Generating Functions and Automatic Differentiation for Photon-Number-Resolved Simulations with Multimode Gaussian States

Erik Fitzke, Florian Niederschuh, and Thomas Walther*
Institute for Applied Physics, Technische Universität Darmstadt, Schloßgartenstraße 7, 64289 Darmstadt, Germany
(Dated: September 13, 2022)

A simple and versatile method to simulate the photon statistics of multimode Gaussian states based on automatic differentiation of generating functions is presented. The generating functions for the photon number distribution, cumulative probabilities, moments, and factorial moments of the photon statistics are derived. Related expressions for multimode photon-added and photon-subtracted Gaussian states are presented. Numerical results are obtained by using the machine learning framework PyTorch for automatic differentiation. It is demonstrated that this approach is well suited for practical simulations of the photon statistics of quantum optical experiments in realistic scenarios with low photon numbers, in which various sources of imperfections have to be taken into account. As an example, the detection probabilities of a recent multipartite time-bin coding quantum key distribution setup are determined and compared with the corresponding experimental values.

Keywords: Gaussian boson sampling, Gaussian states, probability generating function, photon statistics, quantum key distribution, quantum simulation, automatic differentiation

I. INTRODUCTION

Recent progress in the generation, manipulation, and detection of photonic quantum states has led to new applications in the field of photonic quantum information processing and sensing. More and more applications for photonic quantum hardware working at the few-photon level emerge increasing the need for modeling of photonic setups for the development of quantum optical technology. While non-photon number resolving (non-PNR) single-photon detectors such as single-photon avalanche photodiodes are common, detectors allowing for photon-number resolved (PNR) detection have been developed as well [1–4]. PNR detection of quantum states opens the pathways to new experiments and applications requiring the simulation of such experiments. Thus, the photon statistics of the state and the changes in the statistics due to the optical setup need to be modeled.

Common photonic quantum states such as vacuum, coherent states, one- and two-mode squeezed states, as well as thermal states belong to the class of Gaussian states (GSs). The task to find the detection probability \( p(n) = p(n_1, \ldots, n_S) \) to detect \( n_1 \) photons in mode 1, \( n_2 \) photons in mode 2 etc. of a GS with \( S \) modes, i.e. its photon number distribution (PND), is known as the Gaussian boson sampling (GBS) problem [5]. Large-scale GBS experiments have been realized for example in the context of quantum computing and to pursue the demonstration of the computational advantage of quantum computers over classical computers [6–8]. Ref. [9] discusses possible applications for GBS-based quantum computing. The computational complexity of simulating GBS has been investigated as a benchmark for optical quantum computing comparing it to GBS simulations on classical computers [5, 10–17]. The fact that GBS is investigated in the context of computational quantum advantage indicates that calculating the solution of GBS problems can require substantial computational resources. The PND can be calculated by evaluating expressions involving matrix functions called the Hafnian and loop Hafnian for PNR detection as well as of functions called the Torontoonian and loop Torontoonian for non-PNR detection [5, 13, 16, 18, 19]. The operation count for the evaluation of these functions scales exponentially with the number of detected photons and optimization of the algorithms is an active field of research [17].

However, typically, quantum optical setups are operated at low photon numbers, and therefore investigating and minimizing the required computational resources for increasingly high photon numbers, is not always the main concern. Instead, our goal in this paper is to provide a flexible framework enabling simulations of the photon statistics including relevant experimental imperfections of the setup.

The covariance formalism describes GSs by a covariance matrix and a displacement vector and allows the modeling of many common effects on GSs in experiments, such as losses, phase shift, and interference at beam splitters, by relatively simple matrix transformations [20–23]. Therefore, it lends itself to the implementation of simulations of quantum optics experiments. Importantly, the covariance formalism carries the full information about the photon statistics of the state. We briefly describe it in section II.

Two relevant effects in experimental setups are the simultaneous detection of multiple modes by the same detector and noise in the detection process. These effects require the convolution of probability distributions which
can conveniently be expressed by the multiplication of the corresponding probability generating functions (PGFs).

Therefore, we derive generating functions of the photon statistics in terms of the covariance matrix and displacement vector in section III. Our generating-function approach is an alternative to the established expressions for the PND involving Hafnian-type functions. This different perspective on the GBS problem allows us to derive several related expressions for the generating functions of cumulative probabilities, raw or central moments, and rising or falling factorial moments of the detection statistics in section III B, for which so far no systematic method existed. Furthermore, it allows us to derive the same quantities for certain non-Gaussian states called photon-added and photon-subtracted GSs in section III C.

Generating functions need to be differentiated repeatedly to retrieve detection probabilities and moments. We show that these derivatives can be evaluated numerically by automatic differentiation (AD) without much effort. An advantage of AD is that it provides accurate numerical results while hiding the whole complexity of the calculation from the user. We discuss the usage of AD for our application in section III D.

Our simulation method consists of two steps, connected by the generating function: First, the quantum state and the optical setup are modeled in the covariance formalism. Second, the photon statistics are obtained by AD of the generating function, which is expressed in terms of the covariance matrix and displacement vector. Both steps can be easily implemented with widely available software, making the method very practical. However, it is expected that optimized algorithms for the evaluation of Hafnian-type functions will outperform general-purpose AD algorithms for the particular task of calculating the PND. To show that for small photon numbers our method can nevertheless be used to simulate various aspects of Gaussian photon statistics efficiently, we apply it to multiple examples with common GSs in section III I B, for which so far no systematic method existed. Furthermore, it allows us to derive the same quantities for certain non-Gaussian states called photon-added and photon-subtracted GSs in section III C.

In this section, we briefly summarize well-known basic properties of general multi-mode GSs.

A photonic state with $S$ modes can be described by creation and annihilation operators $\hat{a}_s^\dagger$ and $\hat{a}_s$ with $s = 1 \ldots S$. The quadrature operators $\hat{x}_s = (\hat{a}_s + \hat{a}_s^\dagger)/\sqrt{2}$ and $\hat{p}_s = (\hat{a}_s - \hat{a}_s^\dagger)/(i\sqrt{2})$ can be expressed in matrix notation as $\hat{q} = i\Omega \hat{a}$ with

$$\hat{q} = (\hat{x}_1 \ldots \hat{x}_S \hat{p}_1 \ldots \hat{p}_S)^T,$$

$$\hat{a} = (\hat{a}_1 \ldots \hat{a}_S \hat{a}_1^\dagger \ldots \hat{a}_S^\dagger)^T,$$

with the basis-changing matrix $\Omega = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -i & i \\ i & -i \end{pmatrix}$ and with $\mathbb{1}$ denoting the $S \times S$ identity matrix.

The characteristic function of an operator $\hat{O}$,

$$\chi_{\hat{O}}(\xi) = \text{tr}(\hat{O} \exp(i\xi^T \hat{q})),$$

has 2S arguments that can be collected into the vector $\xi^T = (\xi_1, \xi_2, \ldots, \xi_S, \xi_1^*, \xi_2^*, \ldots, \xi_S^*)$. The expectation value $\langle \hat{O} \rangle$ w.r.t. a quantum state $\hat{\rho}$ can be written as [20, 25]

$$\text{tr}(\hat{\rho} \hat{O}) = \frac{1}{(2\pi)^S} \int_{\mathbb{R}^{2S}} \chi_{\hat{\rho}}(\xi) \chi_{\hat{O}}(-\xi) \, d\xi.$$

GSs are the states described by a Gaussian characteristic function [20–22, 25]

$$\chi_\hat{\rho}(\xi) = \exp\left(-\frac{1}{4} \xi^T \Gamma \xi + i d^T \xi\right),$$

with a real, symmetric $2S \times 2S$ covariance matrix $\Gamma$ and a displacement vector $d$ defined by

$$\Gamma_{ij} = \langle \hat{q}_i \hat{q}_j \rangle - 2 \langle \hat{q}_i \rangle \langle \hat{q}_j \rangle \quad \text{and} \quad d_i = \langle \hat{q}_i \rangle.$$

Hamiltonians linear or quadratic in $\hat{a}$ and $\hat{a}^\dagger$ map GSs to GSs [22] and can be expressed in matrix notation as

$$\hat{H}_1 = i\hbar^T \hat{a} \quad \text{with} \quad \hbar = \begin{pmatrix} -\alpha^* \\ \alpha \end{pmatrix} \quad \text{and} \quad (7)$$

$$\hat{H}_2 = \frac{1}{2} \hat{a}^\dagger H \hat{a} \quad \text{with} \quad H = \begin{pmatrix} X & Y \\ Y & X^* \end{pmatrix}. \quad (8)$$

Here, $Y = Y^T$ as well as $X = X^\dagger$ ensure that $\hat{H}$ is Hermitian [23]. A unitary operation $\hat{U}_{1,2} = e^{-i\hat{H}_{1,2}}$ transforms $\hat{q}$ according to

$$\hat{U}_1^\dagger \hat{q} \hat{U}_1 = \hat{q} + \Omega \mathbf{J} \hbar = \hat{q} + \hat{d}, \quad (9)$$

$$\hat{U}_2^\dagger \hat{q} \hat{U}_2 = S \hat{q} \quad \text{where} \quad S = \Omega e^{-i \mathbf{KH} \Omega^T}. \quad (10)$$

As a more complex application, we present in section V a simulation of an entanglement-based quantum key distribution (QKD) system we have presented recently in Ref. 24. It demonstrates the strengths of our method to readily incorporate relevant imperfections such as noise, detection efficiencies, and simultaneous detection of multiple modes. We show that the simulation results are in very good agreement with the experimental values. Furthermore, we show that the contribution of multi-photon pair emission to the quantum bit error rate of the QKD system can be estimated by applying Bayes' theorem to PNR simulation results. Thereby, we demonstrate how PNR simulations enable the calculation of quantities that are not directly accessible in the experiment.

---

1 We use the convention from Refs. [20, 23, 25], as our notation benefits from grouping $x$ and $p$ components together and the covariance of the vacuum states simply becomes $\Gamma = 1$. The expressions we derive will depend on $\Gamma - 1$. Other common conventions are to arrange $x$ and $p$ in alternating order [21, 22], to scale $\Gamma$ by a factor of $1/2$ [5, 14, 22], or to define $\Gamma$ with complex elements with respect to $\hat{a}$ instead of $\hat{q}$ [5, 14, 25].
with GSs to investigate continuous-variable quantum information

and the coupling of two modes by a lossless beam splitter

as well as \( J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \) and \( K = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \).

The characteristic function becomes

\[
\chi_{\hat{\varphi}}(\xi) = \text{tr}(\hat{U} \hat{\varphi} \hat{U}^\dagger \exp(i\xi^T \hat{q} )) = \text{tr}(\hat{\varphi} \exp(i\xi^T \hat{U}^\dagger \hat{q} \hat{U} )) , \tag{11}\]

resulting in transformations of \( \Gamma \) and \( d \) \cite{20, 21, 22, 23}:

\[
\hat{\varphi}' = \hat{U}_1 \hat{\varphi} \hat{U}_1^\dagger \quad \Rightarrow \quad \Gamma' = \Gamma \\
\hat{\varphi}' = \hat{U}_2 \hat{\varphi} \hat{U}_2^\dagger \quad \Rightarrow \quad \Gamma' = S \Gamma S^T \\
d' = d + \hat{d} \tag{12}
\]

The transformation matrices \( S \) are symplectic with respect to \( J \), i.e. they fulfill \( SJS^T = J \) \cite{23}.

Examples for such transformations are the phase shift \( \hat{a} \rightarrow e^{-i\varphi} \hat{a} \) of a single mode, represented by \cite{22, 25}

\[
S_{\text{ph}} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}, \tag{14}\]

and the coupling of two modes by a lossless beam splitter with intensity transmission \( T \), represented by \cite{25}

\[
S_{\text{BS}} = \begin{pmatrix} \sqrt{T} & 0 & 0 \\ -\sqrt{1-T} & \sqrt{1-T} & 0 \\ 0 & 0 & \sqrt{T} \end{pmatrix} \tag{15}
\]

If a beam splitter is introduced into one of the modes of the state, an additional mode containing vacuum needs to be introduced for the second beam splitter input before the beam splitter transformation can be applied. This is performed by inserting two new columns and rows for the \( x \) and \( p \) components with 1 on the diagonal. Conversely, after discarding a subsystem by performing the partial trace operation the covariance matrix of the reduced state \( \hat{\varphi} \) is simply obtained by deleting all rows and columns associated with the traced-out subsystem from \( \Gamma \) and \( d \). Thereby, losses \( L \) can be modeled by introducing an auxiliary mode containing vacuum, coupling the modes with a beam splitter with transmission \( T = 1 - L \) and tracing out the auxiliary mode. Thus, the overall loss transformation for a single-mode GS yields \cite{25}

\[
\hat{\Gamma} = \hat{T} \hat{T}^T + L \mathbb{1} , \quad \hat{d} = T \hat{d} \tag{16}
\]

with the transmission matrix \( T = \sqrt{1-L} \mathbb{1} \).

The covariance formalism has been used, for example, to investigate continuous-variable quantum information with GSs \cite{20, 21, 22, 23} and to analyze multi-photon effects in setups with GSs and non-PNR detectors \cite{25}, QKD with entanglement swapping \cite{26}, optimal conditions for Bell tests \cite{27} and Hong-Ou-Mandel interference involving multiple spectral modes \cite{28}.

In Ref. 25, Takeoka et al. introduce a method representing imperfect interference at a beam splitter with the covariance formalism by considering a mode mismatch: each of the two input modes is split into an interfering and a non-interfering part, only the interfering parts show interference at the beam splitter and the non-interfering parts are treated as separate modes but are guided to the same detectors. The model can be used to describe a mismatch for example in the polarization, temporal or spatial degrees of freedom. We will use it to simulate imperfect two-photon interference in our QKD setup in section V.

### III. Generating Functions for the Detection Statistics of Gaussian States

Our method to simulate the photon statistics of GSs uses generating functions for the photon statistics in terms of the covariance matrix \( \Gamma \) and the displacement vector \( d \). They connect the covariance formalism to the photon statistics and are therefore essential for our simulation method. In appendix A we briefly summarize the relevant, well-known, properties of generating functions for probability distributions. In the following section III A we briefly recap the formulation of the PND in terms of a generating operator and extend it to generating operators for the moments and factorial moments.

#### A. Generating operators for photon detection probabilities and moments

Assuming a photon detection process where photons in a single mode are detected independently of each other with an efficiency \( \eta \), the probability to detect \( n \) photons from a quantum state \( \varphi \) is given by \cite{29, 30, 31, 32}:

\[
p(n) = \left\langle \frac{(\eta \hat{N})^n}{n!} e^{-\eta \hat{N}} \right\rangle. \tag{17}\]

Here, normal order is indicated by \( : : \) and the photon number operator is \( \hat{N} = \hat{a}^\dagger \hat{a} \). By inserting eq. (17) into the defining equation of a PGF \( h(y) \) (cf. eq. (A1)), the generating function of the photon detection probabilities is obtained as the expectation value of a generating operator \( h(y) \) \cite{29, 30, 31, 32}. We extend this procedure to obtain similar generating operators for the moment generating function \( M(\mu, y) \) and rising factorial moment generating function \( R(y) \) as described in appendix A:

\[
h(y) = \sum_{k=0}^{\infty} y^k \frac{(\eta \hat{N})^k}{k!} e^{-\eta \hat{N}} = e^{\eta \hat{N} (y-1)}, \tag{18}\]

\[
M(\mu, y) = \sum_{k=0}^{\infty} e^{\mu(k-\mu)} \frac{(\eta \hat{N})^k}{k!} e^{-\eta \hat{N}} = e^{-\eta \mu} \exp(\eta(e^y - 1) \hat{N}) , \tag{19}\]

\[
R(y) = \sum_{k=0}^{\infty} \frac{(\eta \hat{N})^k e^{-\eta \hat{N}}}{k!(1-y)^{k+1}} = \frac{1}{1-y} e^{\eta y \hat{N}/(1-y)}. \tag{20}\]

These operators can be modified to take into account noise in the detection process. For that, according to
the known multiplication rule for generating functions (cf. appendix A), the generating operator is simply multiplied by the generating function of the noise process. As an example, we consider noise with Poissonian statistics $p_{\text{noise}}(n) = e^{-\nu} \nu^n / n!$ and noise parameter $\nu$. The noise PGF is given by $h_{\text{noise}}(y) = e^{\nu(y-1)}$ so that the generating operators including noise read:

$$\hat{h}_\nu(y) = e^{\nu(y-1)} h(y),$$  

$$\hat{M}_\nu(\mu, y) = \hat{M}(\mu, y) \exp(\nu(e^y - 1)),$$  

$$\hat{R}_\nu(y) = e^{\nu y / (1-y)} \hat{R}(y).$$

Similarly, noise with different statistics could be taken into account by multiplying the generating operators from eqs. (18) to (20) with the generating function of the respective noise process. Another option is to include noise in the covariance formalism. The beam splitter transformation from eq. (15) can be used to model noise by coupling in a thermal state to introduce thermal noise or a matrix can be added to the covariance matrix to represent e.g. classical Gaussian noise or noise from amplification [33].

Equations (18) to (20) have in common that they involve special cases of the generating operator $\hat{g}(w(y)) = \exp(-w(y) a \dag \hat{a})$: for different functions $w(y)$. For example, the $n$-photon detection probability, taking into account the detector efficiency $\eta$ and noise $\nu$, can be expressed by a detection operator $\hat{I}$:

$$p(n) = \langle \hat{I} N = n \rangle = \left. \frac{1}{n!} \frac{d^n}{d y^n} e^{\nu(y-1)} \hat{g}(\eta(1-y)) \right|_{y=0}. \tag{24}$$

In Ref. 34 and 35 the operator $\hat{g}$ has been extended to

$$\hat{g}(u, v, w) = \exp(u \hat{a} + v \hat{a}^\dag - w \hat{a}^\dag \hat{a}); \tag{25}$$

and has been related to the density matrix elements of $\hat{g}$. Here we present an alternative, simple derivation of these relations by expanding the quantum state in the overcomplete basis of coherent states $\hat{\gamma}$ and has been related to the density matrix elements of $\hat{g}$ in the basis of coherent states:

$$\langle \alpha | \hat{g} | \beta \rangle = \int_{\mathcal{C}} P(\gamma) \langle \alpha | \gamma \rangle \langle \gamma | \beta \rangle d^2 \gamma = \int_{\mathcal{C}} P(\gamma) e^{-((|\alpha|^2 + |\beta|^2) / 2 - |\gamma|^2 + \alpha^* \gamma + \gamma^* \beta) d^2 \gamma} = e^{-(|\alpha|^2 + |\beta|^2)/2} \langle \gamma^* \alpha^*, \beta, 1 \rangle. \tag{28}$$

As a special case of eq. (28) we obtain the Q-function $Q(\alpha) = \langle \alpha | \hat{g} | \alpha \rangle / \pi = e^{-|\alpha|^2} \langle \hat{g}(\alpha^*, \alpha) \rangle / \pi$. By using $\langle n | \alpha \rangle = \alpha^n e^{-|\alpha|^2} / \sqrt{n!}$ for photon numbers $n$ and $m$ and comparing eq. (27) to the expression for $\langle n | \hat{g} | m \rangle$, i.e.

$$\langle n | \hat{g} | m \rangle = \int_{\mathcal{C}} P(\alpha) \langle n | \alpha \rangle \langle m | \alpha \rangle d^2 \alpha = \int_{\mathcal{C}} P(\alpha) e^{-|\alpha|^2} \frac{\alpha^n \alpha^m}{\sqrt{n!} \sqrt{m!}} d^2 \alpha, \tag{29}$$

it can be seen that $\hat{g}$ generates matrix elements in the photon number basis via

$$\langle n | \hat{g} | m + \Delta_m \rangle = \left. \frac{1}{\sqrt{n!} m!} \partial_u^n \partial_v^m \hat{g}(u, v, 1) \right|_{u=0, v=0}. \tag{30}$$

Our goal is to evaluate the derivatives numerically. We therefore note that for $m = n + \Delta_m$ and $\Delta_m > 0$, the number of derivatives can be reduced from $n + m$ to $m$ by using

$$\langle n | \hat{g} | n + \Delta_m \rangle = \frac{(-1)^n}{\sqrt{n!} m!} \partial_u^n \partial_v^m \hat{g}(0, v, w) \bigg|_{u=0, v=1}, \tag{31}$$

which is again directly evident from the phase space integral. Similar to eq. (31), $(m + \Delta_m | \hat{g} | m)$ can be calculated for $n > m$ by swapping the roles of $u$ and $v$. Equation (31) is especially useful for matrix elements close to the diagonal as it then considerably reduces the number of derivatives.

### B. Generating functions in terms of the covariance matrix and displacement vector

We now derive the generating function $G(u, v, w) = \langle \hat{g}(u, v, w) \rangle$ for GSs in terms of $\mathbf{T}$ and $\mathbf{d}$.

Inserting $\hat{I} = \pi^{-1} \int \mathcal{C} \left. \langle \alpha | \hat{g} | \alpha \rangle d^2 \alpha \right.$ into eq. (26) allows us to calculate the trace of normally ordered operators via $\langle \mathbf{F}(\hat{a}^\dag, \hat{a}) \rangle = \pi^{-1} \int_{\mathcal{C}} F(\alpha^*, \alpha) d^2 \alpha$. By applying the Baker-Campbell-Hausdorff formula [31] and cyclic permutation of the trace, we obtain the characteristic function from the trace of a normally-ordered operator by

$$\chi_\beta(\xi) = \text{tr}(\hat{g} e^{i(\xi \hat{a}^\dag + \xi^* \hat{a})}) = e^{(\xi^2 + \xi^* \xi)/4} \text{tr}(e^{i\hat{a}^\dag \hat{g} e^\hat{a}}) \tag{32}$$
with \( c = (i\xi_x + \xi_p)/\sqrt{2} \) and \( d = (i\xi_x - \xi_p)/\sqrt{2} \). By separating
the real and imaginary parts of \( \alpha = a + ib \) we obtain
\[
\text{tr}(e^{d\hat{a}^\dagger \hat{a}} e^{i\alpha}) = \frac{1}{\pi} \int_\mathbb{C} e^{(d + iv \xi + vd^2)} d\bar{\alpha} = \frac{1}{\pi} \int_\mathbb{R} e^{(d + iv \xi + vd^2)} da \int_\mathbb{R} e^{i(c + u \xi + ub^2)} db .
\]

(33)

Evaluating the complex Gaussian integrals yields
\[
\chi_{\beta}(\xi) = \frac{1}{w} \exp \left( -\frac{1}{2} \frac{w}{\sqrt{2}} \xi^T \xi - i\xi^T \xi + \frac{uv}{w} \right) \quad (34)
\]

with \( \zeta = \frac{1}{w\sqrt{2}} \left( -(u + v) \quad i(u - v) \right) \) and \( \xi = \left( \xi_x \quad \xi_p \right) \).

In experiments, often multiple modes, e.g. different polarization directions or frequency modes, enter the same detector. For example, calculating the simultaneous detection of multiple modes is required for modeling imperfect interference with the model from Ref. 25 which we have mentioned in section II and used in section V. Furthermore, often joint detection probabilities between multiple detectors such as coincidence probabilities are of interest. Therefore, we generalize the calculation to multi-modes states. By performing the above calculation for each mode separately, we generalize eq. (34) to a multi-mode state with modes \( s = 1 \ldots S \) by defining \( Z = \sum_s u_s \xi_s w_s^{-1} \) and \( \Lambda = \text{diag}(s = 1 \ldots S) \) and extending \( \xi \) and \( \zeta \) accordingly:

\[
\chi_{\beta}(\xi) = w^{-1} \exp \left( -\frac{1}{2} \frac{w}{\sqrt{2}} \xi^T \Lambda^{-1} \Lambda \xi - i\xi^T \xi + Z \right) . \quad (35)
\]

Here, the direct sum \( \Lambda = \Lambda \oplus \Lambda \) is used to apply \( \Lambda \) to the \( \xi \) and \( \xi \) components in \( \xi \). By using eqs. (4) and (5), we calculate \( G(u, v, w) = \langle \hat{g}(u, v, w) \rangle \) for the detection probabilities of a GS with characteristic function \( \chi_{\theta}(\xi) = \exp(-\xi^T T \xi/4 + i\xi^T \chi) \) and obtain

\[
G(u, v, w) = \frac{1}{(2\pi)^S} \int_{R^{2s}} \chi_{\theta}(\xi) \chi_{\theta}(\eta)(-\xi) d\xi \quad (36)
\]

\[
= \frac{w^{-1}}{(2\pi)^S} \int_{R^{2s}} \exp \left( -\frac{1}{2} \xi^T B \xi + i\xi^T Z \right) d\xi \quad (37)
\]

with \( B = \Gamma/2 + (\Lambda + \Lambda) \) and \( z = d + \zeta \).

We absorb the prefactor \( w^{-1} \) into a diagonal matrix \( \Lambda = \text{diag}(u) \oplus \text{diag}(v) \) and obtain

\[
G(u, v, w) = \frac{\exp(-z^T \Lambda^{-1} W z/2 + Z)}{\sqrt{\det \Lambda}} \quad (38)
\]

Equation (37) is the generating function for the multivariate photon statistics from which the probabilities and moments are retrieved by repeated differentiation w.r.t. \( D \) parameters \( y_1 \ldots y_D \) for the detectors \( j = 1 \ldots D \), detecting \( M_j \) modes respectively, with additional Poissonian noise \( \nu \). The total number of modes is \( S = \sum_j M_j \) and the index \( s \) runs over all mode indices, i.e. enumerates \( m_j = 1_j \ldots M_j \) for all \( D \) detectors in the order \( 1_1, 2_1, \ldots, M_1, 1_2, \ldots, M_2, \ldots, M_D \). We obtain:

\[
p(n, \nu, \eta) = \langle \hat{N}_{\leq n} \rangle = \frac{1}{n!} \partial_n^\nu \exp \left( \sum_j (y_j - 1) \nu_j \right) G(0, 0, w) \Big|_{y=0} \quad \text{with} \quad w_s = \eta_m (1 - y_j) \quad (39)
\]

\[
p(N \leq n, \nu, \eta) = \langle \hat{N}_{\leq n} \rangle = \frac{1}{n!} \partial_n^\nu (1 - y)^{-1} \exp \left( \sum_j (y_j - 1) \nu_j \right) G(0, 0, w) \Big|_{y=0} \quad \text{with} \quad w_s = \eta_m (1 - y_j) \quad (40)
\]

\[
\mathcal{M}(\nu, k, \nu, \eta) = \langle (\hat{N} - \mu)^k \rangle = \partial_k^\nu \exp \left( \sum_j (e^{y_j} - 1) \nu_j - \mu_j y_j \right) G(0, 0, w) \Big|_{y=0} \quad \text{with} \quad w_s = \eta_m (1 - e^{y_j}) \quad (41)
\]

\[
n_{(k)}(\nu, \eta) = \langle \hat{N}^{(k)} \rangle = \prod_{j=1}^D \hat{N}_{(k)} \bigg|_{y=0} = \partial_k^\nu \exp \left( \sum_j y_j \nu_j \right) G(0, 0, w) \Big|_{y=0} \quad \text{with} \quad w_s = -\eta_m y_j \quad (42)
\]

\[
n_{(k)}(\nu, \eta) = \langle \hat{N}^{(k)} \rangle = \prod_{j=1}^D \hat{N}_{(k)} \bigg|_{y=0} = \partial_k^\nu (1 - y)^{-1} \exp \left( \sum_j y_j \nu_j \right) G(0, 0, w) \Big|_{y=0} \quad \text{with} \quad w_s = \eta_m y_j \quad (43)
\]

2 Here we use multi-index notation: For tuples \( x \) and \( k \) we write \( x^k = \prod_i x_i^{k_i} \), so that e.g. \( w^{-1} = \prod_i w_i^{-1} \). Below we will also use \( n! = \prod_i n_i! \) and multivariate derivatives \( \partial_n^\nu = \prod_i \partial y_i^{\nu_i} \).

3 The integral in eq. (36) is solved by using the identity

\[
\int_{R^{2s}} \exp \left( -\frac{1}{2} \xi^T B \xi + i\xi^T \xi \right) d\xi = \frac{(2\pi)^S}{\sqrt{\det B}} \exp \left( -\frac{1}{2} \xi^T B^{-1} \xi \right)
\]

and by using \( w^{-1}/\sqrt{\det B} = 1/\sqrt{\det W B} \) and \( B^{-1} = A^{-1} W \).

4 Similarly, the generating function of noise processes with a different, non-Poissonian, statistics could be taken into account by multiplying with the corresponding generating function.
Equations (37) to (43) are central results of this paper. Alternative expressions to calculate $p(n)$, restricted to the noiseless case, can be found in the literature and are discussed and compared to eq. (39) in section VI. Our generating-function approach additionally yields eqs. (40) to (43) for the cumulative probabilities, raw moments (for $\mu=0$), central moments (for $\mu=(N)$), and rising and falling factorial moments, which directly incorporate the detection efficiency and noise. For general multi-mode GSs, to the best of our knowledge, expressions for these quantities have so far not been presented in the literature.

The cumulative probabilities are useful when ranges of photon numbers are to be analyzed jointly, especially when detection events in multiple detectors are considered. For example, the calculation of the coincidence probability to detect $\leq n_1$ photons in one detector and $\leq n_2$ from the PND requires the evaluation of the probabilities for $(n_1+1)/(n_2+1)$ combinations of photon numbers, but only one evaluation of eq. (40). Mean and variance are two of the most important moments of the PND and can be directly calculated from eq. (41). The factorial moments are useful for the calculation of photon statistics of photon-added and photon-subtracted states (cf. section III C).

By using the formulation with detection operators $\Pi$, more complex detection events involving multiple detectors can be calculated. For example, the probability for the detection of $n_A$ photons in detector $A$ and $n_B$ photons in detector $B$ is given by $\langle \hat{\Pi}_{N_A=n_B} \hat{\Pi}_{N_B=n_B} \rangle = \langle \hat{\Pi}_{N_A=n_B} \hat{\Pi}_{N_B=n_B} \rangle$ and the probability for $n_1$ or $n_2$ photons in the same detector with $n_1 \neq n_2$ is given by $\langle \hat{\Pi}_{N_1=n_1} \hat{\Pi}_{N_2=n_2} \rangle = \langle \hat{\Pi}_{N_1=n_1} \hat{\Pi}_{N_2=n_2} \rangle$. Operators for complementary events can be defined as well enabling the calculation of the probability to detect any photon number except for $n$ or for more than $n$ photons with $\hat{\Pi}_{N \neq n} = \hat{1} - \hat{\Pi}_{N=n}$ and $\hat{\Pi}_{N \geq n} = \hat{1} - \hat{\Pi}_{N<n}$. The detection operators for different modes are then joined before the generating function is evaluated. This procedure was used for example in Ref. 25 to relate the detection probabilities for non-PNR detectors to the detection of vacuum by considering operators $\hat{\Pi}_{N>0} = \hat{1} - \hat{\Pi}_{N=0}$. We extend this method to PNR detection. As an example consider the probability to detect not $n_A$ photons in detector $A$ and more than $n_B$ photons in detector $B$:

$$\langle \hat{\Pi}_{N_A \neq n_A} \hat{\Pi}_{N_B>n_B} \rangle = \langle \hat{1} - \hat{\Pi}_{N_A=n_A} \hat{\Pi}_{N_B=n_B} \rangle = 1 - \frac{1}{n_A!} \frac{\partial^n_{a^n_A}}{\sqrt{\det A_A}} \exp(-d_A^{-1} A_A^{-1} W_A d_A/2) \bigg|_{y_A=0}$$

$$\langle \hat{\Pi}_{N_A=n_A} \hat{\Pi}_{N_B<n_B} \rangle = 1 - \frac{1}{n_B!} \frac{\partial^n_{b^n_B}}{\sqrt{\det A_B}} \exp(-d_B^{-1} A_B^{-1} W_B d_B/2) \bigg|_{y_B=0}$$

$$\langle \hat{\Pi}_{N_A=n_A} \hat{\Pi}_{N_B=n_B} \rangle = \frac{1}{n_A! n_B!} \frac{\partial^{n_A+n_B}_{a^n_A b^n_B}}{\sqrt{\det A_{AB}}} \exp(-d_{AB}^{-1} A_{AB}^{-1} W_{AB} d_{AB}/2) \bigg|_{y_A=0, y_B=0}$$

$$\langle \hat{\Pi}_{N_A=n_A} \hat{\Pi}_{N_B>n_B} \rangle = \frac{1}{n_A! (n_B-1)!} \exp(-(d_B^{-1} A_B^{-1} W_B d_B/2) \bigg|_{y_B=0}$$

Here, $d$, $W$ and $A$ contain all rows and columns for modes that enter detector $A$, $B$ or both detectors labeled by the indices $A, B$ and $AB$, respectively.

Finally, we derive multivariate expressions for matrix elements from eqs. (28), (30) and (31):

$$\langle \alpha \mid \hat{\varphi} \mid \beta \rangle = e^{-((\alpha^2+|\beta|^2)/2)} G(\alpha^*, \beta, 1) ,$$

$$\langle n \mid \hat{\varphi} \mid m \rangle = \frac{1}{\sqrt{n!m!}} \frac{\partial^n_{a^n} \partial^m_{b^m}}{\det A} G(u, v, 1) \bigg|_{u=0, v=0}$$

$$\langle n \mid \hat{\varphi} \mid m \rangle = \frac{(-1)^l}{\sqrt{n!m!}} \frac{\partial^n_{a^n} \partial^{\Delta n}_{a^n} \partial^{\Delta m}_{b^m} G(u, v, w) \bigg|_{u=0, v=0, w=1}$$

In eq. (47) we define $l$ by $s = \min(a_s, b_s)$ as well as $\Delta n = n - l$ and $\Delta m = m - l$. In Ref. 38, eq. (46) has already been derived differently and it has been related to an expression involving the loop Hafnian function. Compared to eq. (46), our eq. (47) has the advantage that for each mode the number of derivatives to be evaluated is reduced from $a_s + b_s$ to max($a_s, b_s$). When the derivatives are evaluated numerically, for example by automatic differentiation, our expression is especially advantageous for close-by values of $a_s$ and $b_s$.

The PND of the state itself without taking into account noise and detection efficiencies, $\langle n \mid \hat{\varphi} \mid n \rangle$, can be obtained from eq. (46) with $m = n$. However, this requires twice as many derivatives as eq. (39) or eq. (47). As we evaluate the derivatives numerically and because eq. (39) naturally includes the effects from noise and non-unify detection efficiencies, we prefer eq. (39) over eq. (46) for the calculation of the PND.

**C. Photon statistics of non-Gaussian states derived from Gaussian states**

To extend the range of possible applications for our simulation method of the photon statistics, we now show how it can be adapted to also be applied to certain non-Gaussian states.

A special class of states derived from GSs $\varphi$ are the photon-added GSs $\varphi_{+k}$ and the photon-subtracted GSs $\varphi_{-k}$ [39],

$$\varphi_{+k} = \frac{\hat{\varphi}^k}{m(k)} \quad \text{and} \quad \varphi_{-k} = \frac{\hat{\varphi}_{-k}^k}{m(k)} ,$$

with $k = k_1, k_2, \ldots$ photons added or subtracted from the modes 1, 2, \ldots. Here, the normally and anti-normally ordered moments [34] $m(k) = \text{tr}(\hat{\varphi}^k \hat{\varphi}_{-k}^k) = \langle \hat{\varphi}^k \hat{\varphi}_{-k}^k \rangle$ and $m(k) = \text{tr}(\hat{\varphi}_{-k}^k \hat{\varphi}^k) = \langle \hat{\varphi}_{-k}^k \hat{\varphi}^k \rangle$ ensure normalization. They can be obtained from the falling and rising factorial moments of the GS by setting $\eta = 1$ and $\nu = 0$:

$$m(k) = \sum_{n \geq k} \langle n \mid \hat{\varphi} \mid n \rangle \frac{n!}{(n-k)!} = n_k(\nu = 0, \eta = 1) ,$$

$$m(k) = \sum_n \langle n \mid \hat{\varphi} \mid n \rangle \frac{(n+k)!}{n!} = n_k(\nu = 0, \eta = 1) .$$
Here, our eqs. (42) and (43) for factorial moments are helpful because they allow to calculate \( n_l(k) \) and \( n_l(k) \) directly.

Photon addition and photon subtraction can have non-trivial effects on the PND of quantum states. An example is photon subtraction from a thermal state, which increases the mean photon number of the state [39, 40]. Photon-subtracted states can be approximately realized by inserting a beam splitter with high transmission into the beam and by conditioning the detection of the transmitted quantum state on the detection of a reflected photon [39–41]. Photon addition has been realized by seeding a spontaneous parametric down-conversion (SPDC) process with low gain so that the SPDC signal modes are emitted into the mode of the incident state. The generation of the photon-added state is then conditioned on the detection of an SPDC idler photon [42, 43]. Photon addition and subtraction can be used for example to enhance the signal-to-noise ratio in quantum ghost imaging [44].

The matrix elements of photon-added and photon-subtracted GSs can be directly calculated by using \( \hat{a}^k |l⟩ = \sqrt{l!/|l-k|!} |l-k⟩ \), given that \( l \geq k \), and \( \hat{a}^{lk} |l⟩ = (k+l)/l [k+l]! \):

\[
\begin{align*}
\langle n | \hat{o}_{-k}|l⟩ &= \frac{(n+k)!}{m(k)} \sum_{l} \frac{1}{l! n!} \langle l+k | n+k⟩, \\
\langle n | \hat{o}_{+k}|l⟩ &= \frac{(n-k)!}{m(k)} \sum_{l} \frac{1}{(l-k)! (n-k)!} \langle l-k | n-k⟩.
\end{align*}
\]

The expressions for \( \langle n + k | \hat{o} | l + k⟩ \) and \( \langle n - k | \hat{o} | l - k⟩ \) can be calculated from the GS’s matrix elements by using eq. (46) or eq. (47).

To derive explicit expressions for the photon statistics, we first consider the single mode case of \( \hat{o}_{-k} \). The operator \( \hat{a}^k \hat{g} \hat{a}^\dagger \) is in normal order and we obtain from the optical equivalence theorem in eq. (26) \( \text{tr} (\hat{a}^k \hat{g} \hat{a}^\dagger) = \int \mathcal{D}(\alpha) |\alpha|^2 e^{(\alpha^d + \alpha^e) - |\alpha|^2} \text{d}^2 \alpha \). By comparison to the phase-space representation of \( \hat{g} \) from eq. (27), the expectation value of \( \hat{g} \) with respect to \( \hat{o}_{-k} \) can be expressed in terms of the generating function of the underlying GS \( \hat{g} \):

\[
\text{tr}(\hat{o}_{-k} \hat{g}(u, v, w)) = \frac{(-1)^k}{m(k)} \partial^k_w G(u, v, w).
\]

For the photon-added GS, we consider for simplicity only \( \langle g(0, 0, w) \rangle \), as \( u \) and \( v \) are essentially only required for the calculation of the matrix elements, which has already been discussed above. In appendix B, we derive the expression for the single-mode case which generalizes to

\[
\begin{align*}
\text{tr}(\hat{o}_{+k} \hat{g}(0, 0, w)) &= \frac{1}{m(k)} \partial^k_w G(0, 0, w) \prod_s \left| 1 - r_s(1 - w_s) \right| \bigg|_{r=0}, \\
\end{align*}
\]

with \( \tilde{w}_s = 1 - [(1 - w_s)^{-1} - r_s]^{-1} \).

This means that all the quantities of the photon statistics for which we have derived expressions above can also be calculated for multi-mode photon-added and photon-subtracted GSs. For that, \( G(0, 0, w) = \text{tr}(\hat{g} \hat{g}(0, 0, w)) \) in eqs. (39) to (43) is replaced by the expressions for \( \text{tr}(\hat{o}_{-k} \hat{g}(0, 0, w)) \) and \( \text{tr}(\hat{o}_{+k} \hat{g}(0, 0, w)) \) from eqs. (53) and (54).

To our knowledge, expressions for the PND of photon-added and photon-subtracted GSs have so far only been reported for single-mode states [45–47]. Our expressions enable the calculation of the photon statistics for the general multi-mode case and additionally enable the calculation of moments, factorial moments, and matrix elements. The expressions are as well evaluated by repeated differentiation and can therefore be calculated with the same tools as the expressions for GSs.

Certain non-Gaussian states \( \hat{g} \), such as NOON states and cat states, can be generated from larger GSs \( \hat{g} \) by PNR detection [38, 48–50]. The effects of the optical setup on such states can be modeled using the covariance formalism as described in Ref. 49: consider a GS \( \hat{g} \) with \((M' + M)\) modes, of which \( M \) modes are detected with PNR detectors so that a state \( \hat{g}' \) with \( M' \) modes remains. By conditioning the further analysis of \( \hat{g}' \) on the detected PND, the non-Gaussian states are produced probabilistically [49]. Unitary transformations \( U' \) acting on \( \hat{g}' \) can simply be modeled by applying the unitary transformation \( U = U' \otimes \mathbb{1}_{M \times M} \) to \( \hat{g} \). For the simulation, the required PNR detection pattern on the \( M \) modes is fixed and the photon detection pattern of interest is applied to the remaining \( M' \) modes. Unitary transformations that are linear or quadratic in \( \hat{a} \) and \( \hat{a}^\dagger \) can be applied to \( \hat{g} \) in the covariance matrix formalism before the detection [49]. By using this procedure from Ref. 49, our method can be directly applied to simulate states that can be obtained from GSs via PNR detection.

To summarize, we note that our method to simulate the photon statistics of GSs can also be applied to photon-added and photon-subtracted GSs as well as to non-Gaussian states derived from larger GSs by PNR detection. The class of states that can be simulated and the range of possible applications of the simulation method are thereby greatly extended.

D. Automatic differentiation for retrieval of probabilities and moments

Multiple options exist to evaluate the derivatives of the generating functions. One option used in Ref. 28 is to use finite-difference approximations, but this method can accumulate numerical inaccuracies. Another option to retrieve probabilities from the PGF is to approximate Cauchy’s integral formula on a circle \( \gamma \) around the origin in the complex plane:

\[
h(y) = \sum_{n=0}^{\infty} p_n y^n \Rightarrow p_n = \frac{1}{2\pi i} \int_{\gamma} h(z) z^{-n-1} \, dz.
\]
In Ref. 51 this method is applied to probability generating functions and error bounds for the approximations are given. It is extensively discussed in Ref. 52 and expressions for moments and multi-dimensional distributions are given in Refs. [53–55].

A versatile method to evaluate derivatives of functions numerically is automatic differentiation (AD) [56, 57]. AD breaks down a function to be differentiated into elementary operations such as addition, multiplication, or sin(x). The differentiation rules for these operations are then applied and combined according to the chain rule to the derivative of the overall function, which is finally evaluated. For example, instead of approximating the derivative of sin(x) numerically, automatic differentiation uses the fact that the derivative is given by cos(x) and evaluates the cosine function. Hence, in contrast to finite-difference approximations, the derivatives obtained are accurate to the working precision.

Applying AD to the generating functions allows us to readily calculate all the quantities we have derived. Equations (39) to (43) as well as eqs. (53) and (54) only involve derivatives and basic functions as well as linear algebra that are easily implemented. With modern software tools, multivariate higher-order derivatives can be conveniently obtained with a few lines of code. From a practical point of view, this is convenient because the user is not confronted with the complexity of the calculation, which is hidden in the repeated derivatives. AD thereby greatly improves the practical usability of the generating functions we derived. However, as mentioned in the introduction, the computational complexity to calculate the PND of GBS problems with state-of-the-art algorithms scales exponentially with the number of photons, so that the numerical evaluation of the derivatives of the generating function via AD will become infeasible for large photon numbers.

A variety of AD software exists and an overview of available tools is provided in Ref. 58. AD is widely used in machine learning, for example for training artificial neural networks [59]. Therefore, AD functionalities are a part of popular machine learning libraries such as PyTorch [60, 61] and TensorFlow [62]. Machine learning has gained considerable attention over the last years and the range of its applications is constantly growing. Consequently, software for machine learning will be further developed and it can therefore be expected that AD software will gain further capabilities during the next years, which of course is beneficial for our application. State-of-the-art machine learning libraries provide numerous possibilities to optimize their performance, for example, using parallel computing on graphic processing units (GPUs).

Throughout this paper, we use the autograd function from the machine learning framework PyTorch 1.11.0 for AD. We use PyTorch in a very basic configuration, i.e. we do not use the option for acceleration by GPU computations and the only setting we change is the default precision which we increase from float32 to float64.

All simulations are run on a regular desktop computer to show that for small photon numbers the generating functions can be evaluated without much effort, demonstrating the practicability of this approach.

IV. APPLICATION TO COMMON GAUSSIAN STATES

In this section, we apply our simulation method to two common GSs with well-known photon statistics: the single-mode displaced squeezed thermal state and the two-mode squeezed vacuum with multiple squeezers. These practical examples demonstrate that our simulation method can easily take effects such as non-unity detection efficiencies, noise, and multiple modes entering the same detector into account.

A. Photon number distribution of a single-mode displaced squeezed thermal state

As a first example, we consider the single-mode displaced squeezed thermal state [22]

\[
\hat{\varrho} = D(\alpha)S(\chi)\hat{\varrho}_{\text{th}}(\mu_{\text{th}})S^\dagger(\chi)D^\dagger(\alpha)
\]  

with displacement operator \( D(\alpha) = \exp(\alpha \hat{a}^\dagger - \alpha^* \hat{a}) \) and squeezing operator \( S(\chi) = \exp(\chi \hat{a}^\dagger^2 - \chi^* \hat{a}^2)/2 \). Here, \( \chi = r e^{i \vartheta} \) is the squeezing parameter and

\[
\hat{\varrho}_{\text{th}}(\mu_{\text{th}}) = \sum_{m=0}^{\infty} \frac{\mu_{\text{th}}^m}{(1 + \mu_{\text{th}})^{m+1}} |m\rangle \langle m|
\]  

is a thermal state with mean photon number \( \mu_{\text{th}} \) [22]. The covariance matrix \( \Gamma \) for the state from eq. (56) is given by [22]

\[
(1 + 2\mu_{\text{th}}) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cosh 2r + \begin{bmatrix} \cos \vartheta & \sin \vartheta \\ \sin \vartheta & -\cos \vartheta \end{bmatrix} \sinh 2r
\]  

and the displacement vector is \( \mathbf{d} = \sqrt{2}[\text{Re}(\alpha), \text{Im}(\alpha)]^T \).

By using eqs. (37) and (58), we derive an analytic expression for the generating function:

\[
G(0,0,w) = \frac{1}{\sqrt{D}} \exp\left(\frac{wE}{2D}\right).
\]  

Here, we use the abbreviations

\[
E = |\alpha|^2[w - 2 - w(1 + 2\mu_{\text{th}}) \cosh 2r] + w(1 + 2\mu_{\text{th}}) \text{Re}(\alpha^2 e^{-i\vartheta}) \sinh 2r
\]  

and

\[
D = 1 - w + w^2 \left(\frac{1}{2} + \mu_{\text{th}} + \mu_{\text{th}}^2\right) - (1 + 2\mu_{\text{th}}) \left(\frac{w^2}{2} - w\right) \cosh 2r.
\]  

An explicit expression for the PND can be derived from this generating function, but we do not present it here.
where it was desired to take into account the photon statistics, but we do not resolve the frequency spectrum of the SPDC process. Because the spectrum and the photon statistics are related to each other, we recap basic properties of the multi-mode SPDC photon statistics.

SPDC with one signal mode $s$ and one idler mode $i$ generates the two-mode squeezed vacuum (TMSV) state $|\psi\rangle_{\text{TMSV}} = \exp(r e^{i\theta} \hat{a}_s^\dagger \hat{a}_i^\dagger - r e^{-i\theta} \hat{a}_s \hat{a}_i) |0\rangle$ with mean photon number $\mu = \sinh^2 r$ which can be written as [20]

$$|\psi\rangle_{\text{TMSV}} = \text{sech}(r) \sum_{n=0}^{\infty} (\text{tanh} r)^n |n_s\rangle |n_i\rangle .$$

The photon statistics of either the signal or idler mode alone is therefore given by $p(n) = \text{sech}^2(r) \tanh 2n r$, i.e. resembles that of a thermal state (cf. eq. (57)), with the PGF $h(y) = \text{sech}^2(r)/(1 - y \tanh^2 r)$. In the covariance matrix formalism, TMSV is represented by $d = 0$ and

$$\Gamma_{\text{TMSV}} = \left( \begin{array}{cc} \Gamma_+ & \Gamma_{\text{od}} \\ \Gamma_{\text{od}}^\dagger & \Gamma_- \end{array} \right)$$

with

$$\Gamma_+ = \left( \begin{array}{cc} 0 & \sinh(2r) \sin \vartheta \\ \sinh(2r) \sin \vartheta & 0 \end{array} \right) \quad \text{and} \quad \Gamma_- = \left( \begin{array}{cc} \cosh 2r & \pm \sinh(2r) \cos \vartheta \\ \pm \sinh(2r) \cos \vartheta & \cosh 2r \end{array} \right).$$

In practice, the bandwidths of the parametric process and the pump light result in two-mode squeezing between a number $M$ of pairs of orthogonal Schmidt modes determined by the Schmidt decomposition of the joint spectral amplitude of signal and idler photons [28, 64]. The resulting state is produced by applying $M$ independent two-mode squeezers to vacuum, each with thermal statistics of photon pairs. Hence, the PGF for the number of photon pairs is given by [64]

$$h(y) = \prod_{i=1}^{M} \frac{\text{sech}^2 r_i}{1 - y \tanh^2 r_i}$$

with $r_i e^{i\theta_i} = C \sqrt{\lambda_i}/(i \hbar)$. The constant $C$ is determined by the properties of the nonlinear medium and the pump light. For infinitely many equally strong squeezers, the generating function becomes the PGF of a Poisson distribution [64]. Thus, for an increasing number of squeezers, the PND changes from thermal to Poissonian statistics [64] as shown in fig. 2.

As a further example for our simulation method, we consider the bivariate photon statistics of SPDC while taking into account noise and detection efficiencies. First, we calculate the two-dimensional joint distributions $p(n_A, n_B)$ for different parameters of the detection efficiencies and noise as well as the cumulative distribution. The results obtained via our approach using the covariance matrix are shown in fig. 3. It can be

---

**FIG. 1.** Effect of Poissonian noise $\nu$ and detection efficiency $\eta$ on the detected photon number distribution of a displaced squeezed state $D(|\alpha| \ e^{i\vartheta})S(r \ e^{i\varphi}) |0\rangle$ with $|\alpha|^2 = 1.2$, $\varphi = 50^\circ$, $\vartheta = 30^\circ$, and $r \approx 1.287$ chosen such that $\langle N \rangle = 4$ by $\langle N \rangle = |\alpha|^2 + \sinh^2(r)$ [63]. The upper diagram shows that results derived by AD of the generating function from eq. (59) are in agreement with the analytic distribution from eq. (62). The distributions for noise $\nu > 0$ as well as the bars with efficiency $\eta$ are calculated from eq. (39). For $\eta = 1$, the lower diagram shows that the analytic distribution is in agreement with the results obtained by AD of eq. (39).

The PND has been presented in Ref. 47 where it was derived by using a different approach.

As a simplification we now consider a single-mode displaced squeezed state $|\psi\rangle = D(\alpha)S(\chi) |0\rangle$, i.e. with $\mu_{\text{th}} = 0$, for which the PND is given by [47]

$$p(n) = \exp(-|\alpha|^2 + \text{Re}(\alpha^2 e^{-i\vartheta} \tanh r) \frac{\text{tanh} n r}{n!^2 \cosh r}) \times \left| H_n \left( \alpha e^{-i\vartheta/2} \sinh(r) - \alpha^* e^{i\vartheta/2} \cosh(r) \right) \right|^2,$$

where $H_n$ is the $n$-th Hermite polynomial.

In fig. 1, we verify that deriving the PND of the state from the generating function eqs. (39) and (59) by AD yields the same results as the analytic expression from eq. (62). The analytic formula has the advantage that it can be quickly evaluated even for high photon numbers. But, an advantage of our formulation with a generating function in terms of $A$ and $d$ is that effects such as noise and detection efficiencies can easily be taken into account. We show the influence of both effects on the PND in fig. 1.

**B. Multi-mode spontaneous parametric down-conversion**

A strength of our approach is that the detection statistics of states with multiple modes entering the same detector can be easily evaluated. For example, the spectral degree of freedom can be incorporated into a simulation by discretizing the frequency spectrum and treating each frequency as a separate mode. For the simulation in section V we want to take into account the photon statistics, but we do not resolve the frequency spectrum of the SPDC process. Because the spectrum and the photon statistics are related to each other, we recap basic properties of the multi-mode SPDC photon statistics.
seen that as expected, decreasing the detection efficiency and adding noise blurs the line of delta functions that are observed for the ideal state along the diagonal $n_A = n_B$.

Second, we use eq. (41) to analyze mixed moments of the PND of a TMSV state in fig. 4. For comparison, analytic values for the mean values and the first mixed moment are shown. However, if the photon statistics are more complicated e.g. due to loss, simultaneous detection of multiple modes, or noise, analytic expressions for the moments may not be easy to find, but can be easily calculated by applying AD to the moment generating function eq. (41). In fig. 4, we show the influence of loss and noise on the moments.

V. DETECTION PROBABILITIES FOR ENTANGLEMENT-BASED QUANTUM KEY DISTRIBUTION

In this section, we consider an application to demonstrate that employing our method the photon statistics can be simulated in a relatively simple way, even for complex optical setups. We model a recent multi-user QKD system we have presented in Ref. 24. Although the system uses non-PNR detectors, the photon statistics are relevant as multi-photon pair emission is an important effect limiting the achievable quantum key rates. Detection efficiencies and noise are taken into account. Imperfect interference is modeled by using the mode mismatch model from Ref. 25 we have briefly mentioned in section II. All these effects are readily included with our simulation method.

FIG. 2. Probability distribution of photon pairs produced by SPDC with mean photon number $\mu = 4$ for different numbers of two-mode squeezers $M$ with equal strengths. The distribution changes from thermal statistics for a single two-mode squeezer to Poissonian statistics for infinitely many two-mode squeezers. The distributions were obtained by AD of eq. (67).

FIG. 3. Joint PND $p(n_A, n_B)$ for a state with a mean photon pair number $\mu = 3$ in 16 equally strong TMSV-modes with (a) detection efficiencies $\eta_A = \eta_B = 1$, (b) $\eta_A = 0.8$ and $\eta_B = 0.9$ and (c) efficiencies as in b and additional noise parameters $\nu_A = 1$ and $\nu_B = 2$, scaled by a factor of 4 for better visibility. (d) Cumulative PND $p(N_A = n_A, N_B \leq n_B)$ with noise and efficiencies as in b.

FIG. 4. Moments $M = \langle (\hat{N}_A - \langle \hat{N}_A \rangle)^k A (\hat{N}_B - \langle \hat{N}_B \rangle)^k B \rangle$ of the TMSV photon statistics with mean photon pair number $\mu = 4$ calculated via eq. (41) in logarithmic scale. raw: raw moments; central: central moments; analytic: analytic expressions for the mean values $\langle \hat{N}_A \rangle = \langle \hat{N}_B \rangle = \sinh^2 r$ and for the first mixed moment $\langle (\hat{N}_A - \langle \hat{N}_A \rangle)(\hat{N}_B - \langle \hat{N}_B \rangle) \rangle = \sinh^2(2r)/4$ [63]; imperfect: central moments for an imperfect detector with $\eta_A = 0.5$, $\eta_B = 0.3$, $\nu_A = 1$ and $\nu_B = 2$. 
A. Simulated setup for quantum key distribution

QKD enables the distribution of secure keys between users based on information-theoretic principles from quantum physics [65–67]. We model our entanglement-based QKD system for four users named Alice (A), Bob (B), Charlie (C), and Diana (D) [24]. It implements the BBM92 time-bin protocol [68–73]. Figure 5(a) shows the setup for the users Alice and Bob. The setup for Charlie and Diana is the same, but with different values for parameters such as insertion losses and coupling ratios and with different detector properties.

The central component of the QKD system is a source of time-bin entangled photon pairs. Laser pulses are sent through an imbalanced interferometer, resulting in double pulses with a time- and phase difference given by the interferometer’s path length imbalance. The double pulses produce photon pairs by type-0 SPDC in a nonlinear crystal. The mean photon pair number per double pulse is \( \mu \approx 0.034 \). The spectral width of the pump is much narrower than the photon frequency spectrum so that the photons are frequency-correlated: for a pump frequency \( 2\omega_0 \), energy conservation requires that the frequencies of the signal and idler photon sum up to \( \omega_s + \omega_i = 2\omega_0 \). The photons are separated by wavelength demultiplexing (WDM). The phase matching function is almost constant over several dozen nanometers so that photons are produced in multiple pairs of 100 GHz-wide WDM channels. A pair of users obtain entangled photons only if their wavelength channels are symmetrically arranged around the center frequency \( \omega_0 \). Multiple user pairs can thereby exchange quantum keys simultaneously and independently and the user combinations can be reconfigured by changing the WDM channel assignments.

In the receiver station of each user, the photons pass an interferometer with a path length difference matched to that of the source’s interferometer. They are therefore detected in one of three different time bins, depending on the path combination of the laser pulse and the photons in the interferometers. For detections in the early or late time bin, each user notes down a key bit of 0 or 1, respectively. When one user detects the photon in the central time bin and the other user detects it in the early or late time bin, the results are discarded, which is called key sifting. Furthermore, they assign labels \( i, j \in \{0, 1\} \) to their detectors. When both photons are detected in the central time bin, two-photon Franson interference [74] leads to emission from correlated interferometer outputs. Assuming an ideal setup, the detection probability is then given by [70, 71]

\[
P_{A_i, B_j}(\alpha, \beta, \varphi) = \frac{1}{4} (1 + (-1)^{i+j} \cos(\alpha + \beta - \varphi)) \tag{68}
\]

For QKD, the interferometer phases \( \alpha \) of Alice, \( \beta \) of Bob, and \( \varphi \) of the source are aligned to \( \alpha + \beta - \varphi = 2\pi n \) with \( n \in \mathbb{Z} \). Thereby, Alice and Bob always obtain the same bit values from the detector label.

However, various imperfections of the real setup lead to bit errors. For example, the interferometer delays may not be perfectly matched, the detectors have a non-zero dark count probability and the beam splitters are not ideal 50/50 splitters. The ratio of the number of bit errors to the number of measured bits is called quantum bit error rate (QBER). The higher the QBER, the lower the secure key rate that can be extracted. A simulation of the system can help to quantify the contribution of the different effects to the QBER and to optimize the setup. A higher value of \( \mu \), for example, yields a higher emission probability for photon pairs per laser pulse, resulting in a higher bit rate. However, due to the statistical nature of SPDC, increasing \( \mu \) increases the probability of multi-pair emission and thereby leads to a higher QBER. Therefore, a simulation inevitably needs to take into account the effect of multi-pair emission.

For the simulation, we treat the three time bins as separate modes. The setup from fig. 5(a) can thus be unfolded as depicted in fig. 5(b). The source interferometer is modeled by initializing two TMSV states with covariance matrices \( \Gamma(\chi_S) \) and \( \Gamma(\chi_L) \), where the beam splitter asymmetries and losses in the arms of the source interferometer are absorbed into the squeezing parameters

\[
\chi_S = r_0[T_{P1}(1 - L_{PS})]^{1/2} \quad \text{and} \quad \chi_L = r_0 e^{i\varphi}[(1 - T_{P1})(1 - L_{PL})(1 - T_{P1})]^{1/2}. \tag{69}
\]

The coefficient \( r_0 \) is chosen such that both pump pulses together have a certain mean photon pair number \( \mu \).

In fig. 2, we have shown the effect of the number of two-mode squeezers on the photon statistics. To obtain results representing the correct photon statistics, we would have to replicate the setup from fig. 5(b) in the simulation for the correct number of two-mode squeezers, each with its individual mean photon pair number. The Schmidt number \( K \) can be used to estimate the number of two-mode squeezers based on the shape of the joint spectral amplitude of the SPDC process [64, 75, 76]. For the simulation, we roughly estimate that \( K \) is in the order of magnitude of 100 and we assume equally strong two-mode squeezers resulting in almost Poissonian photon pair statistics. As the beam splitters and losses are applied to all Schmidt modes, the covariance matrix would be block-diagonal with \( K \) identical blocks. Instead of directly calculating the determinant of this large matrix, we apply the rule for block determinants and raise the determinant of one block to the power of \( K = 100 \).

In the setup, we use single-photon avalanche detectors with a dead time of \( \tau_{\text{dead}} = 10 \mu s \), a set efficiency of \( \eta = 20\% \) and dark count rates in the range from 300 cps to more than 3000 cps depending on the detector. In Ref. 77, we have presented a thorough analysis of these detectors. We incorporate efficiencies, dead times, afterpulses, and dark counts into the model. In appendix C we describe in detail how we construct the detection operators including these effects.
FIG. 5. (a) Schematic setup of our BBM92 quantum key distribution system from Ref. 24 consisting of a central entangled photon pair source and receiver modules for two users Alice and Bob each comprising an imbalanced interferometer. Note that here Mach-Zehnder interferometers are shown for clarity, but in practice, Michelson interferometers with Faraday mirrors are used instead [24]. Losses for all connections are modeled by introducing beam splitters coupling a fraction of the signal to an auxiliary vacuum mode and tracing this mode out (cf. eq. (16)). (b) Unfolded setup for the simulation. The SPDC pumped with double pulses from the pump interferometer is modeled by setting up a pair of two-mode squeezed states with different squeezing parameters $\chi_S$ and $\chi_L$. The different time bins are modeled as separate modes and for the early, central, and late times dedicated detectors are introduced. Imperfect interference at the beam splitters is taken into account by using the mode mismatch model from Ref. 25 we briefly reviewed in section II.

B. Comparison of simulated key rates and error rates to measurements

For the simulation, we have measured the values of all relevant experimental parameters, i.e. for the beam splitter transmissions, propagation losses in the fiber links, insertion losses of the components, and interferometer mode mismatches. Then, we have used these values to determine the parameters of the simulation model in fig. 5(b). For the detection efficiencies, dark count rates, and afterpulse probabilities of our detectors we have used values obtained from tomographic measurements which we have presented in Ref. 77. We have included these values in the detection operators as described in appendix C.

We have simulated the sifted key rates and quantum bit error rates by multiplying the source repetition rate with the detection probabilities for key bits and quantum bit errors in the time basis and phase basis. Note that the detection probabilities depend on the repetition rate due to the detector dead time and afterpulses (cf. appendix C). We compare the results with the measurements we have reported in Ref. 24 to check the validity of our simulation. For these experiments, the mean photon pair number had been set to $\mu = 0.034$. For the comparison, we consider the time basis QBER instead of the overall QBER because in the experiment the interferometer phases fluctuate and they are readjusted by tuning the interferometer temperatures as to minimize the phase basis QBER. Therefore, the phase basis QBER is not necessarily always at the minimum. Although it is possible to adjust the phases in the simulation to mimic phase fluctuations, we have chosen the phases for optimal interference and compare only the time QBER, which is unaffected by the phase fluctuations.

A comparison between the measured and simulated sifted key rates and QBERs in the time basis is shown in fig. 6. The simulation represents both the dependence of the key rates and of the time basis QBER on the user combination and distance very well. The key rate decreases with increasing distance between the participants due to the increasing propagation losses. But, the time QBER is almost independent of the transmission distance and transmission losses. The QBER variations can be attributed to different dark count rates of the detectors and to different values for the mode mismatch impairing the two-photon interference in different user combinations. Multiple effects lead to quantum bit errors in the time basis: uncorrelated clicks from two noise counts can lead to coincidence. Or, one photon from a pair may be lost and the other photon can be detected in coincidence with a noise count or a photon from a different pair. Investigating the different contributions helps to determine the effects limiting the secure key rate and optimizing the setup. PNR simulations enable a detailed analysis of the different effects contributing to the QBER, as we demonstrate in the next section.
When each of the two pump pulses produces a photon easily accessible by measurements. For example, the state with a specific photon number? enables such retrodictive analysis of the quantum state, Bayes’ theorem states

\[ p(E_1|E_2) = \frac{p(E_2|E_1)p(E_1)}{p(E_2)} \]  

The ability to simulate the photon number statistics enables such retrodictive analysis of the quantum state, i.e. it enables us to answer the question: Given that a certain measurement result was observed, what is the probability that it was produced by an initial quantum state with a specific photon number?

PNR simulations with GSs as presented above can be used to obtain \( p(E_1|E_2) \) when the probabilities \( p(E_1) \) and \( p(E_2) \) can be obtained via the simulation and \( p(E_2|E_1) \) is also accessible. Such retrodictive analysis enables the investigation of effects that may not be easily accessible by measurements. For example, the influence of multi-photon pair emission on the QBER can be compared to the contribution from noise counts.

When each of the two pump pulses produces a photon pair, Alice can detect a photon of the first pair in the early time bin and Bob can detect a photon of the second pair in the late time bin, causing a quantum bit error in the time basis. Here, we only want to briefly illustrate the method and for simplicity, we neglect the dependency of the time bins in the same repetition cycle due to the dead time among each other. Furthermore, we only consider the contribution from up to one photon pair produced by each pump pulse as well as raw coincidences between the detectors \( A_0 \) and \( B_0 \) instead of all exclusive coincidences that would lead to error bits in the time basis.

Two photon pairs can only lead to an error in the time basis when one was produced by the first pump pulse, which happens with probability \( p(f) \), and the other one is produced by the second pulse, which happens with probability \( p(s) \). Both values are obtained by simulating the probability to obtain one photon pair directly after the SPDC process. From Bayes’ theorem in eq. (71), the probability can be calculated that when a coincidence between \( A_{0,E} \) and \( B_{0,L} \) is measured, actually \( f \) and \( s \) pairs were produced by the first and second pulse, respectively:

\[ p(f,s|A_{0,E}, B_{0,L}) = \frac{p(A_{0,E}, B_{0,L}|f,s)p(f,s)}{p(A_{0,E}, B_{0,L})} \]  

The emission events are independent from each other, so that \( p(f,s) = p(f)p(s) \). The detections are independent as well, so that in eq. (72) we can factorize \( p(A_{0,E}, B_{0,L}|f,s) = p(A_{0,E}|f)p(B_{0,L}|s) \). Here, \( p(A_{0,E}|f) \) is the probability that detector \( A_0 \) is triggered in the early time bin, given that \( f \) pairs have been produced in the first pump pulse. For \( f = 1 \) it can be simply obtained by tracing the path of the photon through the setup to calculate the total photon transmission probability \( T_{A_{0,E}} \) to the detector including the beam splitter transmissions, propagation losses, and the detector efficiency. The value for \( p(B_{0,L}|s) \) is similarly obtained:

\[ p(A_{0,E}|1) = p_{on,A_0,E}(1 - e^{-\mu_{A_0,E}} (1 - T_{A_{0,E}})) \]  

\[ p(B_{0,L}|1) = p_{on,B_0,L}(1 - e^{-\nu_{B_0,L}} (1 - T_{B_{0,L}})) \]  

For \( f = 0 \) and \( s = 0 \), the detection probability is given by the probability for a noise count, which is equivalent to setting \( T = 0 \) in eqs. (73) and (74). Finally, the raw time error coincidence probability \( p(A_{0,E}, B_{0,L}) \) is obtained from the simulation as well. Thereby, all probabilities can be calculated that are required to obtain \( p(f,s|A_{0,E}, B_{0,L}) \) from eq. (72). In fig. 7 we show \( p(f,s|A_{0,E}, B_{0,L}) \) as a function of the produced total mean number of photon pairs per repetition cycle \( \mu \) and for different combinations of \( f \) and \( s \).

For small values of \( \mu \) below \( 1 \times 10^{-3} \), the largest contribution to the time basis error coincidence probability \( p(A_{0,E}, B_{0,L}) \) comes from coincidences between dark counts. Coincidences involving one noise count and one photon pair are relevant for a wide range of \( \mu \)-values from \( \mu < 1 \times 10^{-3} \) to \( \mu > 1 \). The difference between the cases \( f, s = 1, 0 \) and \( f, s = 0, 1 \) is mainly due to different transmission losses from the fiber lengths of 26.9 km to Alice.

![Graph showing sifted key rates and quantum bit error rates](image-url)
and 50.0 km to Bob and due to the much higher dark count probabilities for $A_0$ of 3045 cps compared to $B_0$ with 606 cps. The complementary probability in fig. 7 shows that in the range around $\mu \approx 0.1$ and above, effects from more than one photon pair become relevant. In the range of $\mu = 0.034$ where the QKD setup is operated, a majority of 53% of the $A_0, B_{0, L}$-coincidences occur when each pump pulse produces one photon pair. This underpins the fact that a simulation of the QKD system should take into account effects from multiple photon pairs to produce accurate results for the QBERs.

VI. DISCUSSION

In the previous sections, we have shown the versatile applications of our expressions for the calculation of the photon statistics of GSs. To the best of our knowledge, we are the first to derive expressions for cumulative probabilities, moments and factorial moments for general multi-mode GSs. For the PND $p(n)$, alternative calculation methods can be found in the literature. One approach requires $2n$ derivatives to evaluate the photon number $n$. For that, $\Gamma$ and $d$ are expressed in the complex $(\hat{a}, \hat{a}^\dagger)$-basis instead of the real $(x, p)$-basis via

$$\sigma_Q = \Omega^\dagger \Gamma \Omega / 2 + \mathbb{I} / 2 \text{ and } \tau = \Omega^\dagger d.$$  

In our notation, the expression reads [5, 14]

$$p(n) = \frac{1}{n! \sqrt{\det \sigma_Q}} \exp\left(\frac{1}{2} \tau^\dagger \sigma_Q^{-1} \tau\right) \times \prod_i \left(\frac{\partial^2}{\partial \alpha_i \partial \alpha_i^\dagger}\right)^n \exp\left(\frac{1}{2} \gamma^\dagger A \gamma + \tau^\dagger \sigma_Q^{-1} \gamma\right) \tag{75}$$

with $\gamma^\dagger = (\alpha^\dagger, \alpha^\dagger)$ and $A = \left(\begin{array}{cc}0 & \mathbb{I}_N \\ \mathbb{I}_N & 0 \end{array}\right) (\mathbb{I}_{2N} - \sigma_Q)$. The exponential can be differentiated has been recognized as the generating function for multivariate Hermite polynomials and therefore, the PND can be expressed in terms of these polynomials [84–86]. However, these polynomials are given by complex recursion formulas and the expressions are therefore considered difficult to evaluate for higher numbers of photons [87–89].

For GSs with $d = 0$, eq. (75) can be rewritten in terms of the Hafnian function as

$$p(n) = \text{haf}(A_S) / (n! \sqrt{\det \sigma_Q}), \tag{76}$$

where $A_S$ is derived from $A$ by repeating rows and columns, depending on the number of photons to be detected in a particular mode [5, 13, 14, 16]. For states with $d \neq 0$, a similar expression involves the loop Hafnian function [15, 16, 18, 38]. For non-PNR detection, the probabilities involve the Torontonian and loop Torontonian function [13, 19]. The software library The Walrus [90] provides software related to Gaussian Boson sampling (GBS) and has algorithms implemented to evaluate these functions. In contrast to such highly specialized algorithms, we use the general-purpose tool of automatic differentiation for our computations.

Generating operators for the photon statistics are treated in textbooks such as Refs. [31, 37]. In Refs. [64, 75], Mauerer et al. have applied automatic differentiation to the PGF from eq. (67) in order to obtain the PND resulting from multiple two-mode squeezers. In Ref. 91, a PGF for TMSV states in terms of the covariance matrix has been presented. However, generating functions in combination with automatic differentiation, have rarely been applied for practical calculations in the context of photon statistics. A possible reason is that the evaluation of higher-order derivatives is in general a resource-intensive computational task so that analytical expressions are preferred. But, for the calculation of the photon statistics, this is not an a-priori disadvantage, as the required computational resources for GBS computations scale exponentially even with optimized state-of-the-art algorithms [16, 17]. It can nevertheless be expected that specially tailored algorithms yield performance benefits becoming relevant for higher photon numbers.

Often only the lowest few moments of probability distributions, such as to order 3 or 4 are considered, which are easily evaluated with our method. For the PND, we have been able to calculate photon numbers up to 10 to 12 without any optimizations. The simulations took only a
few minutes but already required several GB of RAM. Ultimately, the 16 GB of RAM of our computer limited the maximum photon number we were able to simulate. By tuning the parameters of PyTorch, by using different AD software or different hardware, even higher numbers of photons can be simulated. In the field of GBS complexity research, PNDs for considerably higher numbers of photons have been computed on much more powerful computers. For example, GBS probabilities have been computed on a workstation with 96 CPUs for 50 photons in Ref. 15 or for up to 92 photons on a 100,000-core supercomputer in Ref. 17.

However, quantum optics applications often consider states with low mean photon numbers, for example when coincidences are considered. An example is entanglement-based QKD where $\mu \ll 1$ is required to avoid quantum bit errors from coincidences between photons from different pairs. The system we have modeled in section V is operated with a mean photon pair number per pulse around $\mu = 0.034$ (cf. Ref. 24). For such low mean photon numbers, only the first few photon numbers are of practical relevance as the probabilities for higher photon numbers decrease rapidly with increasing photon numbers. Thus, for many applications, a computational limitation to low photon numbers is not a significant restriction.

Another method to calculate the PND of GSs has recently been presented in Ref. 28, where one of the major results written in our notation is the expression

$$p(n) = \frac{(-1)^n}{n!} \partial^n_y \det \left( \frac{1 + \frac{Y(I - \Pi)Y}{2}}{Y} \right) \bigg|_{y = 1}$$

(77)

with $Y = \text{diag}(w) \oplus \text{diag}(w)$ and $w_s = \sqrt{\gamma_s}$. The authors of Ref. 28 derive it by applying a fanning-out transformation to infinitely many virtual non-PNR detectors, motivated by the intuition that in this limiting case, the probability that any of these virtual detectors receives more than one photon vanishes. They do not interpret their expression as a PGF. The expression eq. (77) from Ref. 28 can be rewritten as a special case of our eq. (39), which generalizes the result from Ref. 28 in multiple ways: it already incorporates detection efficiencies and noise counts and it can be applied to states with non-zero displacement vector $d$. Besides eq. (39), our derivation yields generating functions that enable the calculation of the moments, cumulative probabilities, density matrix elements, and factorial moments of the PND via eqs. (40) to (43) and eqs. (45) to (47). Furthermore, our expressions require, in contrast to eq. (75), only one derivative per photon number, which facilitates the numerical evaluation. While analytic expressions for the PND of single-mode photon-added and photon-subtracted GSs have been reported in Ref. 45–47, we have presented the more general formulas eqs. (53) and (54) which cover the multi-mode case and which provide, besides the PND, also their cumulative probabilities, moments, and factorial moments.

Conceptually, our generating-function approach provides a new point of view on photon statistics that is different from the Hafnian approach closely related to graph theory, and also different from the approach using the fanning-out transformation. Therefore, it can be a starting point for further theoretical investigation. For example, the analytic generating functions could also be analyzed for the complexity of their computation. Another example is the analytical generating function we have presented for the single-mode displaced squeezed thermal state.

We have shown that the flexibility of our approach enables us to easily incorporate imperfections such as noise of arbitrary statistics, provided its generating functions are known. In Ref. 28 it is noted that in comparison to other formulas for PNR detection, such as eq. (75) or the formulation via the Hafnian, an advantage of eq. (77) is that modes entering the same detector are treated by the same derivative, whereas the alternative formulas treat all modes separately so that terms for all the possible distribution patterns of the photons over the modes in a detector have to be calculated separately. In Ref. 28, this fact has been used to incorporate the spectral modes into calculations of the Hong-Ou-Mandel visibility of SPDC sources. The same argument holds for our eqs. (39) to (43), which involve one differentiation variable per detector, independent of the number of modes entering it. This is because our formulas obey the usual rules for generating functions for probabilities and moments and, therefore, the convolution of probability distributions is simply represented by the multiplication of the generating functions.

VII. CONCLUSIONS

We have presented various generating functions for the photon statistics of multi-mode GSs in terms of the covariance matrix and displacement vector, from which the probabilities and moments are retrieved by repeated automatic differentiation.

An advantage of our expression for the photon number distribution is that common effects in experiments, such as noise, non-unity detection efficiency, and the joint detection of multiple modes can be simply taken into account. Furthermore, our expressions enable calculating the cumulative detection probabilities, raw and central moments, and the rising and falling factorial moments of the photon number distribution. We derive these expressions for multi-mode GSs and also present closely related expressions for the matrix elements in the photon number basis and the coherent state basis. Furthermore, expressions for the generating functions of multi-mode photon-added and photon-subtracted GSs are derived, relating them to the generating functions of the underlying GSs.

We have implemented automatic differentiation by using the machine learning framework PyTorch in order to retrieve probabilities for photon numbers up to val-
ues of 12 on a desktop computer effortlessly, without any optimization for performance. The strength of this approach lies in its versatility and in the simplicity of the implementation, for which one may choose from a variety of available AD tools. The expressions may be evaluated up to higher photon numbers by tuning the parameters of the AD software, by employing different AD software or different hardware.

Our approach allows us to calculate various aspects of the photon number distribution and to incorporate many relevant details of the experimental setup easily. We have demonstrated the method with multiple examples. As an application, we have modeled an entanglement-based system for QKD which we have recently developed and we have observed very good agreement between the simulated and measured key rates and error rates. For this application, we have shown how the detection operators can be modified to include further experimental effects such as afterpulses, dead times, and the detection in different time bins that are relevant for modeling our QKD setup realistically. We have shown that simulated sifted key rates and quantum bit error rates well match the measured values. Finally, we have given an example to show how our simulation of the setup can be used to analyze the fraction of quantum bit errors due to the production of two photon pairs in the same repetition cycle. All these examples underpin the versatility of our simulation method.

Our expressions for the generating functions in eqs. (39) to (43) are formulated in terms of simple operations on the covariance matrix and displacement vector of GSs. Together with AD for the differentiation of the generating functions, this enables practical, easy-to-implement, versatile simulations of quantum-optical setups. We have shown that the method is not limited to GSs, but can be applied to certain classes of non-Gaussian states as well, thus greatly extending the range of its possible applications. Therefore, we expect that with increasing progress and the growing number of applications of photonic quantum technology, our method will become a useful tool for PNR simulations.

The data that support the findings of this study are available from the corresponding author upon request.

ACKNOWLEDGEMENTS

This research has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), under Grant No. SFB 1119–236615297. We thank Philipp Kleinpaß for fruitful discussions about the simulation of the QKD system. We also thank Julian K. Nauth, Vladimir M. Stojanović and Gernot Alber for valuable comments.

Appendix A: Generating functions for probabilities and moments

The probability distribution of a non-negative integer-valued random variable $N$ can be encoded by a probability generating function (PGF) of a parameter $y$ as \[ h(y) = \langle y^N \rangle = \sum_{n=0}^{\infty} p(N = n) y^n \] (A1)
so that the probabilities $p(n)$ and the mean $\langle N \rangle$ can be retrieved from the PGF by
\[ p(n) = \frac{1}{n!} \left. \frac{d^n}{dy^n} h(y) \right|_{y=0} \] and \[ \langle N \rangle = \sum_{n=0}^{\infty} p(N = n)n = \left. \frac{d}{dy} h(y) \right|_{y=1}. \] (A2)

The series in eq. (A1) converges at least for $y$ on the unit disk because all $p(n)$ are non-negative and $\sum_n p(n) = 1$.

The cumulative probabilities $p(N \leq n)$ are encoded by a closely related generating function \[ \frac{h(y)}{1-y} = \sum_{n,k=0}^{\infty} p(N = n) y^{n+k}, \] \[ p(N \leq n) = \sum_{k=0}^{n} p(k) = \frac{1}{n!} \left. \frac{d^n}{dy^n} h(y) \right|_{y=0}. \] (A5)

The moments $M(k, \mu) = \langle (N-\mu)^k \rangle$ of the probability distribution can be encoded in a moment generating function $M(\mu, y) = \langle \exp[y(N-\mu)] \rangle$ \[ M(\mu, y) = \sum_{n=0}^{\infty} p(n) e^{y(n-\mu)} = \sum_{k=0}^{\infty} \frac{y^k}{k!} ((N-\mu)^k), \] (A6)

\[ M(k, \mu) = \left. \frac{d^k}{dy^k} M_\mu(y) \right|_{y=0}. \] (A7)

For $\mu = 0$ or $\mu = \langle N \rangle$, $M(\mu, y)$ generates the raw moments $\langle N^k \rangle$ or the central moments $\langle (N-\langle N \rangle)^k \rangle$.

It is sometimes convenient to study factorial moments [31, 34, 37, 39, 82]. The falling factorial moments \[ n_{(k)} = \langle N_{(k)} \rangle = \langle (N(N-1)\cdots(N-k+1) \rangle \] are generated\(^5\) by $((1+y)^N)$:
\[ h(1+y) = \sum_{n=0}^{\infty} p(n) \sum_{k=0}^{n} y^k \frac{n!}{k!} = \sum_{k=0}^{\infty} \frac{n_{(k)} y^k}{k!}, \] (A8)
\[ n_{(k)} = \left. \frac{d^k}{dy^k} h(1+y) \right|_{y=0}. \] (A9)

\(^5\) In eqs. (A8) and (A10) we use [93]
\[ (1+y)^n = \sum_{k=0}^{n} \frac{n!}{k!} y^k \quad \text{and} \quad \frac{1}{(1-y)^{n+1}} = \sum_{k=0}^{\infty} \frac{(k+n)!}{k! n!} y^k. \]
Similarly, a generating function for the rising factorial moments \( n^{(k)} = \langle N^{(k)} \rangle = ((N + 1) \cdots (N + k)) \) can be defined [31, 39]:

\[
R(y) = \sum_{k=0}^{\infty} \frac{y^k}{k!} \sum_{n=0}^{\infty} p(n) \frac{(n + k)!}{n!} = \frac{1}{1 - y} \langle (1 - y)^{-N} \rangle ,
\]

\[
n^{(k)} = \frac{d^k}{dy^k} R(y) \bigg|_{y=0} . \tag{A10}
\]

In the derivation of the generating functions for multimode GSs, we use two important properties of such generating functions. First, multivariate probability distributions are encoded by multivariate generating functions. For example, the bivariate distribution \( p(n_1, n_2) \) is generated by \( h(y_1, y_2) = \sum_{n_1, n_2} p(n_1, n_2) y_1^{n_1} y_2^{n_2} \).

Second, the PGF or moment generating function for the sum of random variables is the product of the individual moments, sharing the same parameter \( y \):

\[
p(n) = \frac{1}{n!} \frac{d^n}{dy^n} \prod_m h_m(y) \bigg|_{y=0} . \tag{A12}
\]

Note that when multiplying the generating functions for cumulative probabilities and falling factorial moments, the power series in \( N \) is multiplied by the prefactor \((1 - y)^{-1}\) only once, not for each factor individually.

**Appendix B: Derivation of the generating operator for photon-added states**

To derive expressions for the photon statistics of photon-added GSs, we consider \( \text{tr}(\hat{g}^k \hat{a}^{*k}) \) with the abbreviation \( \hat{g} = \hat{g}(0, 0, w) = :\exp(-w\hat{N}):: \) and insert \( \mathbb{I} = \sum_m |m\rangle \langle m|\):

\[
\text{tr}(\hat{g}^k \hat{a}^{*k}) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \langle n|\hat{g}|m\rangle \langle m|\hat{a}^{*k} \hat{a}^{-k}|n\rangle
\]

\[
= \sum_{n=0}^{\infty} \langle n|\hat{g}|n\rangle \langle n + k| n + k \rangle \frac{(n + k)!}{n!} . \tag{B1}
\]

Evaluating the series expansion for \( \hat{g} \) yields

\[
\langle l|\hat{g}|l\rangle = \sum_{j=0}^{l} \frac{(-w)^j l!}{j!(l-j)!} = (1 - w)^l . \tag{B2}
\]

Now, we insert the expression as a coefficient into a generating function, apply the relation from footnote 5 and insert \( \langle n|\hat{g}|n\rangle = \langle \exp(-\hat{N})\hat{N}^n/n! \rangle \) from eq. (17):

\[
\sum_{k=0}^{\infty} \frac{y^k}{k!} \text{tr}(\hat{g}^k \hat{a}^{*k}) = \sum_{n,k=0}^{\infty} \langle n|\hat{g}|n\rangle r^k (1 - w)^{n+k} \frac{(n + k)!}{n!k!}
\]

\[
= \sum_{n=0}^{\infty} \frac{(n|\hat{g}|n)}{1 - r(1 - w)} (1 - w)^n + 1
\]

\[
= \frac{1}{1 - r(1 - w)} \langle \exp\left(\hat{N}(1 - w) - \hat{N}\right) \rangle . \tag{B3}
\]

Finally we obtain with \( \hat{w} = 1 - [(1 - w)^{-1} - r]^{-1} \)

\[
\text{tr}(\hat{g}^k \hat{a}^{*k}) = \frac{d^k}{dr^k} G(0, 0, \hat{w}) \bigg|_{r=0} . \tag{B4}
\]

**Appendix C: Model for photon detection in the QKD setup**

In general, a detector can be described by a positive operator-valued measure (POVM), i.e. a set of positive operators \( \{\hat{F}_i\} \) with \( \sum_i \hat{F}_i = \mathbb{1} \) so that the probability to obtain result \( i \) is \( p(i) = \langle \hat{F}_i \rangle \). Our detectors are non-PNR detectors, i.e. they can only yield two detection results: “click” and “no-click”. Due to the dead time, which is much longer than the time bins and the time separating them, the detector can not click in multiple time bins of the same repetition cycle. For our detectors, we have therefore four mutually exclusive results: Detection in the early (E), central (C), and late (L) time bin as well as no click (no). The probability for a detector click in a particular interval is the probability that non-zero clicks are registered, multiplied by the probability \( p_{\text{on}} \) that the detector is not in the dead time. Hence, for the POVM elements we obtain \( \hat{F} = \hat{F}_E + \hat{F}_C + \hat{F}_L + \hat{F}_{\text{no}} \) with

\[
\hat{F}_E = p_{\text{on}} \hat{I}_{\text{C.L}} (\hat{I}_E - \hat{N}_{E=0}) , \tag{C1}
\]

\[
\hat{F}_C = p_{\text{on}} \hat{I}_{\text{L}} \hat{N}_{E=0} (\hat{I}_C - \hat{N}_{C=0}) , \tag{C2}
\]

\[
\hat{F}_L = p_{\text{on}} \hat{N}_{E=0} (\hat{I}_L - \hat{N}_{L=0}) , \tag{C3}
\]

\[
\hat{F}_{\text{no}} = p_{\text{off}} \mathbb{I} + p_{\text{on}} \hat{N}_{E=0} . \tag{C4}
\]

Here, we abbreviate the index \( N_{E,C,L} = N \), indicating all time bins together. The probability that the detector is live, i.e. not in the dead time \( \tau_{\text{dead}} \) at a given point in time, is given by \( p_{\text{on}} = 1 - p_{\text{off}} \). The operators \( \hat{N}_{E=0} \) in eq. (C2) and \( \hat{N}_{E=0} \) in eq. (C3) take into account that due to the dead time, a click switches the detector off in the subsequent time bins of the same repetition cycle.

To include effects from afterpulses and dark counts occurring with the dark count rate \( r_{\text{dark}} \) and to calculate \( p_{\text{on}} \), we first consider the simple case of a detector showing some Poissonian noise \( \nu \) but exhibiting no dead
time. We define the noise rate $r_{\text{noise}}$, click rate $r$, and click probability $p_{\text{click}}$ by

$$r_{\text{noise}} = r_{\text{dark}} + p_{\text{ap}} r,$$  \hspace{1cm} (C5)

$$r = p_{\text{click}} f_{\text{rep}},$$  \hspace{1cm} (C6)

$$p_{\text{click}} = \langle 1 - \hat{N}_{N=0} \rangle = 1 - \frac{\exp(-\nu)}{\sqrt{\det A}}$$  \hspace{1cm} and (C7)

$$\nu = r_{\text{noise}} / f_{\text{rep}}.$$  \hspace{1cm} (C8)

Here, $f_{\text{rep}}$ is the repetition rate of the photon pair source, i.e. $\nu$ and $p_{\text{click}}$ are defined for a time interval of one repetition cycle. The afterpulse probability $p_{\text{ap}}$ is the probability that a click triggers a subsequent uncorrelated click. Here, only modes of the covariance matrix are kept in $A$ which enter the detector, while the columns and rows corresponding to all other modes are removed. This means that on one hand $r_{\text{noise}}$ depends on the click rate according to eqs. (C5) and (C6) and on the other hand the click rate also depends on the noise according via eqs. (C7) and (C8). A self-consistent value for $r_{\text{noise}}$ can be obtained by solving

$$r_{\text{noise}} = r_{\text{dark}} + p_{\text{ap}} f_{\text{rep}} \left( 1 - \frac{\exp(-r_{\text{noise}} / f_{\text{rep}})}{\sqrt{\det A}} \right).$$  \hspace{1cm} (C9)

for $r_{\text{noise}}$. Linearizing the exponential for $r_{\text{noise}} / f_{\text{rep}} \ll 1$ yields

$$r_{\text{noise}} \approx r_{\text{dark}} + p_{\text{ap}} f_{\text{rep}} \left( 1 - 1 / \sqrt{\det A} \right).$$  \hspace{1cm} (C10)

This approximation is justified because the values for $r_{\text{noise}} / f_{\text{rep}}$ are in the order of $1 \times 10^{-5}$ for our setup operated at values of $\nu \approx 0.034$. From eqs. (C6) and (C7), we obtain the click rate $r$ including afterpulses. The noise parameter for detection in a time bin of width $\Delta T = 1$ ns is then given by $r_{\text{noise}} / f_{\text{rep}} = r_{\text{noise}} \Delta T$.

Finally, we consider the effect of the dead time $r_{\text{dead}}$. Thus, we introduce the measured click rate $r_{m}$ blocking a fraction $p_{\text{off}} = r_{m} r_{\text{dead}}$ of the acquisition time. We can now calculate $p_{\text{on}} = 1 - r_{m} r_{\text{dead}}$. Furthermore, the measured click rate is simply the product $r_{m} = p_{\text{on}} r$ leading to

$$p_{\text{on}} = \frac{1}{1 + r r_{\text{dead}}},$$  \hspace{1cm} (C11)

so that $p_{\text{on}}$ can be directly calculated from the click rate.

For coincidences between detectors, we distinguish between raw and exclusive coincidences. By coincidences, we mean events where at least two detectors click in the same repetition cycle. While raw coincidences only consider that two detectors click, exclusive coincidences also require that the two other detectors do not click. For QKD, we use only those events for which an unambiguous bit value can be derived, i.e. only exclusive coincidences are considered. Note for the security of the key exchange it is recommended that the participant observing clicks in both detectors randomly assigns one of the values $\{0, 1\}$, but for the sake of simplicity of the simulation, we work with exclusive coincidences which exclude such double-clicks. The coincidence probabilities, for example between detectors $A_0$ in time bin $i$ and $B_0$ in time bin $j$ with $i, j \in \{E, C, L\}$, are given by

$$p_{A_0,i,B_0,j,\text{raw}} = \langle \hat{N}_{A_0,i} \hat{N}_{B_0,j} \rangle$$  \hspace{1cm} and (C12)

$$p_{A_0,i,B_0,j,\text{excl.}} = \langle \hat{N}_{A_0,i} \hat{N}_{B_0,j} \hat{N}_{A_1,\text{no}} \hat{N}_{B_1,\text{no}} \rangle.$$  \hspace{1cm} (C13)

As each detection operator consists of two terms (cf. eq. (C7)), expanding eq. (C13) yields 16 terms.

To simplify the comparison of the simulation with the measured key rates, we approximate eq. (C4):

$$\hat{N}_{\text{no}} = \hat{N}_{N=0} + p_{\text{off}} (\hat{1} - \hat{N}_{N=0}) \approx \hat{N}_{N=0}.$$  \hspace{1cm} (C14)

This approximation is valid when $p_{\text{off}} \ll 1$ or when $\hat{N}_{N=0} \approx \hat{1}$, which is the case because the click probability is low due to the detection efficiency of $\eta = 20\%$ and due to the transmission losses. Using this approximation reduces the number of terms from 16 to 4. As an example, consider exclusive coincidences between detectors $A_0$ in time bin $i$ and $B_0$ in time bin $j$ with $i, j \in \{E, C, L\}$. By using eq. (C14), the exclusive coincidence probability becomes

$$p_{\text{click}(A_0, E, B_0, L, \text{excl.})} = p_{A_0,0} p_{B_0,0} \left( \langle \hat{1} A_{0,E} - \hat{N}_{N=(A_0, E)=0} \rangle \times \hat{N}_{(B_{0,C}, E, A_{1,B})=0} - \hat{N}_{(B_{0,L})=0} \right) \approx p_{A_0,0} p_{B_0,0} \exp(-\nu_{B_{0,C}} - \nu_{A_1} - \nu_{B_1}) \times \left( \frac{1}{\sqrt{\det A_{B_{0,C}, E, A_{1,B}}} - \exp(-\nu_{A_{0,E}})} - \frac{\exp(-\nu_{B_{0,L}})}{\sqrt{\det A_{B_{0,E}, B_{0,C}, A_{1,B}}}} + \frac{\exp(-\nu_{A_{0,E}} - \nu_{B_{0,L}})}{\sqrt{\det A_{B_{0,E}, B_{0,C}, A_{1,B}}}} \right).$$  \hspace{1cm} (C15)

In the experiment, the photons have been separated by WDM with an arrayed-waveguide grating, for which we have measured the wavelength-dependent transmission $\tau(\omega)$ of Alice’s and Bob’s channels $C_A$ and $C_B$ with channel widths $\Delta$. In the simulation, we have included the averaged transmissions $\tau_{A/B} = \Delta^{-1} \int_{C_A/B} \tau(\omega) d\omega$ into the losses $L_A$ and $L_B$. Due to the photon frequency correlation, the average transmission probability for a photon pair is not $\tau_{A/B}$ but $\tau_{\text{pair}} = \Delta^{-1} \int_{C_A} \tau(2 \omega_0 - \omega) d\omega$. We, therefore, multiply the key rate by the correction factor $\tau_{\text{pair}} / (\tau_{A/B}) \approx 1.5$ to take the spectral dependence of the channel loss into account.
[1] C. Cahall, K. L. Nicolich, N. T. Islam, G. P. Lafyatis, A. J. Miller, D. J. Gauthier, and J. Kim, Multi-photon detection using a conventional superconducting nanowire single-photon detector, Optica 4, 1534 (2017).

[2] M. Schmidt, M. von Helversen, M. López, F. Gericke, E. Schlottmann, T. Heindel, S. Kück, S. Reitenstein, and J. Beyer, Photon-number-resolving transition-edge sensors for the metrology of quantum light sources, Journal of Low Temperature Physics 193, 1243 (2018).

[3] X. Tao, S. Chen, Y. Chen, L. Wang, X. Li, X. Tu, X. Jia, Q. Zhao, L. Zhang, L. Kang, and P. Wu, A high speed and high efficiency superconducting photon number resolving detector, Superconductor Science and Technology 32, 064002 (2019).

[4] G. He, H. Li, R. Yin, L. Zhang, D. Dong, J. Lv, Y. Fei, X. Wang, Q. Chen, F. Li, H. Li, H. Wang, T. Xu, Q. Zhao, X. Jia, J. Chen, L. Kang, and P. Wu, Simultaneous resolution of photon numbers and positions with series-connected superconducting nanowires, Applied Physics Letters 120, 124001 (2022).

[5] C. S. Hamilton, R. Kruse, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex, Gaussian boson sampling, Phys. Rev. Lett. 119, 170501 (2017).

[6] D. J. Brod, E. F. Galvão, A. Crespi, R. Osellame, N. Spagnolo, and F. Sciarrino, Photonic implementation of boson sampling: a review, Advanced Photonics 1, 1 (2019).

[7] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ying, X. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, Quantum computational advantage using photons, Science 370, 1460 (2020).

[8] H.-S. Zhong, Y.-H. Deng, J. Qin, H. Wang, M.-C. Chen, L.-C. Peng, Y.-H. Luo, D. Wu, S.-Q. Gong, H. Su, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, J. J. Renema, C.-Y. Lu, and J.-W. Pan, Phase-programmable gaussian boson sampling using stimulated squeezed light, Phys. Rev. Lett. 127, 180502 (2021).

[9] T. R. Bromley, J. M. Arrazola, S. Jahangiri, J. Izac, N. Quesada, A. D. Gran, M. Schulz, J. Swinarton, Z. Zabaneh, and N. Killoran, Applications of near-term photonic quantum computers: software and algorithms, Quantum Science and Technology 5, 034010 (2020).

[10] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, in Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing, STOC ’11 (Association for Computing Machinery, New York, NY, USA, 2011) p. 333–342.

[11] A. P. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. L. O’Brien, and T. C. Ralph, Boson sampling from a gaussian state, Phys. Rev. Lett. 113, 100502 (2014).

[12] S. Rahimi-Keshari, A. P. Lund, and T. C. Ralph, What can quantum optics say about computational complexity theory?, Phys. Rev. Lett. 114, 060501 (2015).

[13] N. Quesada, J. M. Arrazola, and N. Killoran, Gaussian boson sampling using threshold detectors, Phys. Rev. A 98, 062322 (2018).

[14] R. Kruse, C. S. Hamilton, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex, Detailed study of gaussian boson sampling, Phys. Rev. A 100, 032326 (2019).

[15] N. Quesada, R. S. Chadwick, B. A. Bell, J. M. Arrazola, T. Vincent, H. Qi, and R. García-Patrón, Quadratic speed-up for simulating gaussian boson sampling, PRX Quantum 3, 010306 (2022).

[16] A. Björklund, B. Gupt, and N. Quesada, A faster hafnian formula for complex matrices and its benchmarking on a supercomputer, ACM J. Exp. Algorithmics 24, 10.1145/3325111 (2019).

[17] J. F. F. Bulmer, B. A. Bell, R. S. Chadwick, A. E. Jones, D. Moise, A. Rigazzi, J. Thorbecke, U.-E. Haus, T. V. Vaenerbergh, R. B. Patel, I. A. Walmsley, and A. Laing, The boundary for quantum advantage in gaussian boson sampling, Science Advances 8, eabl9236 (2022).

[18] N. Quesada, Franck-condon factors by counting perfect matchings of graphs with loops, The Journal of Chemical Physics 150, 164113 (2019).

[19] J. F. F. Bulmer, S. Paesani, R. S. Chadwick, and N. Quesada, Threshold detection statistics of bosonic states (2022).

[20] X.-B. Wang, T. Hiroshima, A. Tomita, and M. Hayashi, Quantum information with gaussian states, Physics Reports 448, 1 (2007).

[21] 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).

[22] S. Olivares, Quantum optics in the phase space, The European Physical Journal Special Topics 203, 3 (2012).

[23] G. Adesso, S. Ragy, and A. R. Lee, Continuous variable quantum information: Gaussian states and beyond, Open Systems & Information Dynamics 21, 140001 (2014).

[24] E. Fitzke, L. Bialowons, T. Dolejsky, M. Tippmann, E. Schlottmann, T. Heindel, S. Kück, S. Reitzenstein, D. Moise, A. Rigazzi, J. Thorbecke, U.-E. Haus, T. V. Vaenerbergh, R. B. Patel, I. A. Walmsley, and A. Laing, Progress towards practical device-independent quantum key distribution with spontaneous parametric down-conversion based photonic quantum information processing, New Journal of Physics 17, 043030 (2015).

[25] M. Takeoka, R.-B. Jin, and M. Sasaki, Full analysis of multi-photon pair effects in spontaneous parametric down conversion based photonic quantum information processing, Phys. Rev. A 93, 042328 (2016).

[26] Y. Tsujimoto, K. Wakui, M. Fujiwara, K. Hayasaka, S. Miki, H. Terai, M. Sasaki, and M. Takeoka, Optimal conditions for the bell test using spontaneous parametric down-conversion sources, on-off photodetectors, and entanglement swapping, Phys. Rev. A 100, 063842 (2018).

[27] O. F. Thomas, W. McCutcheon, and D. P. S. McCutcheon, A general framework for multimode gaussian quantum optics and photo-detection: Application to hong-ou-mandel interference with filtered heralded single photon sources, APL Photonics 6, 040801 (2021).

[28] P. L. Kelley and W. H. Kleiner, Theory of electromagnetic field measurement and photoelectron counting, Phys. Rev. 136, A316 (1964).

[29] M. Lux and M. Zwanziger, Exact photocount statistics: Lasers near threshold, Phys. Rev. A 7, 750 (1973).
[31] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* - (Cambridge University Press, Cambridge, 1995).
[32] D. Walls and G. J. Milburn, *Quantum Optics*, 2nd ed. (Springer, 2008).
[33] J. Eisert and M. M. Wolf, Gaussian quantum channels (Imperial College Press, distributed by Wold Scientific Publishing Co., 2007) pp. 23–42, 0.
[34] G. Adam, Density matrix elements and moments for generalized gaussian state fields, Journal of Modern Optics 42, 1311 (1995).
[35] J. Perina, *Quantum Statistics of Linear and Nonlinear Optical Phenomena*, 2nd ed. (Springer Science & Business Media, 1991).
[36] W. P. Schleich, *Quantum Optics in Phase Space*, 1st ed. (Wiley, 2001).
[37] S. Barnett and P. M. Radmore, *Methods in Theoretical Quantum Optics*, Oxford Series in Optical and Imaging Sciences (Oxford University Press, 2002).
[38] N. Quesada, L. G. Helt, J. Izaac, J. M. Arrazola, R. Shahrokhshahi, C. R. Myers, and K. K. Sabapathy, Simulating realistic non-gaussian state preparation, Phys. Rev. A 100, 022341 (2019).
[39] S. M. Barnett, G. Fеренци, C. R. Gilson, and F. С. Спирис, Statistics of photon-subtracted and photon-added states, Phys. Rev. A 98, 013809 (2018).
[40] Y. I. Bogdanov, K. G. Katamasde, G. V. Avossiopoulos, L. V. Bellsny, N. A. Bogdanova, A. A. Kalinkin, and S. P. Kulik, Multiphoton subtracted thermal states: Description, preparation, and reconstruction, Phys. Rev. A 96, 063803 (2017).
[41] A. Zavatta, V. Parigi, M. S. Kim, and M. Bellini, Subtracting photons from arbitrary light fields: experimental test of coherent state invariance by single-photon annihilation, New Journal of Physics 10, 123006 (2008).
[42] A. Zavatta, S. Viciani, and M. Bellini, Quantum-to-classical transition with single-photon-added coherent states of light, Science 306, 660 (2004).
[43] V. Parigi, A. Zavatta, and M. Bellini, Implementation of single-photon creation and annihilation operators: experimental issues in their application to thermal states of light, Journal of Physics B: Atomic, Molecular and Optical Physics 42, 114005 (2009).
[44] D. Liu, M. Tian, S. Liu, X. Dong, J. Guo, Q. He, H. Xu, and Z. Li, Ghost imaging with non-gaussian quantum light, Phys. Rev. Applied 16, 064037 (2021).
[45] X. xiang Xu, L. yun Hu, and H. yi Fan, Photon-added squeezed thermal states: Statistical properties and its decoherence in a photon-loss channel, Optics Communications 283, 1801 (2010).
[46] Z. Wang, X. guo Meng, and H. yi Fan, Photon-subtracted squeezed coherent state: nonclassicality and decoherence in thermal environment, J. Opt. Soc. Am. B 29, 397 (2012).
[47] S. Wang, H. yi Fan, and L. yun Hu, Photon-number distributions of non-gaussian states generated by photon subtraction and addition, J. Opt. Soc. Am. B 29, 1020 (2012).
[48] 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, Photon-number-resolved detection of photon-subtracted thermal light, Opt. Lett. 38, 2171 (2013).
[49] D. Su, C. R. Myers, and K. K. Sabapathy, Conversion of gaussian states to non-gaussian states using photon-number-resolving detectors, Phys. Rev. A 100, 052301 (2019).
[50] C. N. Gagatsos and S. Guha, Efficient representation of gaussian states for multimode non-gaussian quantum state engineering via subraction of arbitrary number of photons, Phys. Rev. A 99, 053816 (2019).
[51] J. Abate and W. Whitt, Numerical inversion of probability generating functions, Operations Research Letters 12, 245 (1992).
[52] J. Abate and W. Whitt, The fourier-series method for inverting transforms of probability distributions, Queueing Systems 10, 5 (1992).
[53] G. L. Choudhury, D. M. Lucantoni, and W. Whitt, Multi-dimensional transform inversion with applications to the transient m/g/1 queue, The Annals of Applied Probability 4, 719 (1994).
[54] G. L. Choudhury and W. Whitt, Computing distributions and moments in polling models by numerical transform inversion, Performance Evaluation 25, 267 (1996).
[55] G. L. Choudhury and D. M. Lucantoni, Numerical computation of the moments of a probability distribution from its transform, Operations Research 44, 398 (1996).
[56] A. Griewank and A. Walther, *Evaluating Derivatives*, 2nd ed. (Society for Industrial and Applied Mathematics, 2008).
[57] U. Naumann, *The Art of Differentiating Computer Programs - An Introduction to Algorithmic Differentiation* (SIAM, Philadelphia, 2012).
[58] Community portal for automatic differentiation, https://www.autodiff.org (2022), accessed: 2022-07-09.
[59] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, Automatic differentiation in machine learning: a survey, Journal of Machine Learning Research 18, 1 (2018).
[60] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, Automatic differentiation in pytorch, in NIPS-W (2017).
[61] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, Pytorch: An imperative style, high-performance deep learning library, in *Advances in Neural Information Processing Systems 32*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019) pp. 8024–8035.
[62] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, *TensorFlow: Large-scale machine learning on heterogeneous systems* (2015), software available from tensorflow.org.
[63] C. Gerry and P. Knight, *Introductionary Quantum Optics* (Cambridge University Press, 2004).
[64] W. Mauerer, M. Avenhaus, W. Helwig, and C. Silberhorn, How colors influence numbers: Photon statistics of parametric down-conversion, Phys. Rev. A 80, 053815 (2009).
[65] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Quantum cryptography, Rev. Mod. Phys. 74, 145 (2002).
[66] 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).
[67] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, Secure quantum key distribution with realistic devices, Rev. Mod. Phys. 92, 025002 (2020).
[68] C. H. Bennett, G. Brassard, and N. D. Mermin, Quantum cryptography without bell’s theorem, Phys. Rev. Lett. 68, 557 (1992).
[69] J. Brendel, N. Gisin, W. Tittel, and H. Zbinden, Pulsed energy-time entangled twin-photon source for quantum communication, Phys. Rev. Lett. 82, 2594 (1999).
[70] W. Tittel, J. Brendel, H. Zbinden, and N. Gisin, Quantum cryptography using entangled photons in energy-time bell states, Phys. Rev. Lett. 84, 4737 (2000).
[71] I. Marcikic, H. de Riedmatten, W. Tittel, H. Zbinden, M. Legré, and N. Gisin, Distribution of time-bin entangled qubits over 50 km of optical fiber, Phys. Rev. Lett. 93, 180502 (2004).
[72] T. Honjo, S. W. Nam, H. Takesue, Q. Zhang, H. Kamada, Y. Nishida, O. Tadanaga, M. Asobe, B. Baek, R. Hadfield, S. Miki, M. Fujiwara, M. Sasaki, Z. Wang, K. Inoue, and Y. Yamamoto, Long-distance entanglement-based quantum key distribution over optical fiber, Phys. Rev. Lett. 103, 190502 (2009).
[73] H. Takesue, K. ichi Harada, K. Tamaki, H. Fukuda, T. Tsuchizawa, T. Watanabe, K. Yamada, and S. ichi Itabashi, Long-distance entanglement-based quantum key distribution experiment using practical detectors, Opt. Express 18, 16777 (2010).
[74] J. D. Franson, Bell inequality for position and time, Phys. Rev. Lett. 62, 2295 (1989).
[75] W. Mauerer, On Colours, Keys, and Correlations: Multimode Parametric Downconversion in the Photon Number Basis, doctoral thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) (2009).
[76] A. Christ, K. Laiho, A. Eckstein, K. N. Cassemiro, and C. Silberhorn, Probing multimode squeezing with correlation functions, New Journal of Physics 13, 033027 (2011).
[77] E. Fitzke, R. Krebs, T. Haase, M. Mengler, G. Alber, and T. Walther, Time-dependent POVM reconstruction for single-photon avalanche photo diodes using adaptive regularization, New Journal of Physics 10.1088/1367-2630/ac5004 (2022).
[78] D. T. Pegg and S. M. Barnett, Retrodiction in quantum optics, Journal of Optics B: Quantum and Semiclassical Optics 1, 442 (1999).
[79] S. M. Barnett, D. T. Pegg, and J. Jeffers, Bayes’ theorem and quantum retrodiction, Journal of Modern Optics 47, 1779 (2000).
[80] S. M. Barnett, Quantum retrodiction, in Quantum Information and Coherence, edited by E. Andersson and P. Öhberg (Springer International Publishing, Cham, 2014) pp. 1–30.
[81] S. M. Barnett, J. Jeffers, and D. T. Pegg, Quantum retrodiction: Foundations and controversies, Symmetry 13, 10.3390/sym13040586 (2021).
[82] S. M. Barnett, L. S. Phillips, and D. T. Pegg, Imperfect photodetection as projection onto mixed states, Optics Communications 158, 45 (1998).
[83] O. Jedrkiewicz, R. Loudon, and J. Jeffers, Retrodiction for optical attenuators, amplifiers, and detectors, Phys. Rev. A 70, 033805 (2004).
[84] V. V. Dodonov, O. V. Man’ko, and V. I. Man’ko, Multidimensional hermite polynomials and photon distribution for polyme mixed light, Phys. Rev. A 50, 813 (1994).
[85] P. Kok and S. L. Braunstein, Multi-dimensional hermite polynomials in quantum optics, Journal of Physics A: Mathematical and General 34, 6185 (2001).
[86] J. Huh, Multimode bogoliubov transformation and husimi’s q-function, Journal of Physics: Conference Series 1612, 012015 (2020).
[87] S. Berkowitz and F. J. Garner, The calculation of multidimensional hermite polynomials and gram-charlier coefficients, Mathematics of Computation 24, 537 (1970).
[88] G. Cariolaro and G. Pierobon, Fock expansion of multimode pure gaussian states, Journal of Mathematical Physics 56, 122109 (2015).
[89] D. S. Phillips, M. Walschaers, J. J. Renema, I. A. Walmsley, N. Treps, and J. Sperling, Benchmarking of gaussian boson sampling using two-point correlators, Phys. Rev. A 99, 023836 (2019).
[90] B. Gupt, J. Izaac, and N. Quesada, The walrus: a library for the calculation of hafnians, hermite polynomials and gaussian boson sampling, Journal of Open Source Software 4, 1705 (2019).
[91] J. K. Nauth, Full time-dependent counting statistics of highly entangled biphoton states (2022).
[92] R. Chattanvell and R. Shannugam, Generating Functions in Engineering and the Applied Sciences (Morgan & Claypool Publishers, 2019).
[93] H. S. Wilf, Generatingfunctionology, 2nd ed. (Elsevier, 2014).
[94] N. Lütkenhaus, Security against individual attacks for realistic quantum key distribution, Phys. Rev. A 61, 052304 (2000).