined by the settings of the wave plates,

$$\alpha = 4\tilde{\alpha} - 2\tilde{\beta}, \quad \beta = 2\tilde{\beta}. \quad (4)$$

An arbitrary Stokes operator $\hat{S}_{\alpha\beta}$ can be represented in Cartesian coordinates $(\hat{S}_1,\hat{S}_2,\hat{S}_3)$ as

$$\hat{S}_{\alpha\beta} = (\hat{S}_1\cos\alpha + \hat{S}_2\sin\alpha)\cos\beta + \hat{S}_3\sin\beta. \quad (5)$$

It is clear that this operator possesses inversion symmetry $\hat{S}_{(\alpha+\pi)(-\beta)} = -\hat{S}_{\alpha\beta}$, thus measurements only on half of the Poincaré sphere suffice for the full reconstruction of any state.

In the experiment, for each point on the Poincaré sphere (for each α and β), acquisition of many $S_{\alpha\beta}$ values is needed. From these values we calculate the probabilities $W_{\alpha\beta}(n)$ that $S_{\alpha\beta}$ are equal to n. From these probabilities we restore the polarization characteristic function $\chi_{\alpha\beta}(\lambda)$ in spherical coordinates (λ,α,β) [11]:

$$\chi_{\alpha\beta}(\lambda) = \sum_{n=-\infty}^{\infty} W_{\alpha\beta}(n)e^{i\lambda n}, \quad \lambda \in [0,\infty). \quad (6)$$

These spherical coordinates (λ,α,β) are related to the Cartesian ones (u_1,u_2,u_3) by the following transformations:

$$u_1 = \lambda\cos\alpha\cos\beta, \quad u_2 = \lambda\sin\alpha\cos\beta, \quad u_3 = \lambda\sin\beta. \quad (7)$$

Thus, using these transformations, Eq. (1) can be rewritten as

$$W(S_1,S_2,S_3) = -\frac{1}{(2\pi)^2}\int_0^{2\pi} d\alpha \int_0^{\pi/2} d\beta\cos\beta \times \sum_{n=-\infty}^{\infty} W_{\alpha\beta}(n)\delta^{(2)}(S_{\alpha\beta}-n), \quad (8)$$

where $\delta^{(2)}(x)$ is the second derivative of the Dirac delta function. Here we exploit the symmetry of $\hat{S}_{\alpha\beta}$ and perform

integration over the radial coordinate λ. As a result, we obtain the equation for reconstructing the PQPD $W(S_1,S_2,S_3)$ from the experimentally measured probabilities $W_{\alpha\beta}(n)$.

The reconstruction of PQPD $W_\epsilon(S_1,S_2,S_3)$ from the experimentally acquired data set [Eq. (8)] requires some approximation $\delta_\epsilon(x)$ for the Dirac delta function $\delta(x)$. Here ϵ is the smoothing parameter. We choose the Gaussian approximation,

$$\delta_\epsilon(x) = \frac{1}{2\epsilon\sqrt{\pi}} e^{-x^2/4\epsilon^2}, \tag{9}$$

and similarly for the derivatives of $\delta(x)$. The value of the smoothing parameter ϵ should be chosen properly depending on the reconstructed state, the acquired data set, and the interpolation method (see below). On the one hand, it has to be small enough to represent all features of the PQPD (a large ϵ masks them), but on the other hand, small values of ϵ lead to a lot of artifacts in the reconstructed distribution (the so-called reconstruction noise).

IV. EXPERIMENT AND DATA PROCESSING

We have performed the polarization tomography of a horizontally polarized weak coherent state $|\gamma\rangle$. This state was produced by strongly attenuating a coherent beam at the wavelength 532 nm generated by a pulsed Nd:YAG laser (Nd:YAG 2ω) with 10-ns pulse duration and 10-kHz repetition rate (see Fig. 1). The laser power stability was about 2%.

Attenuation (or any other linear losses) does not change the statistical properties of a coherent state: the state remains coherent, but the mean number of photons $|\gamma|^2$ is reduced.

The attenuation to a single-photon level was performed by a neutral density filter (NDF). The resulting probability of a single-photon detection event $p_1 \approx |\gamma|^2$ was equal to 0.189. In this case, p_1 was at least one order of magnitude higher than the probabilities of two-photon and higher-order detection events. Therefore we ignored such events and considered only single-photon and no-photon detection events (with the probability p_0).

We used avalanche photodiodes as single-photon detectors (D_1 and D_2). The photodiodes were gated electronically synchronously with the laser pulses. The gating led to considerable reduction of dark count rate of the detectors, so the latter was only 0.1% from the total mean count rate. Nevertheless it was subtracted from the measured count rate. The data for each tomographic measurement were acquired for 120 s.

The points (α_k,β_l) on the Poincaré sphere where tomographic measurements were performed cover the upper hemisphere ($\beta \geqslant 0$) with a step of 8° (see Fig. 1). These points have been accessed by different combinations of the settings for the quarter- and half-wave plates with the steps equal to 4° and 2°, respectively (and for $\tilde{\beta} = 45°$, the "north" pole of the Poincaré sphere was accessed). For each point from this discrete set we calculated the experimental probabilities $\tilde{W}_{\alpha_k\beta_l}(n)$, where $n = \{-1,0,1\}$.

The full experimental data set $\tilde{W}_{\alpha_k\beta_l}(n)$ is not suitable for the final integration over α and β in Eq. (8), because it is defined on a discrete set $\{\alpha_k,\beta_l\}$. Thus it should be interpolated by a continuous function. The interpolated function $W_{\alpha\beta}(n)$ is given by the convolution sum of the data points $\tilde{W}_{\alpha_k\beta_l}(n)$ with the interpolation kernel $u(\alpha,\beta)$,

$$W_{\alpha\beta}(n) = \sum_{\alpha_k,\beta_l} \tilde{W}_{\alpha_k\beta_l}(n) u(\alpha-\alpha_k,\beta-\beta_l). \tag{10}$$

Various interpolation kernels can be used. The simplest one is a rectangular function $u(\alpha,\beta) = \Pi(\alpha)\Pi(\beta)$, where

$$\Pi(x) = \begin{cases} 1, & |x| < 1/2, \\ 0, & |x| \geqslant 1/2. \end{cases} \tag{11}$$

The integration of such an interpolated function (e.g., as part of the Fourier or Radon transform) gives exactly the same result as when the integration is replaced by the summation. Such a replacement was always used for reconstruction in polarization tomography [23–26]. Unfortunately, with this interpolation, the transformations are accompanied by rather high noise. One can overcome this problem by collecting more experimental points (α_k,β_l) or by using different interpolation kernels. Interpolation methods are well-developed for image resampling [29,30]. It has been shown that several interpolation kernels could suppress the reconstruction noise by more than 30 dB better than the rectangular-function kernel.

In our case, the probabilities $W_{\alpha\beta}(n)$ could not be negative; hence we needed a strictly positive kernel. We chose a positive cubic spline kernel $u(\alpha,\beta) = u(\alpha)u(\beta)$ [29], where

$$u(x) = \begin{cases} 2|x|^3 - 3|x|^2 + 1, & |x| \leqslant 1, \\ 0, & |x| > 1. \end{cases} \tag{12}$$

This kernel suppresses the noise very well and is at the same time quite simple. For each interval between the data points, e.g., $[x_k,x_{k+1}]$, the interpolation requires only the experimental data from the endpoints of the interval (x_k and x_{k+1}). Hence this kernel has the same simplicity as the linear interpolation kernel, but a better performance.

V. RESULTS

Using this interpolation and the approximation (9) with $\epsilon = 0.02$, we have reconstructed the PQPD $W_\epsilon(S_1,S_2,S_3)$. Its cross sections along the (S_2,S_3) plane at different values of S_1 and the one-dimensional (1D) cut along S_2 ($S_1 = 0$ and $S_3 = 0$) are shown in Fig. 2.

In general, each distribution contains a central peak at the origin of the Stokes space ($S \equiv \sqrt{(S_1)^2 + (S_2)^2 + (S_3)^2} = 0$) and a jump from negative values to positive ones at $S = 1$. The central peak, which appears because of the no-photon detection events, is more than two orders of magnitude higher than the jump, which happens because of the single-photon ones. At values $S > 1$ there is only the reconstruction noise [Fig. 2(g)].

The reconstructed distribution $W_\epsilon(S_1,S_2,S_3)$ is in agreement with the theoretical one for our case (see the Appendix for derivation of the full theoretical distribution):

$$W(S,\theta,\phi) = p_0\delta_3(S) + \frac{p_1\cos\theta}{4\pi S^2}\delta(S-1) - \frac{p_1(1+\cos\theta)}{4\pi S}\delta'(S-1), \tag{13}$$

where $\delta_3(S) = \delta(S_1)\delta(S_2)\delta(S_3)$, $\delta'(x)$ is the first derivative of the Dirac delta function. Here we use spherical coordinates

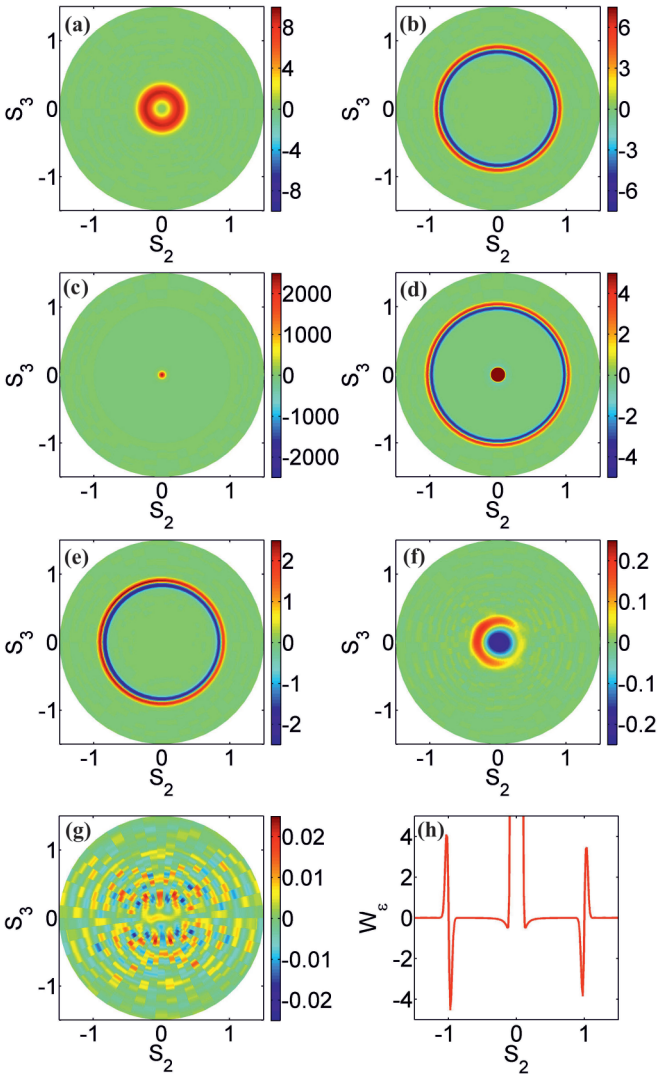

FIG. 2. Cross sections of the reconstructed PQPD $W_\epsilon(S_1,S_2,S_3)$ (with $\epsilon = 0.02$) along the (S_2,S_3) plane at $S_1 = 1$ (a), $S_1 = 0.5$ (b), $S_1 = 0$ (c,d), $S_1 = -0.5$ (e), $S_1 = -1$ (f), and $S_1 = -1.5$ (g). In panel (d), the color code is changed to highlight the jump at $S = 1$. Panel (h) presents the 1D cut of the reconstructed PQPD along S_2 ($S_1 = 0$ and $S_3 = 0$).

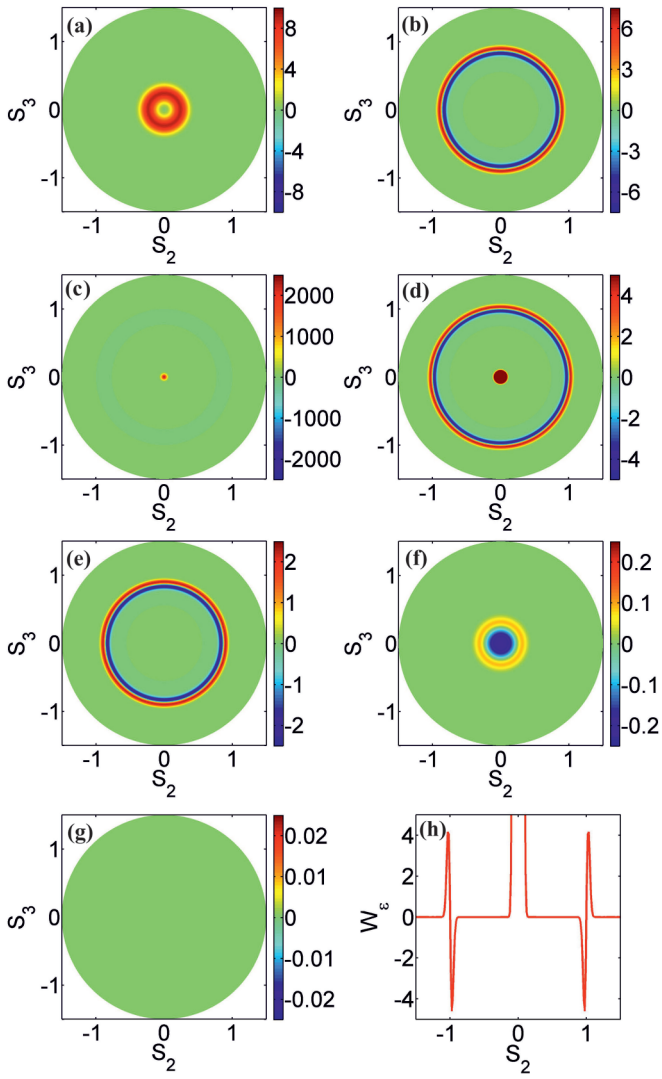

FIG. 3. Same as Fig. 2, but for the theoretical PQPD $W_\epsilon(S_1,S_2,S_3)$ smoothed by $\epsilon = 0.02$.

(S,θ,ϕ):

$$S_1 = S\cos\theta, \quad S_2 = S\sin\theta\cos\phi, \quad S_3 = S\sin\theta\sin\phi. \tag{14}$$

From these formulas we calculated the theoretical PQPD $W_\epsilon(S_1,S_2,S_3)$ for the same probabilities of single-photon ($p_1 = 0.189$) and no-photon detection events ($p_0 = 0.811$) as in the experimental case. We used the same approximation (9) and the same value of the smoothing parameter $\epsilon = 0.02$. The same cross sections are shown for both distributions (Fig. 3). The experimental and theoretical distributions are almost indistinguishable. The only differences are caused by the reconstruction noise [Fig. 2(g)] and imperfections of the half- and quarter-wave plates [Fig. 2(f)]. We want to stress here that the plotted theoretical PQPD is not a simulation of any kind, it is just a smoothed exact distribution [Eq. (13)].

It is clear that the distribution $W_\epsilon(S_1,S_2,S_3)$ possesses a rotation symmetry in the plane (S_2,S_3). Thus it is convenient to use cylindrical coordinates (S_1,S_{23},ϕ), with the radial coordinate $S_{23} \equiv \sqrt{(S_2)^2 + (S_3)^2} = S\sin\theta$, instead of the Cartesian ones (S_1,S_2,S_3). Due to this symmetry, up to experimental imperfections a cross section at some angle ϕ (e.g., $\phi = 0$) demonstrates all features of the PQPD (Fig. 4).

Our experimental reconstruction shows that in realistic cases we are interested only in the main δ' contributions of Eq. (A24). Moreover in most cases p_0 is much higher than p_n for $n \geqslant 2$, therefore [see Eqs. (A9) and (A26)]

$$W(S,\theta,\phi) \approx p_0\delta_3(S) - \frac{1}{2\pi S}\sum_{n=1}^{\infty}\frac{p_n}{2^n}\sum_{k=0}^{n}\binom{n}{k} \times \sum_{\tilde{n}=0}^{\lfloor\frac{n-1}{2}\rfloor} B^n_{k\tilde{n}}\delta'(S-n+2\tilde{n})\cos^k\theta, \tag{15}$$

where $B^n_{k\tilde{n}}$ is a constant value defined in Eq. (A25).

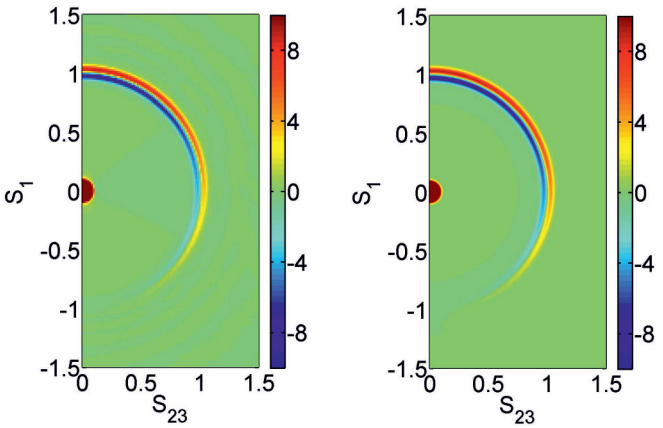

FIG. 4. Cross sections of the experimental (left) and theoretical (right) PQPD $W_\epsilon(S_1,S_{23},\phi)$ (with $\epsilon = 0.02$) at $\phi = 0$. In all figures, the color code is changed to highlight the jump at $S = 1$.

From Eq. (15) it is evident that an n-photon state gives a contribution not only around $S \approx n$, but also around $S \approx n - 2\tilde{n}, 0 \leqslant \tilde{n} \leqslant \lfloor (n-1)/2 \rfloor$. This has a simple physical meaning: n photons split by a beam splitter into $(n - \tilde{n})$ and $\tilde{n}$ photons give $S_{\alpha\beta} = n - 2\tilde{n}$.

This property provides a possibility to infer probability p_n by comparing the experimental and theoretical PQPDs in an experimental setup that does not allow measuring $S_{\alpha\beta} = n$. In the case of significant p_n the experimental PQPD will have a contribution proportional to a noticeable $\cos^n\theta$ factor around $S \approx n - 2\tilde{n}$, $\forall\, \tilde{n} \geqslant 1$, which cannot be produced by contributions from the lower photon-number probabilities $p_{n-2\tilde{n}}$.

Furthermore Eq. (15) can be even more simplified in the case of strongly decaying photon-number distribution, e.g., for coherent or thermal light with low mean number of photons, $p_n \gg p_{n+j}, \{\forall\, n \geqslant 0, j \geqslant 2\}$:

$$W(S,\theta,\phi) \approx p_0\delta_3(S) - \frac{1}{2\pi S}\sum_{n=1}^{\infty} p_n \delta'(S-n)\cos^{2n}(\theta/2). \tag{16}$$

In such form it looks simple and extremely convenient for the comparison with experimental results.

VI. CONCLUSION

We have shown experimentally the full reconstruction of the PQPD from measurements using photon-number resolving detectors. As a result we observed the intrinsic negativity of PQPD originating from the discrete nature of the Stokes observables. For our reconstruction we have elaborated a procedure that leads to a high-quality PQPD from a relatively small data set. Finally, we derived the theoretical PQPD distribution that is valid for any linearly polarized state, which is especially interesting for states with low photon numbers.

The PQPD reconstruction with photon-number resolving detectors is very promising because these detectors can resolve up to tens of photons with more than 90% quantum efficiency [31–35]. These detectors can advance polarization tomography in this direction and make it a useful tool for quantum state characterization.

ACKNOWLEDGMENTS

We acknowledge the financial support of the Russian Foundation for Basic Research (RFBR) Grants No. 14-02-31030 and No. 14-02-00399 and the joint DFG-RFBR project CH1591/2-1-16-52-12031 NNIOa. The work of F.Ya.K. was supported by LIGO NSF Grant No. PHY-1305863.

APPENDIX: CALCULATION OF THE THEORETICAL PQPD

To describe the reconstructed PQPD $W(S_1,S_2,S_3)$ we calculate the theoretical distribution. In the general case of linearly (horizontally) polarized state the density operator is

$$\hat{\rho} = \sum_{n=0}^{\infty} p_n |n\rangle\langle n|_H \otimes |0\rangle\langle 0|_V. \tag{A1}$$

Similar to the case of the experimental data processing, the starting point is the definition of the PQPD and the polarization characteristic function [Eqs. (1) and (2)]. The derivation of the theoretical distribution requires the use of spherical coordinates (λ,ξ,ρ) that differ from (λ,α,β) used in the experimental data processing:

$$u_1 = \lambda\cos\xi, \quad u_2 = \lambda\sin\xi\cos\rho, \quad u_3 = \lambda\sin\xi\sin\rho. \tag{A2}$$

The polarization characteristic function for the density operator (A1) is equal to [11]

$$\chi_{\xi\rho}(\lambda) = \sum_{n=0}^{\infty} p_n(\cos\lambda + i\sin\lambda\cos\xi)^n. \tag{A3}$$

We expand the binomial and replace the sine and cosine by exponentials:

$$\chi_{\xi\rho}(\lambda) = \sum_{n=0}^{\infty}\frac{p_n}{2^n}\sum_{k=0}^{n}\binom{n}{k}\sum_{l=0}^{n-k}\sum_{m=0}^{k}\binom{n-k}{l}\binom{k}{m} \times (-1)^m(\cos\xi)^k e^{i\lambda r}, \tag{A4}$$

where $r = n - 2l - 2m$ and $\binom{n}{k}$ is a binomial coefficient.

In contrast to previous works, e.g., Ref. [16], here we use the characteristic function without an approximation of large n. Therefore the obtained PQPD will be valid for any n.

For the future integration of Eq. (A4) in Eq. (1) we specify three different groups of summation terms. In the first one $r = 0$ and $k = 0$, in the second one $r = 0$ and $k > 0$, and in the third one $r \neq 0$. We denote the corresponding parts of PQPD as W_0, $\bar{W}_0$, and $\bar{W}$. So

$$W(S_1,S_2,S_3) = W_0 + \bar{W}_0 + \bar{W}. \tag{A5}$$

As numerical calculations show, W_0 is a trivial δ contribution around $S = 0$, $\bar{W}_0$ is nontrivial around $S = 0$, and $\bar{W}$ is the main (most interesting) contribution for $S > 0$.

The first group can be directly integrated in Eq. (1):

$$W_0(S_1,S_2,S_3)=\sum_{f=0}^{\infty}\frac{p_{2f}}{2^{2f}}\binom{2f}{f}\delta(S_1)\delta(S_2)\delta(S_3). \quad \text{(A6)}$$

For the second contribution we obtain

$$\bar{W}_0(S_1,S_2,S_3)=\sum_{f=0}^{\infty}\frac{p_{2f}}{2^{2f}}\sum_{g=1}^{f}\binom{2f}{2g}\sum_{\substack{m=\\ \max(0,2g-f)}}^{\min(2g,f)}\times\binom{2f-2g}{f-m}\binom{2g}{m}(-1)^m I_{r=0}^{k=2g}, \quad \text{(A7)}$$

where

$$I_r^k=\int_{-\infty}^{\infty}(\cos\xi)^k e^{i\lambda r}\exp\left(-i\sum_{j=1}^{3}u_j S_j\right)\frac{du_1\,du_2\,du_3}{(2\pi)^3}. \quad \text{(A8)}$$

Here we removed combinations that give zero contribution. Only terms with even n and k give nonzero contribution around $S=0$.

Unfortunately, the integral $I_{r=0}^{k=2g}$ cannot be expressed in any simple analytical form. However, in the cases relevant to polarization tomography with photon-number resolving detectors, this fact does not play a significant role. First, the most interesting is the behavior of PQPD at $S>0$. Second, in these cases quantum states have large probability of no-photon detection events (p_0) that is much bigger than probabilities of two-photon and higher-order detection events (p_n for $n\geqslant 2$). Therefore, the first group of summation terms [Eq. (A6)] can be simplified and the second group does not play a significant role. Thus we obtain for the contributions around $S=0$

$$W_0(S,\theta,\phi)+\bar{W}_0(S,\theta,\phi)\approx p_0\delta_3(S). \quad \text{(A9)}$$

The other terms in Eq. (A4) form the main contribution

$$\bar{W}(S_1,S_2,S_3)=\sum_{n=1}^{\infty}\frac{p_n}{2^n}\sum_{k=0}^{n}\binom{n}{k}\sum_{lm}, \quad \text{(A10)}$$

where

$$\sum_{lm}=\sum_{l=0}^{n-k}\sum_{m=0}^{k}\binom{n-k}{l}\binom{k}{m}(-1)^m I_{r\neq 0}^{k}. \quad \text{(A11)}$$

We can perform integration over λ in Eq. (A8) by recombining the summation members in Eq. (A11) and using the Leibnitz integral rule. So we obtain

$$\sum_{lm}=\sum_{l=0}^{n-k}\sum_{m=0}^{k}\frac{(-1)^{m+1}}{2(2\pi)^2}\binom{n-k}{l}\binom{k}{m}\frac{\partial^2 I_\xi}{(\partial y)^2}\bigg|_{y=r\neq 0}, \quad \text{(A12)}$$

where

$$I_\xi=\int_0^{\pi}d\xi\,\frac{(\cos\xi)^k}{S\sin\theta}I_{\bar{\rho}}, \quad \text{(A13)}$$

$$I_{\bar{\rho}}=\int_0^{2\pi}d\bar{\rho}\,\delta(P-\cos\bar{\rho}), \quad \text{(A14)}$$

and

$$P=\frac{y-S\cos\xi\cos\theta}{S\sin\xi\sin\theta}. \quad \text{(A15)}$$

At first we perform integration over $\bar{\rho}$ in Eq. (A14),

$$I_{\bar{\rho}}=\frac{2}{\sqrt{1-P^2}}\,\Pi\left(\frac{P}{2}\right). \quad \text{(A16)}$$

If $S<|y|$ then $P>1$ regardless of ξ and θ, thus $I_{\bar{\rho}}=0$ and $I_\xi=0$. If $S>|y|$ then $I_{\bar{\rho}}\neq 0$ only if $|P|<1$. The last condition restricts the limits of integration in I_ξ.

Using this result Eq. (A13) can be rewritten as

$$I_\xi=H(S-|y|)\int_{-\bar{z}}^{\bar{z}}\left(z+\frac{y\cos\theta}{S}\right)^k\frac{2dz}{S\sqrt{\bar{z}^2-z^2}}, \quad \text{(A17)}$$

where $z=\cos\xi-y\cos\theta/S$, $\bar{z}=\sin\theta\sqrt{1-(y/S)^2}$, and

$$H(x)=\begin{cases}0, & x\leqslant 0,\\ 1, & x>0\end{cases} \quad \text{(A18)}$$

is the Heaviside step function. We expand the binomial and perform integration over z,

$$I_\xi=\frac{2\pi}{S}\sum_{t=0}^{\lfloor k/2\rfloor}\binom{k}{2t}\frac{(2t)!}{(2^t\,t!)^2}(\sin\theta)^{2t}(\cos\theta)^{k-2t}F(y), \quad \text{(A19)}$$

where

$$F(y)=H(S-|y|)\left(1-\frac{y^2}{S^2}\right)^t\left(\frac{y}{S}\right)^{k-2t}. \quad \text{(A20)}$$

Again from the summation we remove the terms that give zero contribution.

To take the second derivative in Eq. (A12) we open the brackets in Eq. (A20) and obtain

$$\frac{\partial^2 F(y)}{(\partial y)^2}=[\mathrm{sgn}(r)]^k\sum_{v_1=0}^{k-2t}\sum_{v_2=0}^{t}\binom{k-2t}{v_1}\binom{t}{v_2}(-1)^{\bar{v}}\frac{2^{t-v_2}}{S^{\bar{v}+t}}\frac{d^2}{(d\bar{y})^2}[(\bar{y})^{\bar{v}+t}H(\bar{y})], \quad \text{(A21)}$$

where $\bar{v}=v_1+v_2$, $\mathrm{sgn}(x)$ is the sign function and $\bar{y}=S-|y|$.

Therefore, using the fact that

$$\frac{d^2}{(dx)^2}[x^j H(x)]=\begin{cases}\delta'(x), & j=0,\\ \delta(x), & j=1,\\ j(j-1)x^{j-2}H(x), & j\geqslant 2,\end{cases} \quad \text{(A22)}$$

we obtain an expression for $\sum_{lm}$:

$$\sum_{lm} = -\frac{(\mathrm{sgn}(r))^k}{4\pi S}\sum_{l=0}^{n-k}\sum_{m=0}^{k}(-1)^m\binom{n-k}{l}\binom{k}{m}\Bigg[\left(\delta'(\bar{y}) - k[2-(k-1)\tan^2\theta]\frac{\delta(\bar{y})}{2S}\right)\cos^k\theta + \sum_{t=0}^{\lfloor k/2\rfloor}\sum_{v_1=0}^{k-2t}\sum_{v_2=0}^{t}\binom{k}{2t}\binom{2t}{t}\binom{k-2t}{v_1}\binom{t}{v_2}\frac{(-1)^{\bar{v}}}{2^{t+v_2}}(\sin\theta)^{2t}(\cos\theta)^{k-2t}(\bar{v}+t)(\bar{v}+t-1)\left(\frac{\bar{y}}{S}\right)^{\bar{v}+t-2}\frac{H(\bar{y})}{S^2}\Bigg]\Bigg|_{\bar{y}=S-|r|,\, r\neq 0}. \tag{A23}$$

It is convenient to rewrite Eq. (A23) using summation over $\tilde{n} = l + m$. In this case Eq. (A10) in spherical coordinates (S,θ,ϕ) becomes

$$\bar{W}(S,\theta,\phi) = -\frac{1}{2\pi S}\sum_{n=1}^{\infty}\frac{p_n}{2^n}\sum_{k=0}^{n}\binom{n}{k}\sum_{\tilde{n}=0}^{\lfloor\frac{n-1}{2}\rfloor}B^n_{k\tilde{n}}\Bigg[\left(\delta'(S_{\tilde{n}}) - k[2-(k-1)\tan^2\theta]\frac{\delta(S_{\tilde{n}})}{2S}\right)\cos^k\theta + \sum_{t=0}^{\lfloor k/2\rfloor}\sum_{v_1=0}^{k-2t}\sum_{v_2=0}^{t}\binom{k}{2t}\binom{2t}{t}\binom{k-2t}{v_1}\binom{t}{v_2}\frac{(-1)^{\bar{v}}}{2^{t+v_2}}(\sin\theta)^{2t}(\cos\theta)^{k-2t}(\bar{v}+t)(\bar{v}+t-1)\left(\frac{S_{\tilde{n}}}{S}\right)^{\bar{v}+t-2}\frac{H(S_{\tilde{n}})}{S^2}\Bigg], \tag{A24}$$

where $S_{\tilde{n}} = S - (n - 2\tilde{n})$ and

$$B^n_{k\tilde{n}} = \sum_{\substack{m=\\ \max(0,k+\tilde{n}-n)}}^{\min(k,\tilde{n})}\binom{n-k}{\tilde{n}-m}\binom{k}{m}(-1)^m \tag{A25}$$

is a constant value that depends on n, k, and $\tilde{n}$.

If we are interested only in the main δ' contributions, $\bar{W}(S,\theta,\phi)$ can be simplified

$$\bar{W}(S,\theta,\phi) \approx -\frac{1}{2\pi S}\sum_{n=1}^{\infty}\frac{p_n}{2^n}\sum_{k=0}^{n}\binom{n}{k}\sum_{\tilde{n}=0}^{\lfloor\frac{n-1}{2}\rfloor}B^n_{k\tilde{n}}\delta'(S_{\tilde{n}})\cos^k\theta. \tag{A26}$$

Finally, the full PQPD $W(S,\theta,\phi)$ under approximation $p_0 \gg p_n$, $n \geqslant 2$, is obtained by combining Eqs. (A9) and (A24) [or (A26) for δ' contributions only]. Note that for our experimental case (single-photon and no-photon detection events) the full PQPD is obtained rigorously without any approximations.

[1] E. Wigner, Phys. Rev. **40**, 749 (1932).
[2] R. J. Glauber, Phys. Rev. Lett. **10**, 84 (1963).
[3] E. C. G. Sudarshan, Phys. Rev. Lett. **10**, 277 (1963).
[4] K. Husimi, Proc. Math. Soc. Jpn. **22**, 264 (1940).
[5] Y. Kano, J. Math. Phys. **6**, 1913 (1965).
[6] W. Schleich, *Quantum Optics in Phase Space* (Wiley-VCH, Berlin, 2001).
[7] R. L. Hudson, Rep. Math. Phys. **6**, 249 (1974).
[8] F. Soto and P. Claverie, J. Math. Phys. **24**, 97 (1983).
[9] A. I. Lvovsky, H. Hansen, T. Aichele, O. Benson, J. Mlynek, and S. Schiller, Phys. Rev. Lett. **87**, 050402 (2001).
[10] M. G. Raymer and M. Beck, Lect. Notes Phys. **649**, 235 (2004).
[11] M. V. Chekhova and F. Ya. Khalili, Phys. Rev. A **88**, 023822 (2013).
[12] G. Björk, A. B. Klimov, and L. L. Sánchez-Soto, Prog. Opt. **51**, 469 (2008).
[13] R. Simon and G. S. Agarwal, Opt. Lett. **25**, 1313 (2000).
[14] I. Rigas, L. L. Sánchez-Soto, A. B. Klimov, J. Řeháček, and Z. Hradil, Phys. Rev. A **78**, 060101(R) (2008).
[15] P. A. Bushev, V. P. Karassiov, A. V. Masalov, and A. A. Putilin, Opt. Spectrosc. (USSR) **91**, 526 (2001).
[16] V. P. Karassiov and A. V. Masalov, Laser Phys. **12**, 948 (2002).
[17] M. G. Raymer, A. Funk, and D. F. McAlister, in *Quantum Communication, Computing, and Measurement 2*, edited by P. Kumar (Plenum, New York, 2000).
[18] V. P. Karassiov and A. V. Masalov, J. Exp. Theor. Phys. **99**, 51 (2004).
[19] G. Björk, A. B. Klimov, P. de la Hoz, M. Grassl, G. Leuchs, and L. L. Sánchez-Soto, Phys. Rev. A **92**, 031801(R) (2015).
[20] C. R. Müller, L. S. Madsen, A. B. Klimov, L. L. Sánchez-Soto, G. Leuchs, C. Marquardt, and U. L. Andersen, Phys. Rev. A **93**, 033816 (2016).
[21] A. Royer, Phys. Rev. A **15**, 449 (1977).
[22] K. Banaszek and K. Wódkiewicz, Phys. Rev. A **58**, 4345 (1998).
[23] B. Kanseri, T. Iskhakov, I. Agafonov, M. Chekhova, and G. Leuchs, Phys. Rev. A **85**, 022126 (2012).
[24] Ch. Marquardt, J. Heersink, R. Dong, M. V. Chekhova, A. B. Klimov, L. L. Sánchez-Soto, U. L. Andersen, and G. Leuchs, Phys. Rev. Lett. **99**, 220401 (2007).
[25] I. N. Agafonov, M. V. Chekhova, T. Sh. Iskhakov, B. Kanseri, and G. Leuchs, JETP Lett. **96**, 496 (2012).

[26] 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, New J. Phys. **14**, 085002 (2012).
[27] N. Namekata, Y. Takahashi, G. Fujii, D. Fukuda, S. Kurimura, and Sh. Inoue, Nat. Photonics **4**, 655 (2010).
[28] R. Dong, A. Tipsmark, A. Laghaout, L. A. Krivitsky, M. Ježek, and U. L. Andersen, J. Opt. Soc. Am. B **31**, 1192 (2014).
[29] E. Maeland, IEEE Trans. Med. Imaging **7**, 213 (1988).
[30] J. A. Parker, R. V. Kenyon, and D. E. Troxel, IEEE Trans. Med. Imaging **2**, 31 (1983).
[31] A. E. Lita, A. J. Miller, and S. W. Nam, Opt. Express **16**, 3032 (2008).
[32] D. Fukuda, G. Fujii, T. Numata, K. Amemiya, A. Yoshizawa, H. Tsuchida, H. Fujino, H. Ishii, T. Itatani, S. Inoue, and T. Zama, IEEE Trans. Appl. Supercond. **21**, 241 (2011).
[33] S. Miki, T. Yamashita, Z. Wang, and H. Terai, Opt. Express **22**, 7811 (2014).
[34] M. S. Allman, V. B. Verma, M. Stevens, T. Gerrits, R. D. Horansky, A. E. Lita, F. Marsili, A. Beyer, M. D. Shaw, D. Kumor, R. Mirin, and S. W. Nam, Appl. Phys. Lett. **106**, 192601 (2015).
[35] P. C. Humphreys, B. J. Metcalf, T. Gerrits, T. Hiemstra, A. E. Lita, J. Nunn, S. W. Nam, A. Datta, W. S. Kolthammer, and I. A. Walmsley, New J. Phys. **17**, 103044 (2015).