Non-Gaussian Quantum States and Where to Find Them

Mattia Walschaers

Tutorial

Non-Gaussian Quantum States and Where to Find Them

Mattia Walschaers*

Laboratoire Kastler Brossel, Sorbonne Université, CNRS, ENS-Université PSL, Collège de France, 4 place Jussieu, Paris F-75252, France

(Received 2 May 2021; published 28 September 2021)

Gaussian states have played an important role in the physics of continuous-variable quantum systems. They are appealing for the experimental ease with which they can be produced, and for their compact and elegant mathematical description. Nevertheless, many proposed quantum technologies require us to go beyond the realm of Gaussian states and introduce non-Gaussian elements. In this Tutorial, we provide a roadmap for the physics of non-Gaussian quantum states. We introduce the phase-space representations as a framework to describe the different properties of quantum states in continuous-variable systems. We then use this framework in various ways to explore the structure of the state space. We explain how non-Gaussian states can be characterized not only through the negative values of their Wigner function, but also via other properties such as quantum non-Gaussianity and the related stellar rank. For multimode systems, we are naturally confronted with the question of how non-Gaussian properties behave with respect to quantum correlations. To answer this question, we first show how non-Gaussian states can be created by performing measurements on a subset of modes in a Gaussian state. Then, we highlight that these measured modes must be correlated via specific quantum correlations to the remainder of the system to create quantum non-Gaussian or Wigner-negative states. On the other hand, non-Gaussian operations are also shown to enhance or even create quantum correlations. Finally, we demonstrate that Wigner negativity is a requirement to violate Bell inequalities and to achieve a quantum computational advantage. At the end of the Tutorial, we also provide an overview of several experimental realizations of non-Gaussian quantum states in quantum optics and beyond.

DOI: 10.1103/PRXQuantum.2.030204

CONTENTS

*mattia.walschaers@lkb.upmc.fr

I. INTRODUCTION

Gaussian states have a long history in quantum physics, which dates back to Schrödinger's introduction of the coherent state as a means to study the harmonic oscillator [1]. In later times, Gaussian states rose to prominence due to their importance in the description of Bose gases [2–4] and in the theory of optical coherence [5,6]. With the advent of quantum-information theory, the elegant mathematical structure of Gaussian states made them important objects in the study of continuous-variable (CV) quantum-information theory [7–9]. In this Tutorial, we focus on bosonic systems, which means that the continuous variables of interest are field quadratures. Gaussian quantum states are then defined as the states for which measurement statistics of these field quadratures is Gaussian.

Gaussian states can be fully described by their mean field and covariance matrix, and, due to Williamson's decomposition [10], the latter can be studied using a range of tools from symplectic vector spaces. As such, one can directly relate quadrature squeezing to Gaussian entanglement via the Bloch-Messiah decomposition [11]. In the full state space of CV systems, Gaussian states are furthermore known to play a specific role: of all possible states with the same covariance matrix, the Gaussian state will always have the weakest entanglement [12] and the highest entropy [13]. From a theoretical point of view, Gaussian quantum states provide, thus, an elegant and highly relevant framework for quantum-information theory. On an experimental level, CV quantum information has long been motivated by advances in quantum optics, due to the capability of on-demand generation of ever larger entangled states using either spatial modes [14–16] or time-frequency modes [17–22]. Furthermore, Gaussian states also play a key role in the recent demonstration of a quantum advantage with Gaussian Boson sampling [23]. These developments have made the CV quantum optics an important platform for quantum computation [24].

Regardless of all the experimental and theoretical successes of Gaussian states, they have a major shortcoming in the context of quantum technologies: all Gaussian measurements of such states can be efficiently simulated [25]. In pioneering work on CV quantum computation, it is already argued that a non-Gaussian operation is necessary to implement a universal quantum computer in CV [26]. Later works that laid the groundwork for CV measurement-based quantum computing have left the question of this non-Gaussian operation somewhat in the open [27–29]. Common schemes, based on the cubic phase gate, turn out to be particularly hard to implement in realistic setups [30]. Furthermore, these protocols require highly non-Gaussian states, such as Gottesman-Kitaev-Preskill (GKP) states [31], to encode information. Even though such states could also serve as a non-Gaussian resource for implementing non-Gaussian gates [32], these states remain notoriously challenging to produce. In spite of the practical problems involved with non-Gaussian states, one is obliged to venture into non-Gaussian territory to reach a quantum computational advantage in the CV regime [33]. This emphasizes the importance of a general understanding of non-Gaussian states and their properties. In this Tutorial, we attempt to provide a roadmap to navigate within this quickly developing field.

In Sec. II, we take an unusual start to introduce CV systems. We first present some elements of many-boson physics, by treating Fock space. This mathematical environment is probably familiar to most readers to describe photons. We then explain how such a Fock space can also be described in phase space, which is the more natural framework from CV quantum optics. We introduce phase-space representations of states and observables in CV systems such as the Wigner function, and to familiarize the reader with the language of multimode systems. By first reviewing the basics of Fock space, we can make interesting connections between what is known as the discrete-variable (DV) approach and the CV approach to quantum optics. We see that there is often a shady region between these two frameworks, where techniques that are typically associated with one framework can be applied in the other. We finally argue that the main distinction lies in whether one measures photons (DV) or field quadratures (CV).

In Sec. III, we provide the reader with an introduction to some of the different structures that can be identified in the space of CV quantum states. When pure states are considered, all non-Gaussian states are known to have a nonpositive Wigner function [34,35], but this no longer holds when mixed states enter the game [36]. In the entirety of the state space, non-Gaussian states occupy such a vast territory that it is impossible to describe all of them within one single formalism. Nevertheless, there has recently been considerable progress in the classification of non-Gaussian states [37,38]. We introduce some key ideas behind quantum non-Gaussianity, the stellar rank, and Wigner negativity as tools to characterize non-Gaussian states.

Section IV introduces two main families of techniques to create non-Gaussian states starting from Gaussian inputs. The first approach concentrates on deterministic

methods, which rely on the implementation of non-Gaussian unitary transformations. We show how such transformations can be built by using a specific non-Gaussian gate. We then introduce the second class of techniques, which are probabilisitic and rely on performing non-Gaussian measurements on a Gaussian state and conditioning on a certain measurement result. We introduce our recently developed approach to describe these systems [39] and present mode-selective photon subtraction as a case study.

Then all the pieces are set to discuss the interplay between non-Gaussian effects and quantum correlations in Sec. V. First, we consider the resources that are required to conditionally prepare certain non-Gaussian states. The conditional scheme relies on performing a non-Gaussian measurement on one part of a bipartite Gaussian state, and will show that the nature of the quantum correlations in this bipartite state is essential. We show that we can only generate quantum non-Gaussian states if the initial bipartite state is entangled. Furthermore, to conditionally generate Wigner negativity we even require quantum steering. In the second part of Sec. V, we show how non-Gaussian operations can in return enhance or create quantum correlations. Finally, we show that Wigner negativity (in either the state or the measurement) is necessary to violate Bell inequalities in CV systems.

In a similar fashion, we spend most of Sec. VI explaining the result of Ref. [33], which shows that Wigner negativity is also necessary to reach a quantum advantage. To show this, we explicitly construct a protocol to efficiently simulate the measurement outcomes of a setup with states, operations, and detectors that are described by positive Wigner functions. In the remainder of the section, we provide comments on the quantum computational advantage reached with Gaussian Boson sampling.

Finally, in Sec. VII, we provide a quick overview of non-Gaussian states in CV experiments. Due to the author's background, the first half of this overview focuses on quantum optics. In the second part, we also discuss some key developments in other branches of experimental quantum physics. Readers should be warned that this is by no means an extensive review of all the relevant experimental progress. A more general conclusion and outlook on what the future may have in store is presented in Sec. VIII.

II. CONTINUOUS-VARIABLE QUANTUM STATES

Before we can start our endeavor to classify non-Gaussian states of CV systems and study their properties, we must develop some basic formalism for dealing with multimode bosonic systems. At the root of bosonic systems lies the canonical commutation relation, $[\hat{x},\hat{p}] \sim i\mathbb{1}$, which can be traced back to the early foundations of quantum mechanics. The study of the algebra of such noncommuting observables has given birth to rich branches of mathematics and mathematical physics that ponder on the subtleties of these observables and their associated states. In this Tutorial, we keep a safe distance from the representation theory of the associated C^* algebras that describe bosonic field theories in their most general sense. We do refer interested readers to a rich but technical literature [4,40–42].

In this Tutorial, we do exclusively work within the Fock representation, which implies that we consider systems with a finite expectation value for the number of particles. In quantum optics, this assumption translates to the logical requirement that energies remain finite. There are many approaches to mathematically construct such systems (luckily for us they are all equivalent [43–46]). Here, we briefly present two such approaches that nicely capture one of the key dualities on quantum physics. First we take the particle approach by introducing the Fock space that describes identical bosonic particles in Sec. A. Subsequently, in Sec. B, we take the approach that starts out from a wave picture, by concentrating on the phase-space representation of the electromagnetic field. Here we also introduce the phase-space representations of CV quantum states that proves to be crucial tools in the remainder of this Tutorial. We show how these approaches are quite naturally two sides of the same coin. In Sec. C, we briefly discuss the concept of modes and the role they play in CV quantum systems. This subsection is both intended to provide some clarification about common jargon and to eliminate common misconceptions. We finish this section by presenting a brief case study of Gaussian states in Sec. D, reviewing some key results. After all, it is difficult to appreciate the subtleties of non-Gaussian states without having a flavor from their Gaussian counterparts.

A. Fock space

In typical quantum mechanics textbooks, the story of identical particles usually starts by considering a set of n particles, which are each described by a quantum state vector in a single-particle Hilbert space $\mathcal{H}$, thus for the ith particle we ascribe a state vector $|\psi_i\rangle \in \mathcal{H}$. The joint state of these n particles is then given by the tensor product of the state vectors $|\psi_1\rangle, \ldots, |\psi_n\rangle$. However, if the particles are identical in all their internal degrees of freedom, we should be free to permute them without changing the observed physics. Formally, such permutation is implemented by a unitary operator U_σ, for the permutation $\sigma \in S_n$, which acts as

$$U_\sigma |\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle = \left|\psi_{\sigma(1)}\right\rangle \otimes \cdots \otimes \left|\psi_{\sigma(n)}\right\rangle. \quad (1)$$

Invariance of physical observables under such permutations can be achieved by either imposing the n-particle state vector to be fully symmetric (bosons) or fully antisymmetric (fermions) under these permutations of particles. In this Tutorial, we focus exclusively on bosons, and

thus the condition that must be imposed to obtain a bosonic n-particle state is

$$U_\sigma \left|\Psi^{(n)}\right\rangle = \left|\Psi^{(n)}\right\rangle. \tag{2}$$

Because these are the only states that are permitted to describe the bosonic system, we commonly use the Hilbert space $\mathcal{H}_s^{(n)}$, which is a subspace of $\mathcal{H}^{\otimes n}$ that contains only those states that fulfil Eq. (2). It is usually convenient to generate these spaces with a set of elementary tensors, known as Fock states, which we define as

$$|\psi_1\rangle \vee \cdots \vee |\psi_n\rangle := \sum_{\sigma\in S_n} \left|\psi_{\sigma(1)}\right\rangle \otimes \ldots \left|\psi_{\sigma(n)}\right\rangle, \tag{3}$$

such that

$$\mathcal{H}_s^{(n)} = \overline{\text{span}\left\{\, |\psi_1\rangle \vee \cdots \vee |\psi_n\rangle \mid |\psi_i\rangle \in \mathcal{H}\right\}}, \tag{4}$$

where we refer to the Appendix for some further details on the span. This fully describes a system of n bosonic particles in what is often referred to as first quantization. It is interesting to note that these identical particles appear to be entangled with respect to the tensor product structure of $\mathcal{H}^{\otimes n}$. There is still debate on whether this is a mathematical artefact of our description or rather a genuine physical feature of identical particles. Even though there is still debate about how to exactly define entanglement between indistinguishable particles [47,48], several authors have shown how these symmetrizations can [49–52] induce useful entanglement. Furthermore, it is undeniable that this structure leads to physical interference phenomena that do not exist for distinguishable particles [53].

The name "first quantization" suggests the existence of a second quantization, which turns out to be more appropriate for this Tutorial. Second quantization finds its origins in models where particle numbers are not fixed or conserved. This formalism is largely based on creation and annihilation operators, denoted $\hat{a}^\dagger$ and $\hat{a}$, respectively, that add or remove particles. To accommodate these operators in our mathematical framework, we must equip our Hilbert space to describe a varying number of particles. Therefore, we introduce the Fock space

$$\Gamma(\mathcal{H}) := \mathcal{H}_s^{(0)} \oplus \mathcal{H}_s^{(1)} \oplus \mathcal{H}_s^{(2)} \oplus \ldots, \tag{5}$$

where the single-particle Hilbert space is given by $\mathcal{H}_s^{(1)} = \mathcal{H}$. Furthermore, we retrieve a peculiar component $\mathcal{H}_s^{(0)}$, which describes the fraction of the system that contains no particles at all. On its own, $\mathcal{H}_s^{(0)}$ is thus populated by a single state $|0\rangle$ that we refer to as the vacuum. This implies that technically $\mathcal{H}_s^{(0)} \cong \mathbb{C}$ the zero-particle Hilbert space is just described by a complex number that corresponds to the overlap of the state with the vacuum. A general pure state in Fock space $|\Psi\rangle \in \Gamma(\mathcal{H})$ can then be described using the structure, Eq. (5), as

$$|\Psi\rangle = \Psi^{(0)} \oplus \Psi^{(1)} \oplus \Psi^{(2)} \oplus \ldots, \tag{6}$$

where $\Psi^{(i)} \in \mathcal{H}_s^{(i)}$ are non-normalized vectors (and therefore we omit the $|.\rangle$) in the i-particle Hilbert space. Because $|\Psi\rangle$ is a state, we must impose the normalization condition $\|\Psi\|^2 = \sum_{i=0}^{\infty} \|\Psi^{(i)}\|^2 = 1$

We can now define a creation operator $\hat{a}^\dagger(\varphi)$ for every $\varphi \in \mathcal{H}$ [54], which acts as

$$\begin{aligned}\hat{a}^\dagger(\varphi)\,|\Psi\rangle = 0 \oplus \left(\Psi^{(0)}\,|\varphi\rangle\right) \oplus \left(|\varphi\rangle \vee \Psi^{(1)}\right) \\ \oplus \left(|\varphi\rangle \vee \Psi^{(2)}\right) \oplus \ldots\end{aligned} \tag{7}$$

In the same spirit, it is possible to provide an explicit construction of the annihilation operators $\hat{a}(\varphi)$, but here we content ourselves by just introducing the annihilation operator as the hermitian conjugate of the creation operator. Just as the creation operator that literally adds a particle to the system, the annihilation operator literally removes one. One additional property of the annihilation operators is that they destroy the vacuum state:

$$\hat{a}(\varphi)\,|0\rangle = 0. \tag{8}$$

We can now use creation and annihilation operators to build an arbitrary Fock state by creating particles on the vacuum state

$$|\psi_1\rangle \vee \cdots \vee |\psi_n\rangle = \hat{a}^\dagger(\psi_1)\hat{a}^\dagger(\psi_2)\ldots\hat{a}^\dagger(\psi_n)\,|0\rangle \tag{9}$$

and by considering superpositions of such Fock states, we can ultimately generate the entire Fock space. By considering any basis of the single-particle Hilbert space $\mathcal{H}$ and constructing all possible Fock states of all possible lengths that can be formed by generating particles in these basis vectors we construct a basis of the Fock space $\Gamma(\mathcal{H})$. We refer to this basis as the Fock basis.

The beauty of second quantization lies in the natural appearance of states, which have no fixed particle number. The most important example is the coherent state

$$|\alpha\rangle := e^{-\frac{\|\alpha\|^2}{8}} \sum_{j=0}^{\infty} \frac{[\hat{a}^\dagger(\alpha)]^j}{2^j j!}\,|0\rangle\,, \tag{10}$$

where $\alpha \in \mathcal{H}$ is a non-normalized vector in the single particle Hilbert space. One can, indeed, simply generalise (7) to non-normalized vectors in $\mathcal{H}$ which we use explicitly in Eq. (10). Second, we note that an unusual factor 2 is included to make the definition consistent with Eq. (62).

In quantum optics, these coherent states are crucial objects as they describe perfectly coherent light [5]. It is

important to remark that a coherent state is always generated by a single vector in the single-particle Hilbert space. Coherent states often provide a good approximation for the state that is produced by a single-mode laser far above threshold [55]. More generally, the study of laser light is a whole field in its own right and often the light deviates from the fully coherent approximation.

The creation and annihilation operators are not only important objects because they populate the Fock space; they are also of key importance for describing observables in a many-boson system. These operators are the generators of the algebra of observables that represents the canonical commutation relations on Fock space. This implies that any observable can ultimately be approximated by a polynomial of creation and annihilation operators. At the heart of this mathematical formalism lies the canonical commutation relation (CCR):

$$[\hat{a}(\varphi), \hat{a}^\dagger(\psi)] = \langle \varphi \mid \psi \rangle, \tag{11}$$

which describes the algebra of observables. Note that this relation holds for any vectors $|\varphi\rangle$ and $|\psi\rangle$ in the single-particle Hilbert space $\mathcal{H}$. These vectors should not form a basis, nor should they be orthogonal. When $|\varphi\rangle = |\psi\rangle$, we find that $[\hat{a}(\psi), \hat{a}^\dagger(\psi)] = 1$. On the other hand, when the single-particle states $|\varphi\rangle$ and $|\psi\rangle$ are fully orthogonal, we find that $[\hat{a}(\varphi), \hat{a}^\dagger(\psi)] = 0$. In these cases we recover the typical creation and annihilation operators for harmonic oscillators. However, by introducing the creations and annihilation operators through Eq. (7), we can also deal with more general cases. Furthermore, all definitions and the form of the CCR are still valid when φ and ψ are unnormalized vectors in $\mathcal{H}$. A more detailed discussion can be found in Ref. [56].

When we leave the realm of pure states, the description of quantum states becomes tedious. Commonly, one uses a density operator $\hat{\rho}$ with $\mathrm{tr}\,\hat{\rho} = 1$ to formally describe a state. However, we can generally think of these density operators as infinite-dimensional matrices with an infinite number of components in the Fock basis. In other words, this is not necessarily a convenient description. In an operational sense, any state is considered to be characterized when we know all the moments of all the possible observables. Because the creation and annihilation operators generate the algebra, one knows all the moments of all the observables if one knows all the correlation functions $\mathrm{tr}[\hat{\rho}\,\hat{a}^\dagger(\psi_1)\ldots\hat{a}^\dagger(\psi_n)\hat{a}(\varphi_1)\ldots\hat{a}(\varphi_m)]$, for all possible lengths n and m. Even though this might seem like an equally challenging endeavor, much of quantum statistical mechanics boils down to finding expressions of the correlation functions for relevant classes of states.

B. Phase space

In the previous subsection, we started our analysis by extending a system of one quantum particle to a system of many quantum particles. Here we follow a different route, where we start by considering the classical electric field. With some effort, we can apply such an analysis to any bosonic field, but in this Tutorial we focus on quantum optics as our main field of application. For a more extensive introduction from a quantum optics perspective we recommend Refs. [55,57,58], whereas a general introduction to quantum physics in phase space can be found in Ref. [59].

A traveling electromagnetic wave is described by a solution of Maxwell's equations. As is commonly the case in optics, we focus on the complex representation of the electric field, which is generally given by $\mathbf{E}^{(+)}(\mathbf{r}, t)$. It is related to the real-valued electric field $\mathbf{E}(\mathbf{r}, t)$ that is encountered in standard electrodynamics textbooks by $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}^{(+)}(\mathbf{r}, t) + \left[\mathbf{E}^{(+)}(\mathbf{r}, t)\right]^*$. To express the electric field, it is useful to introduce an orthonormal mode basis $\{\mathbf{u}_i(\mathbf{r}, t)\}$. These modes are solutions to Maxwell's equations

$$\nabla \cdot \mathbf{u}_i(\mathbf{r}, t) = 0, \tag{12}$$

$$\left(\Delta - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\mathbf{u}_i(\mathbf{r}, t) = 0. \tag{13}$$

The orthogonalization property is implemented by the following condition:

$$\frac{1}{V}\int_V \mathrm{d}^3\mathbf{r}\,[\mathbf{u}_i(\mathbf{r}, t)]^*\,\mathbf{u}_j(\mathbf{r}, t) = \delta_{i,j}, \tag{14}$$

where V is some large volume that contains the entire physical system. This assumption serves the practical purpose of allowing us to consider a discrete mode basis and on top it makes physical sense. Note that we do not integrate over t, which implies that at every instant of time t we consider a mode basis that is normalized with respect to the spatial degrees of freedom. It is practical to assume that all relevant physics can be described by a (possibly large) finite number of modes m. These modes now form a basis in which we can expand any solution to Maxwell's equations and thus we may write

$$\mathbf{E}^{(+)}(\mathbf{r}, t) = \sum_{j=1}^{m} \mathcal{E}_j\,\mathbf{u}_j(\mathbf{r}, t), \tag{15}$$

where $\mathcal{E}_i$ are a set of complex numbers, which can be written in terms of the real and imaginary parts

$$\mathcal{E}_j = E_j^{(x)} + iE_j^{(p)}. \tag{16}$$

These real and imaginary parts of the field are known as the amplitude and phase quadrature, respectively. We can interpret these quantities $\vec{E} := (E_1^{(x)}, E_1^{(p)}, \ldots, E_m^{(x)}, E_m^{(p)}) \in \mathbb{R}^{2m}$ as the coordinate in optical phase space that describes the light field.

The space of solutions of Maxwell's equations forms a Hilbert space, which we call the mode space $\mathcal{M}$ and the mode basis chosen to describe this space is far from unique. As with all Hilbert spaces, we can define unitary transformations and use them to change from one basis to another. As such, let us introduce the unitary operator U to change between bases

$$\mathbf{u}_i(\mathbf{r},t) = \sum_{j=1}^{m} U_{ji}\mathbf{v}_j(\mathbf{r},t), \tag{17}$$

$$\mathbf{v}_i(\mathbf{r},t) = \sum_{j=1}^{m} U_{ji}^{\dagger}\mathbf{u}_j(\mathbf{r},t), \tag{18}$$

where we can in principle obtain U as an infinite-dimensional matrix with

$$U_{ji} = \frac{1}{V}\int_V \mathrm{d}^3\mathbf{r}\left[\mathbf{v}_j(\mathbf{r},t)\right]^* \mathbf{u}_i(\mathbf{r},t), \tag{19}$$

which remarkably does not depend on time due to the normalization properties of the mode bases. We can analogously expand the electric field in the new mode basis

$$\mathbf{E}^{(+)}(\mathbf{r},t) = \sum_i \mathcal{E}_i'\mathbf{v}_i(\mathbf{r},t), \tag{20}$$

where $\mathcal{E}_i' = \sum_j U_{ij}\mathcal{E}_j$. This observation is of great importance when we quantize the electric field. The change of mode basis also imposes a change of coordinates in the optical phase space. Like in Eq. (16) the new components can also be divided in real and imaginary parts, which leads to a new coordinate $\vec{E}'$. Because the coordinate vectors in optical phase space are real $2m$-dimensional vectors, we obtain

$$\vec{E}' = O\vec{E}, \tag{21}$$

where O is an orthonormal transformation. However, the orthogonal transformation O on the phase space must correspond to the unitary transformation U on the modes, which imposes the constraint

$$O_{2i-1,2j-1} = \frac{1}{2}(U_{ij} + U_{ij}^*), \tag{22}$$

$$O_{2i-1,2j} = -\frac{1}{2i}(U_{ij} - U_{ij}^*), \tag{23}$$

$$O_{2i,2j-1} = \frac{1}{2i}(U_{ij} - U_{ij}^*), \tag{24}$$

$$O_{2i,2j} = \frac{1}{2}(U_{ij} + U_{ij}^*). \tag{25}$$

This imposes a symplectic structure to the transformation O such that the optical phase space, just like the phase space of analytical mechanics, can be treated as a symplectic space. The conserved symplectic structure associated with this space is given by

$$\Omega = \bigoplus_{j=1}^{m}\omega, \text{ with } \omega = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{26}$$

such that $O^T\Omega O = \Omega$. Note that Ω can be interpreted as a matrix representation of the imaginary i, in the sense that it has the properties $\Omega^T = -\Omega$ and $\Omega^2 = -\mathbb{1}$ [60].

In quantum optics, the electric field of light is treated as a quantum observable $\hat{\mathbf{E}}^{(+)}(\mathbf{r},t)$. In this quantization, the modes, i.e., the normalized solutions to Maxwell's equations, remain classical objects and all the quantum features are absorbed in the coefficients. We can thus write

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r},t) = \sum_i \mathcal{E}_i^{(1)}\frac{\hat{x}_i + i\hat{p}_i}{2}\mathbf{u}_i(\mathbf{r},t), \tag{27}$$

where $\mathcal{E}_i^{(1)}$ is a constant that carries the dimensions of the field, which can be interpreted as the electric field of a single photon. Glossing over many subtleties of the quantization of the electromagnetic field, we remind the reader that any system that is described on phase space can be quantized through canonical quantization. The quadrature operators $\hat{x}_j$ and $\hat{p}_k$ therefore follow the canonical commutation relations $[\hat{x}_j,\hat{p}_k] = 2i\delta_{j,k}$, such that they satisfy the Heisenberg relation $\Delta\hat{x}\Delta\hat{p} \geqslant 1$. As they are introduced above, the quadrature operators are specifically related to the specific mode basis. Indeed, $\hat{x}_j$ and $\hat{p}_j$ are the quadrature operators that describe the field in mode $\mathbf{u}_j(\mathbf{r},t)$. Thus, when we change the basis of modes, we should change the quadrature operators accordingly in line with Eq. (21). To overcome these difficulties, it is often convenient to introduce a basis-independent expression for the quadrature operators, which can be done by mapping any point in the optical phase space $\vec{f} \in \mathbb{R}^{2m}$ to an observable $\hat{q}(\vec{f})$, given by

$$\hat{q}(\vec{f}) := \sum_{j=1}^{m} f_{2j-1}\hat{x}_j + f_{2j}\hat{p}_j. \tag{28}$$

These quadrature operators follow a generalized version of the CCR, given by

$$[\hat{q}(\vec{f}_1),\hat{q}(\vec{f}_2)] = -2i\vec{f}_1{}^T\Omega\vec{f}_2, \text{ for all } \vec{f}_1,\vec{f}_2 \in \mathbb{R}^{2m}. \tag{29}$$

We highlight the particular case where $[\hat{q}(\vec{f}),\hat{q}(\Omega\vec{f})] = 2i\|\vec{f}\|^2$, such that we recover the typical form of the CCR for $\|\vec{f}\| = 1$. This highlights that Ω maps an amplitude quadrature to its associated phase quadrature. From a mathematical point of view, everything is perfectly well

defined for arbitrary $\vec{f} \in \mathbb{R}^{2m}$ and no normalization conditions have to be imposed. From Eq. (28) we can see that the norm of $\vec{f}$ can be factored out, such that it serves as a general rescaling factor of the quadrature operator. In a physical context, when a quadrature is measured, it is common to renormalize measurements to units of vacuum noise, which practically means that we set $\|\vec{f}\| = 1$. Unless explicitly stated otherwise, we assume that $\|\vec{f}\| = 1$ throughout this Tutorial. We can use these general quadrature operators to express electric field operator as

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r},t) = \sum_{j=1}^{m} \mathcal{E}_j^{(1)} \frac{\hat{q}(\vec{e}_j) + i\hat{q}(\Omega\vec{e}_j)}{2} \mathbf{u}_j(\mathbf{r},t), \tag{30}$$

and we can use the basis transformation (21) to equivalently express the electric field operator in a different mode basis as

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r},t) = \sum_{j=1}^{m} \mathcal{E}_j^{(1)} \frac{\hat{q}(O\vec{e}_j) + i\hat{q}(\Omega O\vec{e}_j)}{2} \mathbf{v}_j(\mathbf{r},t). \tag{31}$$

This procedure shows us that optical elements, that change the mode basis, change the associated quadrature operators accordingly.

Equation (30) shows us explicitly that quadrature operators $\hat{q}(\vec{f})$ and $\hat{q}(\Omega\vec{f})$ correspond to the same mode, regardless of the mode basis. This reflects the fact that $\vec{f}$ generates one axis in the optical phase space and $\Omega\vec{f}$ generates the second axis that corresponds to the same mode. As such, any arbitrary mode comes with an associated two-dimensional phase space that mathematically can be denoted as $\mathrm{span}(\vec{f}, \Omega\vec{f})$. Because this phase space is uniquely associated with a specific mode, we introduce the notation

$$\mathbf{f} = \mathrm{span}(\vec{f}, \Omega\vec{f}), \tag{32}$$

and we refer to this as "mode **f**." This allows us to concentrate on the multimode quantum states within this Tutorial, while the specifications of the modes can be left ambiguous. The modes can be seen as the physical implementations of the quantum system and are of major importance in the experimental setting as multimode quantum optics experiments rely on the manipulation of these modes.

Multimode quantum states define expectation values of the field, and when we consider CV quantum optics, we primarily focus on the expectation values of the quadrature operators $\hat{q}(\vec{f})$. These operators are unbounded and have a continuous spectrum. The measurement of a field quadrature thus leads to a continuum of possible outcomes and the continuous-variable approach to quantum optics implies that this characterizes quantum properties of light through the measurement of such quadrature operators.

Formally, we can again describe a quantum state on such a system by a density operator $\hat{\rho}$ but this description is rather inconvenient. It turns out that the quadrature operators $\hat{q}(\vec{f})$ generate the algebra of observables for the quantum system that is comprised within our multimode light. In other words, any observable can be approximated by a polynomial of quadrature operators. This generally implies that we can fully characterize the quantum state $\hat{\rho}$ by correlation functions of the type $\mathrm{tr}[\hat{\rho}\hat{q}(\vec{f}_1)\ldots\hat{q}(\vec{f}_n)]$. When we know these correlation functions for all lengths n and normalized vectors in phase space, we have fully characterized the state.

To go beyond the information that is contained in correlation functions, it is often convenient to consider probability distributions as a whole. For a single quadrature $\hat{q}(\vec{f})$ we can introduce the characteristic function for any $\lambda \in \mathbb{R}$ as

$$\chi(\lambda) = \mathrm{tr}[\hat{\rho}e^{i\lambda\hat{q}(\vec{f})}] = \sum_{n=0}^{\infty} \frac{(i\lambda)^n}{n!} \mathrm{tr}[\hat{\rho}\hat{q}(\vec{f})^n], \tag{33}$$

which is clearly related to the moments $\mathrm{tr}[\hat{\rho}\hat{q}(\vec{f})^n]$. The characteristic function is the Fourier transform of the probability distribution of the outcomes of observable $\hat{q}(\vec{f})$. We can thus obtain the probability distribution as

$$p(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{d}\lambda\, \chi(\lambda) e^{-ix\lambda}. \tag{34}$$

This approach can be readily generalized to the joint probability distribution for a set of commuting quadrature operators. We thus consider $\vec{f}_1, \ldots \vec{f}_n$ with $[\hat{q}(\vec{f}_j), \hat{q}(\vec{f}_k)] = 0$ for all j, k, and we define for all $\vec{\lambda} = \lambda_1\vec{f}_1 + \lambda_2\vec{f}_2 + \cdots + \lambda_n\vec{f}_n$ (note that $\vec{\lambda}$ is not normalized). We can then use the properties of the quadrature operators to construct $\hat{q}(\vec{\lambda}) = \sum_{k=1}^{n} \lambda_k \hat{q}(\vec{f}_k)$ and define the function

$$\chi(\vec{\lambda}) = \mathrm{tr}[\hat{\rho}e^{i\hat{q}(\vec{\lambda})}]. \tag{35}$$

This function generates all the correlations between observables $\hat{q}(\vec{f}_1), \ldots, \hat{q}(\vec{f}_n)$ and it can be used to obtain the multivariate probability distribution

$$p(\vec{x}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \mathrm{d}\vec{\lambda}\, \chi(\vec{\lambda}) e^{-i\vec{\lambda}^T\vec{x}}, \tag{36}$$

where $\mathrm{d}\vec{\lambda} = \mathrm{d}\lambda_1 \ldots \mathrm{d}\lambda_n$. The function $p(\vec{x})$ describes the probability density to jointly obtain $x_1, \ldots, x_n$ as measurement outcomes for the measurements of $\hat{q}(\vec{f}_1), \ldots, \hat{q}(\vec{f}_n)$, respectively. This approach relies on the fact that commuting observables can be jointly measured and what we presented to derive Eqs. (34) and (36) is ultimately just classical probability theory. However, not all quadrature operators commute such that joint measurements are not

always possible. This implies that a quantum state cannot be straightforwardly defined by a probability distribution of the optical phase space.

Intriguingly, we can carry out the same procedure for a full multimode system over a set of m modes. To this goal, let us define the quantum characteristic function

$$\chi : \mathbb{R}^{2m} \to \mathbb{C} : \vec{\lambda} \mapsto \chi(\vec{\lambda}) := \mathrm{tr}[\hat{\rho} e^{i\hat{q}(\vec{\lambda})}]. \tag{37}$$

This function, defined on the full optical phase space, can be used to generate all correlation functions between all quadrature operators. As such, it does characterize the full quantum state, but it is common practice to rather study its inverse Fourier transform, which is known as the Wigner function [61–63]

$$W(\vec{x}) := \frac{1}{(2\pi)^{2m}} \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\lambda}\ \chi(\vec{\lambda}) e^{-i\vec{\lambda}^T \vec{x}}. \tag{38}$$

This function has many appealing properties even though it is not a probability distribution but rather a quasiprobability distribution. First of all, the Wigner function is normalized, i.e., $\int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\ W(\vec{x}) = 1$. Furthermore, its marginals consistently describe all the joint probability distributions for sets of commuting quadratures in the system. Formally, this implies that $p(\vec{x})$ of Eq. (36) can be obtained by integrating over all the phase-space axes that are not contained within $\mathrm{span}(\vec{f}_1, \ldots \vec{f}_n)$. To do so, let us introduce the n-dimensional vector $\vec{x}_M$ that is associated with the measured quadratures, and the $2m - n$ dimensional vectors $\vec{x}_c$, which describe all other axes in phase space. An arbitrary point in phase space can thus be written as $\vec{x} = \vec{x}_M \oplus \vec{x}_c$. Then we find that

$$P(\vec{x}_M) = \int_{\mathbb{R}^{2m-n}} \mathrm{d}\vec{x}_c\ W(\vec{x}_M \oplus \vec{x}_c). \tag{39}$$

Finally, the Wigner function also produces the correct expectation values

$$\int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\, \vec{f}_1{}^T \vec{x} \ldots \vec{f}_n{}^T \vec{x} W(\vec{x}) = \mathrm{Re}\{\mathrm{tr}[\hat{\rho}\hat{q}(\vec{f}_1) \ldots \hat{q}(\vec{f}_n)]\}. \tag{40}$$

Note that considering the real part of $\mathrm{tr}[\hat{\rho}\hat{q}(\vec{f}_1) \ldots \hat{q}(\vec{f}_n)]$ is essentially equivalent to considering symmetric ordering of the operators. Regardless of these nice properties the Wigner function is by itself not a well-defined probability distribution. Due to complementarity, the function can reach negative values for some states. This Wigner negativity is consistent with the impossibility to jointly describe the measurement statistics of all quadratures while also complying with the laws of quantum physics (notably the Heisenberg relation). The profound relation between negativity of the Wigner functions and joint measurability is perhaps most strikingly illustrated by its connection to contextuality [64]. The formalism of Wigner functions can be used to construct phase-space representations of arbitrary observables by introducing

$$\chi_A(\vec{\lambda}) = \mathrm{tr}[\hat{A} e^{i\hat{q}(\vec{\lambda})}], \tag{41}$$

such that the Wigner representation is given by

$$W_A(\vec{x}) = \frac{1}{(2\pi)^{2m}} \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\lambda}\ \chi_A(\vec{\lambda}) e^{-i\vec{\lambda}^T \vec{x}}. \tag{42}$$

These Wigner representations have the appealing property that

$$\mathrm{tr}[\hat{A}\hat{\rho}] = (4\pi)^m \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\ W_A^*(\vec{x}) W(\vec{x}). \tag{43}$$

When $\hat{A}$ is an observable and thus has $\hat{A} = \hat{A}^\dagger$, its Wigner function will be real such that $W_A^*(\vec{x}) = W_A(\vec{x})$. However, it may sometimes be useful to extend the formalism to more general operators. As such the entire theory of continuous-variable quantum systems can be developed using Wigner functions.

Several aspects of the phase-space representations in this section are reminiscent of earlier results in Sec. A, which was fully developed in a language of particles (also known as a discrete-variable approach). Indeed, the algebra of operators that is generated by the creation and annihilation operators is actually the same as the algebra generated by the quadrature operators. To formalize this, we must first stress that the optical phase space is isomorphic to an m-dimensional complex Hilbert space, which can equally be interpreted as the single-particle Hilbert space of a photon. Formally, this equivalence is constructed through the bijection (see also the Appendix)

$$\vec{f} \in \mathbb{R}^{2m} \mapsto \sum_j (f_{2j-1} + i f_{2j}) \left|\varphi_j\right\rangle \in \mathcal{H}, \tag{44}$$

where $\{\left|\varphi_j\right\rangle\}$ is an arbitrary basis of $\mathcal{H}$. We can introduce the operators

$$\begin{aligned} \hat{a}(\vec{f}) &= \frac{1}{2}[\hat{q}(\vec{f}) + i\hat{q}(\Omega\vec{f})], \\ \hat{a}^\dagger(\vec{f}) &= \frac{1}{2}[\hat{q}(\vec{f}) - i\hat{q}(\Omega\vec{f})]. \end{aligned} \tag{45}$$

By using Eq. (44), we can naturally associate these operators to creation and annihilation operators on the single-particle Hilbert space. We retrieve the canonical commutation relation

$$[\hat{a}(\vec{f}_1), \hat{a}^\dagger(\vec{f}_2)] = \vec{f}_1{}^T \vec{f}_2 - i\vec{f}_1{}^T \Omega \vec{f}_2, \tag{46}$$

which can be connected to the inner product on the Hilbert space $\mathcal{H}$ via Eq. (44).

The definition of creation and annihilation operators allows us to make sense of the vacuum state in our phase-space picture. The vacuum state is completely characterized by the property

$$\hat{a}(\vec{f})\,|0\rangle = 0, \text{ for all } \vec{f} \in \mathbb{R}^{2m}. \tag{47}$$

This simple fact can be used to evaluate the quantum characteristic function

$$\chi_0(\vec{\lambda}) = \mathrm{tr}(|0\rangle\langle 0|\, e^{i[\hat{a}^\dagger(\vec{\lambda})+\hat{a}(\vec{\lambda})]}) \tag{48}$$

$$= \sum_{n=0}^{\infty} \frac{i^n}{n!} \langle 0|\,[\hat{a}^\dagger(\vec{\lambda}) + \hat{a}(\vec{\lambda})]^n\,|0\rangle \tag{49}$$

$$= \sum_{n=0}^{\infty} -\frac{\|\vec{\lambda}\|^{2n}}{2^n n!} \tag{50}$$

$$= \exp\left[-\frac{\|\vec{\lambda}\|^2}{2}\right]. \tag{51}$$

To obtain Eq. (50) we need a considerable amount of combinatorics to evaluate $\langle 0|\,[\hat{a}^\dagger(\vec{\lambda}) + \hat{a}(\vec{\lambda})]^n\,|0\rangle$. In general, it can be shown that $\langle 0|\,\hat{a}^\dagger(\vec{\lambda}_1)\ldots\hat{a}^\dagger(\vec{\lambda}_k)\hat{a}(\vec{\lambda}_{k+1})\ldots\hat{a}(\vec{\lambda}_{k+l})\,|0\rangle = 0$. Thus it suffices to cast $[\hat{a}^\dagger(\vec{\lambda}) + \hat{a}(\vec{\lambda})]^n$ in normal ordering and extract the term proportional to identity. Even though straightforward, this calculation is quite cumbersome and thus we do not present the details.

From Eq. (51) the Wigner function can be obtained via an inverse Fourier transformation that leads to

$$W_0(\vec{x}) = \frac{e^{-\frac{1}{2}\|\vec{x}\|^2}}{2\pi}. \tag{52}$$

This Wigner function describes a Gaussian distribution on the phase space with unit variance along every axis. We can thus use Eq. (40) to see that the vacuum state saturates Heisenberg's inequality, i.e., $\Delta\hat{q}(\vec{f})\Delta\hat{q}(\Omega\vec{f}) = 1$.

The quadrature operators thus generate the same algebra of observables as the creation and annihilation operators. However, both sets of observables tend to cause mathematical problems because they are unbounded operators [42,65–67]. The unboundedness means that, when $|\Psi\rangle$ is contained in the Fock space, there is no guarantee that $\hat{q}(\vec{f})\,|\Psi\rangle$ will also be contained in the Fock space. One way of solving this problem explicitly is by only considering states for which $\langle\Psi|\,\hat{q}(\vec{f})^2\,|\Psi\rangle < \infty$, such that $\hat{q}(\vec{f})\,|\Psi\rangle$ is a well-defined state. Physically this assumption makes sense, as it ultimately implies that we consider only states with finite energies. However, the unboundedness of quadrature operators also disqualifies them as well-defined generators of the C^* algebra of observables (since elements of such algebras must be bounded). C^* algebras are essential tools as they allow reconstruction of the whole framework of Hilbert spaces based on representation theory of abstract algebras (which is essentially the idea of canonical quantization). A highly formal and detailed treatment that considers all these subtleties for bosonic systems is found in Ref. [41]. The key idea is to rather consider a set of bounded operators that describes the same algebra of observables [65,66] and are known as the displacement operators:

$$\hat{D}(\vec{\alpha}) = e^{-i\hat{q}(\Omega\vec{\alpha})/2}, \tag{53}$$

where $\vec{\alpha} \in \mathbb{R}^{2m}$ need not be normalized. Again, we can use the isomorphism (44) to identify the displacement operator on the quantized phase space to a displacement operator on the Fock space. These operators can be seen as generators of the quadrature operators and they act in a very natural way on them:

$$\hat{D}^\dagger(\vec{\alpha})\hat{q}(\vec{f})\hat{D}(\vec{\alpha}) = \hat{q}(\vec{f}) + \vec{\alpha}^T\vec{f}, \tag{54}$$

which means that the value $\vec{\alpha}^T\vec{f}$ is added to the measurement outcomes of $\hat{q}(\vec{f})$. The displacement operator can be combined according to the rule

$$\hat{D}(\vec{\alpha}_1)\hat{D}(\vec{\alpha}_2) = \hat{D}(\vec{\alpha}_1 + \vec{\alpha}_2)e^{\frac{i}{4}\vec{\alpha}_1^T\Omega\vec{\alpha}_2}. \tag{55}$$

This rule is yet another representation of the canonical commutation relation and it generates the same algebra of observables. This implies that any observable $\hat{A}$ can be written as a linear combination of displacement operators. We use the Hilbert-Schmidt inner product $\langle\hat{A},\hat{B}\rangle_{\mathrm{HS}} = \mathrm{tr}[\hat{A}^\dagger\hat{B}]$ to make this explicit

$$\hat{A} = \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\lambda}\,\langle\hat{D}(2\Omega\vec{\lambda}),\hat{A}\rangle_{\mathrm{HS}}\hat{D}(2\Omega\vec{\lambda}),$$

$$= \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\lambda}\,\mathrm{tr}[\hat{A}\hat{D}(-2\Omega\vec{\lambda})]\hat{D}(2\Omega\vec{\lambda}), \tag{56}$$

and we can readily identify that

$$\mathrm{tr}[\hat{A}\hat{D}(-2\Omega\vec{\lambda})] = \chi_A^*(\vec{\lambda}). \tag{57}$$

It can then directly be seen that

$$\mathrm{tr}[\hat{A}\hat{\rho}] = \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\lambda}\,\chi_A^*(\vec{\lambda})\mathrm{tr}[\hat{D}(2\Omega\vec{\lambda})\hat{\rho}], \tag{58}$$

$$= \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\lambda}\,\chi_A^*(\vec{\lambda})\chi(\vec{\lambda}). \tag{59}$$

And we immediately obtain Eq. (43) via Plancherel's theorem [67,68].

The displacement operators also implement a unitary operation on a quantum state. This unitary operation has a remarkably simple effect when it is expressed on the level

of the Wigner function. Via the property, Eq. (55), we can calculate that

$$\hat{\rho} \mapsto \hat{D}(\vec{\alpha})\hat{\rho}\hat{D}^{\dagger}(\vec{\alpha}) \implies \chi(\vec{\lambda}) \mapsto \chi(\vec{\lambda})e^{-i\vec{\alpha}^T\vec{\lambda}}. \tag{60}$$

Performing the inverse Fourier transform of these quantum characteristic functions leads to

$$W(\vec{x}) \overset{D(\vec{\alpha})}{\mapsto} W(\vec{x}-\vec{\alpha}). \tag{61}$$

The displacement operator thus literally implements a displacement of the Wigner function by a vector $\vec{\alpha} \in \mathbb{R}^{2m}$ in phase space.

Displacement operators are also well known as the generators of the coherent states that were introduced in Eq. (10). We can combine the bijection between phase space and Hilbert space, Eq. (44), the expression of creation and annihilation operators in terms of quadratures, Eq. (45), and the definition of the displacement operator, Eq. (53), to derive that

$$|\vec{\alpha}\rangle = \hat{D}(\vec{\alpha})\,|0\rangle\,. \tag{62}$$

By combining Eqs. (52) and (61), we immediately see that the Wigner function for such a coherent state is given by

$$W_{\alpha}(\vec{x}) = W_0(\vec{x}-\vec{\alpha}) = \frac{e^{-\frac{1}{2}\|\vec{x}-\vec{\alpha}\|^2}}{2\pi}. \tag{63}$$

We emphasize that there is a slight difference between the coherent states as defined here, and coherent states as sometimes introduced in the literature. The difference is a factor of 2, which appears because we normalized the shot noise to 1 rather than to $1/2$. As such, our coherent states have an energy in mode $\mathbf{f}$, which is given by

$$\langle\vec{\alpha}|\,\hat{a}^{\dagger}(\vec{f}\,)a(\vec{f}\,)\,|\vec{\alpha}\rangle = \frac{1}{4}[(\vec{\alpha}^T\vec{f}\,)^2 + (\vec{\alpha}^T\Omega\vec{f}\,)^2]. \tag{64}$$

The coherent states lead us to two other representations of quantum states and observables: the Q function and P function. The definition of the P function is related to the idea that coherent states form an overcomplete basis of Fock space. This implies, notably, that for a m-dimensional single-particle Hilbert space

$$\frac{1}{(4\pi)^m}\int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\alpha}\,|\vec{\alpha}\rangle\,\langle\vec{\alpha}| = \mathbb{1}, \tag{65}$$

with $\mathbb{1}$ the identity operator. We can then show that any observable can be written as [5,6]

$$\hat{A} = \frac{1}{(4\pi)^m}\int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\alpha} P_A(\vec{\alpha})\,|\vec{\alpha}\rangle\,\langle\vec{\alpha}|\,, \tag{66}$$

where we refer to $P_A(\vec{\alpha})$ as the P function of the observable $\hat{A}$. Similarly, we can represent a density of operator $\hat{\rho}$ by its P function $P(\vec{\alpha})$. The reader should be warned that P functions often have rather unpleasant mathematical properties. In particular, they often are not actual functions and can be highly singular.

The P function naturally comes with a dual representation that is known as the Q function. As often in quantum physics, what actually counts is the expectation value of an observable in a specific state. We can use the P function to write

$$\mathrm{tr}[\hat{A}\hat{\rho}] = \frac{1}{(4\pi)^m}\int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\alpha} P_A(\vec{\alpha})\,\langle\vec{\alpha}|\,\hat{\rho}\,|\vec{\alpha}\rangle \tag{67}$$

$$= \frac{1}{(4\pi)^m}\int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\alpha} P(\vec{\alpha})\,\langle\vec{\alpha}|\,\hat{A}\,|\vec{\alpha}\rangle\,. \tag{68}$$

This naturally introduces the Q function, given by

$$Q_A(\vec{\alpha}) = \frac{1}{(4\pi)^m}\,\langle\vec{\alpha}|\,\hat{A}\,|\vec{\alpha}\rangle\,, \tag{69}$$

and, in particular, for the quantum state $\hat{\rho}$ we find that

$$Q(\vec{\alpha}) = \frac{1}{(4\pi)^m}\,\langle\vec{\alpha}|\,\hat{\rho}\,|\vec{\alpha}\rangle\,. \tag{70}$$

The latter is of particular interest because it represents the quantum state $\hat{\rho}$ as an actual probability distribution. This leads us to the general identity that

$$\mathrm{tr}[\hat{A}\hat{B}] = \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\alpha} P_A(\vec{\alpha})Q_B(\vec{\alpha}) = \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\alpha} Q_A(\vec{\alpha})P_B(\vec{\alpha}). \tag{71}$$

Thus finishing our introduction to the various descriptions of the quantum states and observables of bosonic many-particle systems.

The Q function has a clear physical interpretation. It is directly proportional to the fidelity of the state $\hat{\rho}$ with respect to a target coherent state $|\vec{\alpha}\rangle$. Furthermore, it is always positive, which implies that it is a well-defined probability distribution. Because we can write $\langle\vec{\alpha}|\,\hat{\rho}\,|\vec{\alpha}\rangle = \mathrm{tr}[\hat{\rho}\,|\vec{\alpha}\rangle\,\langle\vec{\alpha}|]$, we can use Eqs. (43) and (52) to express the Q function in terms of the Wigner function as

$$Q(\vec{\alpha}) = \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\, W(\vec{x})W_0(\vec{x}-\vec{\alpha}). \tag{72}$$

To satisfy both Eqs. (43) and (71), we find that

$$W(\vec{x}) = \frac{1}{(4\pi)^m}\int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\alpha}\; P(\vec{\alpha})W_0(\vec{x}-\vec{\alpha}), \tag{73}$$

$$= \frac{1}{(4\pi)^m}\int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\alpha}\; P(\vec{\alpha})\frac{e^{-\frac{1}{2}\|\vec{x}-\vec{\alpha}\|^2}}{(2\pi)^m}. \tag{74}$$

In turn, this implies that

$$Q(\vec{\alpha}) = \frac{1}{(8\pi)^m} \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\beta}\, P(\vec{\beta}) \frac{e^{-\frac{1}{4}\|\vec{\beta}-\vec{\alpha}\|^2}}{(2\pi)^m}. \tag{75}$$

These results thus show that all these phase-space representations are ultimately related to one another through convolution or deconvolution with a Gaussian [recall that the Wigner function of the vacuum Eq. (52) is a Gaussian distribution on phase space]. One can now follow Ref. [69] to define a continuous family of phase-space representations $\mathcal{W}_\sigma$ for $\sigma \in [-1, 1]$

$$\mathcal{W}_\sigma(\vec{\alpha}) = \left(\frac{1}{4\pi[1-\sigma]}\right)^m \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\beta}\, P(\vec{\beta}) \frac{e^{-\frac{1}{2[1-\sigma]}\|\vec{\beta}-\vec{\alpha}\|^2}}{(2\pi)^m}, \tag{76}$$

where we convolute the P function with an ever-increasing Gaussian, smoothening its features. We can then see that

$$\mathrm{tr}[\hat{A}\hat{\rho}] = (4\pi)^m \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\alpha}\, \mathcal{W}_{A,-\sigma}(\vec{\alpha}) \mathcal{W}_\sigma(\vec{\alpha}). \tag{77}$$

We find, notably, that $W(\vec{x}) = \mathcal{W}_{\sigma=0}(\vec{x})$, $Q(\vec{\alpha}) = \mathcal{W}_{\sigma=-1}(\vec{\alpha})$, and $P(\vec{\alpha}) = (4\pi)^m \mathcal{W}_{\sigma=1}(\vec{\alpha})$. This shows that the phase-space representation becomes more regular when decreasing σ. Other generalized probability distributions have been considered in the literature [70,71], often to circumvent the unappealing properties of the P function. In this Tutorial, we mainly use the Wigner function and (to a lesser extent) the Q function, as they are suitable tools to classify non-Gaussian quantum states. The P function is often used in the literature to characterize the nonclassicality of a state, where the intuition is that classical light is a mixture of coherent states (and thus its P function is a probability distribution) [5,6].

Before we close this introductory section on the phase-space description of CV quantum systems, we introduce one final tool that often comes in handy. The Wigner function can itself be obtained as the expectation value of an operator [72]. Formally, we write

$$W(\vec{x}) = \frac{1}{(2\pi)^m} \mathrm{tr}[\hat{\rho}\hat{\Delta}(\vec{x})]. \tag{78}$$

Using linearity and Eq. (38), we obtain the special case

$$\hat{\Delta}(\vec{0}) = \frac{1}{(2\pi)^m} \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{\lambda}\, e^{i\hat{q}(\vec{\lambda})}. \tag{79}$$

By using techniques based on Eqs. (54) and (55), we can show that

$$\hat{\Delta}(\vec{0})\hat{q}(\vec{f})\hat{\Delta}(\vec{0}) = -\hat{q}(\vec{f}). \tag{80}$$

This means that $\hat{\Delta}(\vec{0})$ is the parity operator. Its eigenstates are the Fock states, since

$$\hat{\Delta}(\vec{0})a^\dagger(\vec{f}_1)\ldots a^\dagger(\vec{f}_n)\,|0\rangle = (-1)^n a^\dagger(\vec{f}_1)\ldots a^\dagger(\vec{f}_n)\,|0\rangle\,. \tag{81}$$

Thus we can formally identify

$$\hat{\Delta}(\vec{0}) = (-\mathbb{1})^{\hat{N}}, \tag{82}$$

where $\hat{N}$ is the number operator. We define this operator by introducing a mode basis $\{\vec{e}_1, \Omega\vec{e}_1, \ldots, \vec{e}_m, \Omega\vec{e}_m\}$ of the optical phase space, such that $\hat{N} := \sum_{j=1}^m a^\dagger(\vec{e}_j)a(\vec{e}_j)$. This definition can be combined with the properties of the displacement operator to obtain that

$$\hat{\Delta}(\vec{x}) = \hat{D}(-\vec{x})(-\mathbb{1})^{\hat{N}}\hat{D}(\vec{x}). \tag{83}$$

Note that what we just obtained is the operator equivalent of a δ function, which becomes even more explicit when we explicitly write down its Wigner representation

$$W_{\Delta(\vec{x}')}(\vec{x}) = \frac{1}{(4\pi)^m}\delta(\vec{x}-\vec{x}'), \tag{84}$$

which follows directly from Eqs. (43) and (78).

This result may seem somewhat artificial, but it turns out to be extremely useful. The observable $\hat{\Delta}(\vec{x})$ can be measured experimentally by counting photons, which means that the combination of photon counting and displacements directly allows us to reconstruct the Wigner function of the quantum state [73]. Until recently, the lack of good photon-number-resolving detectors in the optical frequency range has long made this method unfeasible for most states. Even though there was an early demonstration of the method for coherent states [74], it is only due to recent developments in detector technologies that the method can be applied to more general states [75,76]. The idea was also applied in other settings [77], and was used in pioneering CV experiments with trapped ions, such as Ref. [78], and in cavity QED [79,80].

C. Discrete and continuous variables

In Sec. A, we have introduced a many-boson system, regardless of the physical realization of these bosons. Such a many-boson system and its Fock space are built upon the structure that is determined by the single-particle Hilbert space $\mathcal{H}$. The Fock space that is constructed accordingly has a rich structure that is further explored in the Tutorial [56]. In optics, the bosons that we consider are photons, and quantum optics can thus be seen as the theory of a many-boson system in the context of Sec. A. This approach to quantum optics is referred to as the DV approach.

In Sec. B, we contrast this with the CV approach to quantum optics. This approach relies on the measurement of the field quadratures, and can thus be seen as a bosonic quantum field theory. Therefore we started this approach by introducing the classical electric field and its modes, which we subsequently quantized through canonical quantization. We introduced the notion of optical phase space as a general way of describing CV quantum systems. The phase space is directly related to the modes of the field and manipulations of the modes also cause changes in the optical phase space. Nevertheless, any system with a phase space can be described by these techniques.

Both of these approaches are ultimately equivalent. Bosonic creation and annihilation operators describe the same algebra of observables as bosonic quadrature operators, which means that on the level of mathematical structure, both approaches can be interchanged and even mixed. This is strikingly clear when the Wigner function, i.e., the phase-space representation of quantum states and observables that is most naturally associated with field quadratures, turns out to be directly measurable by counting photons. Notably, this implies that when it comes to mathematical structures, bosonic particles such as atoms can also be described on phase space.

The real difference between CV and DV approaches is of an experimental nature. What is important is not the observables that are technically present in the quantum system, but the observables that are practically measured in the lab. For the CV approach we typically use homodyne detection to measure quadratures [7,81], whereas in DV approaches we count photons [82].

A common source of misunderstanding between the DV and CV community stems from the role they attribute to the single-particle Hilbert space and optical phase space, respectively. As we argued, both spaces are (at least for a finite-dimensional number of modes) isomorphic, see the Appendix for some additional mathematical intuition. However, the Hilbert space of a photon, which is inherently a quantum particle, is often interpreted as a quantum object. At the same time, the optical phase space represents the field quadratures of optical modes and is thus rather considered to be a classical object. The origin of this confusion lies in the fact that the optical modes, i.e., normalized solutions of Maxwell's equations, also form a Hilbert space that has its origins entirely in classical physics.

The optical modes are the vessels that contain photons much in the same way as a set of electrons contains spins. The crucial difference is that optical modes are not uniquely defined, we can manipulate them, transform them from one mode basis to another with an interferometer and thus consider new superpositions of modes. In typical experimental settings, one would not consider a superposition of two electrons a new well-defined electron.

Because creation and displacement operators always act in one specific mode (i.e., they are generated by a single vector on the single-photon Hilbert space), single-photon states and coherent states are always single-mode states. We may expand this single mode in a different mode basis, which can even be done physically by sending the state through a beam splitter, to create some form of entanglement in the quantum states. However, this entanglement is just a manifestation of the fact that we are not considering the optimal mode basis. In the CV approach, this has led to the notion of "intrinsic" properties [58], which are those properties of quantum states that are independent of the chosen mode basis. The purity and entropy of a state are notable examples, but one can also introduce a notion of "intrinsic entanglement" to refer to a state that is entangled in any possible mode basis. In the next subsection, we introduce Gaussian states, which will later be shown to never be intrinsically entangled.

D. Gaussian states

Now that we have introduced phase-space representations for states and observables of CV quantum systems, we still need one building block before we can tackle multimode non-Gaussian states: a good understanding of Gaussian states. It is not the goal of this subsection to delve deep into decades worth of research on Gaussian states. We rather highlight a few key results that set apart Gaussian quantum states from the rest of the vast states' space. For more extended reviews, we refer the reader to Refs. [8,9]. These states are also extensively studied in the mathematical physics literature under the name "quasifree states of the CCR algebra."

Gaussian states are by definition states that have a Wigner function, which is a Gaussian:

$$W_G(\vec{x}) = \frac{e^{-\frac{1}{2}(\vec{x}-\vec{\xi})^T V^{-1}(\vec{x}-\vec{\xi})}}{(2\pi)^m\sqrt{\det V}}, \tag{85}$$

where $\vec{\xi}$ is referred to as the mean field (or displacement) and V is known as the covariance matrix. With Eq. (40), we can verify that the mean field indeed corresponds to the expectation value of the field quadrature

$$\mathrm{tr}[\hat{\rho}\hat{q}(\vec{f})] = \vec{\xi}^T\vec{f}\,, \tag{86}$$

similarly, we find for the covariance matrix

$$\begin{aligned}&\mathrm{tr}[\hat{\rho}\hat{q}(\vec{f}_1)\hat{q}(\vec{f}_2)] - \mathrm{tr}[\hat{\rho}\hat{q}(\vec{f}_1)]\mathrm{tr}[\hat{\rho}\hat{q}(\vec{f}_2)]\\ &= \vec{f}_1^{\,T} V \vec{f}_2 - i\vec{f}_1^{\,T}\Omega\vec{f}_2.\end{aligned} \tag{87}$$

These quantities can thus be obtained for arbitrary quantum states, but for Gaussian states the covariance matrix and the mean field also determine all higher-order expectation values. The most elegant way to see this is via the

multivariate cumulants (also known as truncated correlation functions), which vanish beyond order two [4]. This fact implies that all properties of Gaussian states can ultimately be deduced from their mean field and—often more importantly—from their covariance matrix.

Hitherto, we have encountered the vacuum state $|0\rangle$ and the coherent states $|\vec{\alpha}\rangle$ as examples of Gaussian states. Both of the examples have a covariance matrix $V = \mathbb{1}$. However, there is a much larger range of possible covariance matrices available and they have to satisfy certain constraints [83]. At first instance, we note that a covariance matrix must be positive. An additional constraint is obtained by imposing that the variance $\Delta^2\hat{q}(\vec{f}) \geqslant 0$ for all $\vec{f}$ in phase space. Equation (87) then directly yields that $\vec{f}^{\,T}(V - i\Omega)\vec{f} \geqslant 0$, which implies that $V \geqslant 0$ and suggests that $(V - i\Omega) \geqslant 0$. However, the latter is not obvious, since $\vec{f}$ are real vectors, whereas $(V - i\Omega)$ is a complex matrix. We thus need an additional ingredient: the Heisenberg inequality. Formally, this inequality can be obtained through Robertson's more general inequality [84], such that we find

$$\Delta^2\hat{q}(\vec{f}_1)\Delta^2\hat{q}(\vec{f}_2) \geqslant \frac{1}{4}\left|\mathrm{tr}\{\hat{\rho}[\hat{q}(\vec{f}_1), \hat{q}(\vec{f}_2)]\}\right|^2. \tag{88}$$

We can now apply the CCR (29) to obtain the general form

$$\Delta^2\hat{q}(\vec{f}_1)\Delta^2\hat{q}(\vec{f}_2) \geqslant \left|\vec{f}_1^{\,T}\Omega\vec{f}_2\right|^2. \tag{89}$$

On the other hand, the definition (87) of the covariance matrix can be used to translate this result to

$$\vec{f}_1^{\,T}V\vec{f}_1\vec{f}_2^{\,T}V\vec{f}_2 \geqslant \left|\vec{f}_1^{\,T}\Omega\vec{f}_2\right|^2. \tag{90}$$

This identity can then be used to prove $(\vec{f}_1^{\,T} - i\vec{f}_2^{\,T})(V - i\Omega)(\vec{f}_1 + i\vec{f}_2) \geqslant 0$ for all $\vec{f}_1, \vec{f}_2 \in \mathbb{R}^{2m}$. As a consequence, we find that

$$V - i\Omega \geqslant 0, \tag{91}$$

an important constraint on the covariance matrix V, which can be understood as combining the positivity conditions and the Heisenberg inequality.

To further understand the structure of covariance matrices and the Gaussian states that they describe, we highlight some important results on symplectic matrices. The first of these results is Williamson's decomposition [10], which states that any positive-definite real matrix V can be diagonalized by a symplectic matrix S (i.e., a matrix with $S^T\Omega S = \Omega$):

$$V = S^T N S, \text{ with } N = \mathrm{diag}[\nu_1, \nu_1, \nu_2, \nu_2, \ldots, \nu_m, \nu_m]. \tag{92}$$

The values $\nu_1, \ldots \nu_m$ are also known as the symplectic spectrum of V. From Heisenberg's relation, we then find the additional constraint that $\nu_1, \ldots, \nu_m \geqslant 1$, in other words, the values in the symplectic spectrum are larger than shot noise. It now becomes straightforward to see the Heisenberg's relation also implies that

$$\det V \geqslant 1. \tag{93}$$

It thus becomes apparent that the Heisenberg inequality is saturated when $\det V = 1$. The states for which this is the case must have a covariance matrix $V = S^T S$.

The Gaussian states for which the Heisenberg inequality is saturated turn out to be the pure Gaussian states. Recall that the purity of a quantum state is given by $\mu = \mathrm{tr}[\hat{\rho}^2]$. This quantity can be directly calculated from the Wigner function via Eq. (43). We then find for an arbitrary Gaussian state

$$\mu_G = (4\pi)^m \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\, W_G(\vec{x})^2 = \frac{1}{\sqrt{\det V}}. \tag{94}$$

Alternatively, we may use the symplectic spectrum to express $\mu_G = \prod_{k=1}^m \nu_k^{-1}$. This shows us that a Gaussian state is pure if and only if its covariance matrix is a positive symplectic matrix, i.e., it can be written as $V = S^T S$.

The class of states with a covariance matrix given by $V = S^T S$ is much larger than just the vacuum and coherent states with $V = \mathbb{1}$. The additional states turn out to have asymmetric noise in their quadratures, and because the Heisenberg inequality is saturated this implies that some quadratures have less noise than the vacuum state. The states with such covariance matrices are therefore known as squeezed states. To formalize this intuition, we consider the Bloch-Messiah decomposition (which is known in mathematics and classical mechanics as Euler's decomposition) [11,85]. Any symplectic matrix S can be decomposed as follows:

$$S = O_1 K O_2, \text{ with } K = \mathrm{diag}[s_1^{1/2}, s_1^{-1/2}, \ldots, s_m^{1/2}, s_m^{-1/2}], \tag{95}$$

where O_1 and O_2 are orthogonal symplectic matrices, i.e., $O_j^T O_j = \mathbb{1}$ and $O_j^T \Omega O_j = \Omega$. We can then see that for any pure Gaussian state, we find

$$V = S^T S = O^T K^2 O. \tag{96}$$

We have already encountered orthogonal symplectic transformations in Eq. (21), where we associated them with transformations of mode bases. Thus, if we find a set of optical modes that are prepared in a pure Gaussian state, we can always find a different mode basis in which the state

is given by

$$V' = O\,VO^T = \begin{pmatrix} s_1 & & & & \\ & 1/s_1 & & & \\ & & \ddots & & \\ & & & s_m & \\ & & & & 1/s_m \end{pmatrix}. \quad (97)$$

This means that we can always find a set of symplectic eigenvectors $\{\vec{e}_1, \Omega\vec{e}_1, \ldots, \vec{e}_m, \Omega\vec{e}_m\}$ of a pure Gaussian state's covariance matrix, which have the properties that $\Delta^2\hat{q}(\vec{e}_j) = s_j$ and $\Delta^2\hat{q}(\Omega\vec{e}_j) = 1/s_j$, such that the Heisenberg relation is saturated: $\Delta^2\hat{q}(\vec{e}_j)\Delta^2\hat{q}(\Omega\vec{e}_j) = 1$. At the same time, we find that clearly either $\Delta^2\hat{q}(\vec{e}_j)$ of $\Delta^2\hat{q}(\Omega\vec{e}_j)$ is smaller than one (and thus below shot noise).

Gaussian states naturally come with the notion of Gaussian channels [86], they are the completely positive trace-preserving transformations that map Gaussian states into other Gaussian states. We have already seen that the displacement operators are unitary transformations that fulfil this condition. Because any Gaussian transformation Γ preserves the general shape of the Wigner function (85), we can simply describe the Gaussian channel Γ in terms of its actions on the mean field and the covariance matrix:

$$V \overset{\Gamma}{\mapsto} XVX^T + V_c, \quad (98)$$

$$\vec{\xi} \overset{\Gamma}{\mapsto} X\vec{\xi} + \vec{\alpha}. \quad (99)$$

The vector $\vec{\alpha}$ simply serves to displace the entire Gaussian to a different location in phase space. On the level of the covariance matrix, X transforms and reshapes the initial covariance matrix, whereas V_c describes the addition of Gaussian classical noise. Both can *a priori* be any real matrices, as long as they satisfy the constraint

$$V_c - i\Omega + iX\Omega X^T \geqslant 0. \quad (100)$$

This constraint derives from the demand that $XVX^T + V_c$ is a well-defined covariance matrix, and therefore $XVX^T + V_c - i\Omega \geqslant 0$. Because V is a well-defined covariance matrix, $X(V - i\Omega)X^T \geqslant 0$ and thus it can be seen that $XVX^T + V_c$ is also a well-defined covariance matrix whenever Eq. (100) holds. This simple argument proves Eq. (100) is a sufficient condition for Γ to transform the covariance matrix of the initial state into a new bona fide covariance matrix.

An important case is obtained when we impose that Γ conserves the purity of the state and is thus a unitary transformation. It then immediately follows that $V_c = 0$, since there cannot be any classical noise. The displacement $\vec{\alpha}$ is simply implemented by a displacement operator, and the constraint (100), combined with the demand that purity is conserved implies that X is a symplectic matrix. In other words, a Gaussian unitary transformation $\hat{U}_G$ satisfies $V \mapsto S^TVS$. Another relevant example is the case of uniform Gaussian losses, where we set $X = \sqrt{1-\eta}\mathbb{1}$, $V_c = \eta\mathbb{1}$ and $\vec{\alpha} = 0$, with the positive value $\eta \leqslant 1$ denoting the amount of loss.

More generally, the action of a Gaussian channel on an arbitrary state can be understood from its action on $\exp[i\hat{q}(\vec{\lambda})]$, which can be proven to take the form

$$\exp[i\hat{q}(\vec{\lambda})] \overset{\Gamma}{\mapsto} \exp\left[i\hat{q}(X^T\vec{\lambda}) + i\vec{\alpha}^T\vec{\lambda} - \frac{1}{2}\vec{\lambda}^T V_c\vec{\lambda}\right]. \quad (101)$$

We can then calculate the quantum characteristic function and use some properties of Fourier transforms to find that the Wigner function transforms as

$$W(\vec{x}) \overset{\Gamma}{\mapsto} \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{y}\; W(X^{-1}\vec{x} - \vec{y}) \frac{e^{-\frac{1}{2}(\vec{y}-\vec{\alpha})^T V_c^{-1}(\vec{y}-\vec{\alpha})}}{(2\pi)^m\sqrt{\det V_c}}. \quad (102)$$

For Gaussian unitary transformations we find the appealing result that $W(\vec{x}) \mapsto W[S^{-1}(\vec{x} - \vec{\alpha})]$. This means that a Gaussian unitary transformation is simply a coordinate transformation on phase space.

Proving that any completely positive Gaussian channel Γ is of the form Eq. (102) with condition (100), is a challenging task. The result was first obtained in Refs. [87,88], using the language of C^* algebras. The proof is rather technical and we do not go into details here.

The paradigm of Gaussian channels is also useful to structure general Gaussian states. One may, for example, wonder which Gaussian channel would transform the vacuum state into the Gaussian state with covariance matrix V. In general, there is no unique solution to this question, but there is a straightforward route to find an answer. First, take any symplectic matrix S that satisfies $V - S^TS \geqslant 0$ (the Williamson decomposition guarantees that this is always possible). This implies that there is a positive-definite matrix V_c such that $V = S^TS + V_c$. As such, an arbitrary Gaussian state can always be decomposed as

$$W_G(\vec{x}) = \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{y}\; W_0(S^{-1}\vec{x} - \vec{y}) e^{-\frac{1}{2}(\vec{y}-\vec{\alpha})^T V_c^{-1}(\vec{y}-\vec{\alpha})}. \quad (103)$$

The symplectic operation that is applied to the vacuum is known as multimode squeezing in optics. These transformations are fully equivalent to Bogoliubov transformations that are regularly used in condensed-matter physics [4,8]. Combined with a displacement, this operation provides the most general operation that maps quadrature operators into well-defined quadrature operators.

Now that we have introduced the basic concepts of Gaussian states, we are equipped to start exploring their

non-Gaussian counterparts. Several other important properties of Gaussian states will be introduced along the way to stress just how peculiar these Gaussian states are compared to the rest of state space.

Figure 1 provides an overview that can be used as a brief guide to understand the structure of non-Gaussian states. We attempt to highlight how the different quantities used to structure the non-Gaussian part of state space are interconnected.

III. NON-GAUSSIAN QUANTUM STATES

Contrary to Gaussian states with their elegant Wigner function and properties that can be nicely deduced from the covariance matrix, the set of non-Gaussian states is vast and wild. Literally all states with Wigner functions that are not Gaussian are contained within this class. To give an idea of the enormous variety, one can consider that highly exotic states such as Gottesman-Kitaev-Preskill states [31] and Schrödinger cat states inhabit the set of non-Gaussian states together with the states that describe single photons and even certain convex mixtures of Gaussian states. Throughout the years, there have been considerable efforts to structure the set of non-Gaussian states. We introduce the notion of quantum non-Gaussian states [89] and then extend it to a hierarchy based on stellar rank [38]. A different approach is provided by considering that the negativity of the Wigner function can be used as a genuine signature of nonclassicality [90]. However, before we attack these different measures to structure non-Gaussian quantum states, we contrast some properties of Gaussian and non-Gaussian states.

A. Gaussian versus Non-Gaussian

Gaussian states have many extraordinary properties that set them apart from non-Gaussian states. First of all, pure Gaussian states turn out to be the only quantum states that saturate the uncertainty relation. The easiest way to see this is by describing arbitrary pure states in terms of their wave functions. The wave functions associated with amplitude quadratures $\hat{q}(\vec{f})$ and those associated with phase quadratures $\hat{q}(\Omega\vec{f})$ are related by a Fourier transform. This fact can then be used to show that only Gaussian wave functions saturate the Heisenberg inequality. The extension to arbitrary mixed states can be achieved via Jensen's inequality, which emphasizes that no mixed states can saturate the uncertainty relation. Let us consider a mixed state $\hat{\rho} = \sum_k p_k |\Psi_k\rangle \langle\Psi_k|$ with variances $\Delta^2\hat{q}(\vec{f})$. We also introduce the variances $\Delta_k^2\hat{q}(\vec{f})$ for the pure states $|\Psi_k\rangle$. From Jensen's inequality [91], it follows that

$$\Delta^2\hat{q}(\vec{f}) \geqslant \sum_k p_k \Delta_k^2\hat{q}(\vec{f}). \tag{104}$$

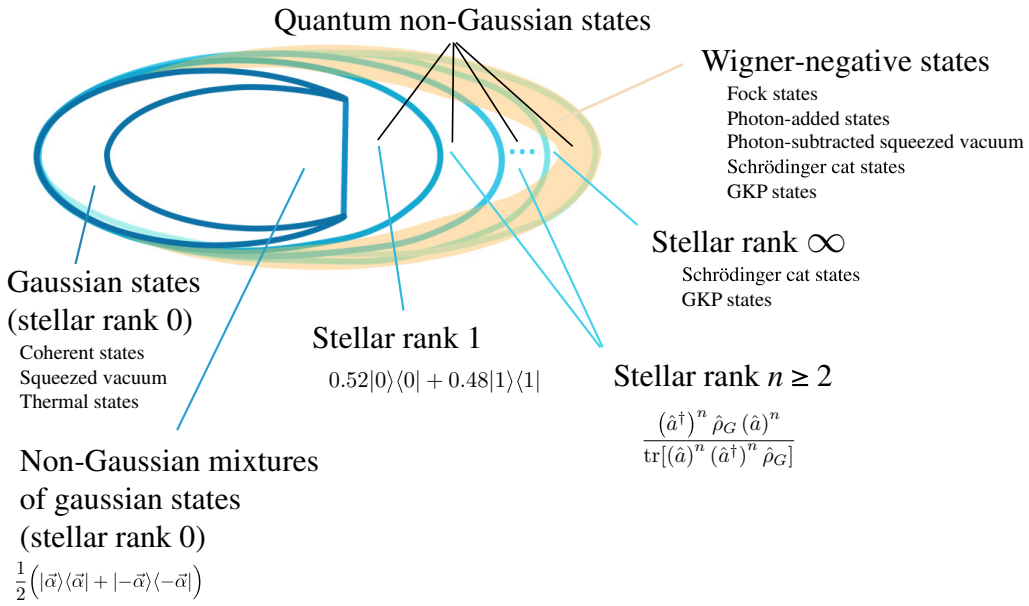

FIG. 1. Overview of the different types of non-Gaussian states that can be found in state space. The different aspects will all be considered throughout Sec. III. Here we attempt to show the stellar hierarchy and how it differentiates itself from the convex hull (mixtures) of Gaussian states. Furthermore, we emphasize that the stellar rank and Wigner negativity are different quantifiers of non-Gaussianity. It should be noted that all non-Gaussian pure states are Wigner-negative states, but we can find states that are not mixtures of Gaussian states without Wigner negativity. For all the classes, we provide examples of states that belong to this group, the Wigner functions for several of these examples are shown in Fig. 2.

For Heisenberg's inequality, we calculate

$$\Delta^2\hat{q}(\vec{f})\Delta^2\hat{q}(\Omega\vec{f}) \geqslant \sum_k p_k^2 \Delta_k^2\hat{q}(\vec{f})\Delta_k^2\hat{q}(\Omega\vec{f}) + \sum_{k\neq l} p_k p_l \Delta_k^2\hat{q}(\vec{f})\Delta_l^2\hat{q}(\Omega\vec{f}) \geqslant 1. \tag{105}$$

The presence of cross terms highlights that even when all the pure states in the mixture saturate the inequality, the mixture does not. The only possible exception is the case where the state is pure.

That only pure Gaussian states saturate the Heisenberg inequality may seem like an innocent observation, but it has an important implication for non-Gaussian states. The Heisenberg inequality can be formulated entirely in terms of the covariance matrix. We showed in Eq. (92) that the inequality is saturated if and only if the covariance matrix is symplectic, i.e., $V = S^T S$. Furthermore, we showed in Eq. (94) that a Gaussian state is pure if and only if its covariance matrix is symplectic $V = S^T S$. The fact that no non-Gaussian states can saturate the inequality thus implies that non-Gaussian states can never have a symplectic covariance matrix $V = S^T S$. This is a first hint of the special role played by Gaussian states.

A more general result along these lines states that for all states $\hat{\rho}$ with the same covariance matrix V, the Gaussian state always has the highest von Neumann entropy [13]. First of all, note that entropy $-\mathrm{tr}[\hat{\rho}\log\hat{\rho}]$ is conserved under unitary transformations. Due to the Williamson decomposition (92), we can write any Gaussian state as

$$\hat{\rho}_G = \hat{U}_G \bigotimes_{j=1}^{m} \hat{\rho}_{\overline{n}_j} \hat{U}_G^\dagger, \tag{106}$$

where $\hat{\rho}_{\overline{n}_j}$ is a thermal state of the Hamiltonian $\hat{a}_j^\dagger \hat{a}_j$ with average particle number $\overline{n}_j = (\nu_j - 1)/2$. From statistical mechanics, we know that thermal states are the quantum states that maximize the von Neumann entropy for a given temperature (here fixed by the occupations $\overline{n}_j$).

It turns out that Gaussian states are limiting cases for many quantities [12]. This result shows that for a range of functionals f on the state space, we find that $f(\hat{\rho}) \geqslant f(\hat{\rho}_G)$, where $\hat{\rho}_G$ is the Gaussian state with the same covariance matrix as $\hat{\rho}$. Apart from some more technical aspects such as continuity, f must have two important features: it must be conserved under (a certain class of) unitary operations $f(\hat{U}\hat{\rho}\hat{U}^\dagger) = f(\hat{\rho})$ and it must be strongly superadditive $f(\hat{\rho}) \geqslant f(\hat{\rho}_1) + f(\hat{\rho}_2)$ (note that $\hat{\rho}_1$ and $\hat{\rho}_2$ are marginals of $\hat{\rho}$). The equality must be saturated for product states, i.e., $f(\hat{\rho}_1 \otimes \hat{\rho}_2) = f(\hat{\rho}_1) + f(\hat{\rho}_2)$. For strongly subadditive functions with $f(\hat{\rho}) \leqslant f(\hat{\rho}_1) + f(\hat{\rho}_2)$ the same result implies $f(\hat{\rho}) \leqslant f(\hat{\rho}_G)$, (after all, in that case $-f$ is a strongly superadditive function). It is clear that the von Neumann entropy fulfils the latter conditions and is maximized for Gaussian states. For superadditive entanglement measures, this result can be used to show that for all states with the same covariance matrix, Gaussian states are the least entangled ones (entanglement is much more extensively discussed in Sec. V). However, several common entanglement measures, e.g., the logarithmic negativity [92] and the entanglement of formation [93], are not superadditive.

At the heart of these extremal properties lies the central limit theorem [4,94–96]. There are many versions of the central limit theorem in quantum physics, but we stick to what is probably the simplest one. As always, we consider our optical phase space $\mathbb{R}^{2m}$, but this time, we take N copies of it, which implies that we are dealing with a phase space $\mathbb{R}^{2Nm} = \mathbb{R}^{2m} \oplus \cdots \oplus \mathbb{R}^{2m}$ for the full system. We can then embed a vector $\vec{\lambda} \in \mathbb{R}^{2m}$ in the jth of these N copies via $\vec{\lambda}_j := \vec{0} \oplus \cdots \oplus \vec{0} \oplus \vec{\lambda} \oplus \vec{0} \cdots \oplus \vec{0}$ and introduce the new averaged operator

$$\overline{q}_N(\vec{\lambda}) := \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \hat{q}(\vec{\lambda}_j). \tag{107}$$

It is rather straightforward to see that these observables follow the canonical commutation relation. We can now restrict ourselves to studying the algebra that is generated entirely by such averaged quadrature operators. When we then assume that the different copies of the system are "independently and identically distributed" we must set the overall state to be $\hat{\rho}^{(N)} = \hat{\rho}^{\otimes N}$. We then find the characteristic function of the algebra of averaged observables by

$$\chi_N(\vec{\lambda}) = \mathrm{tr}[\hat{\rho}^{\otimes N} e^{i\overline{q}_N(\vec{\lambda})}]. \tag{108}$$

The following pointwise convergence can be shown:

$$\chi_N(\vec{\lambda}) \overset{N\to\infty}{\to} \chi_G(\vec{\lambda}), \tag{109}$$

where $\chi_G(\vec{\lambda})$ is the characteristic function of the Gaussian state $\hat{\rho}_G$ that has the same covariance matrix as $\hat{\rho}$. This means that the non-Gaussian features in any state $\hat{\rho}$ can be coarse grained away by averaging sufficiently many copies of the state. Note that this result considers N copies of an arbitrary m-mode state. The single-mode version of this result was proven in Ref. [94], whereas a much more general versions are derived in Refs. [95,96]. In Ref. [12] the central limit theorem is combined with invariance under local unitary transformations to prove the final extremality result, we do not review these points in detail.

The extremality of Gaussian states and the associated central limit theorem highlight why Gaussian states are important in quantum-information theory and quantum

statistical mechanics. It also shows that Gaussian states have some particular properties compared to non-Gaussian states. It thus should not come as a surprise that some of these properties can be used to measure the degree of non-Gaussianity of the state [97–99]. As we mentioned before, for a fixed covariance matrix V the von Neumann entropy is maximized by the Gaussian state. This suggest that we can use the difference in von Neumann entropy as a measure for non-Gaussianity. To formalize things, let us consider an arbitrary state $\hat{\rho}$ with covariance matrix V and mean field $\vec{\xi}$ (this quantities can be derived, respectively, for the second and first moments of the quadrature operators). We then construct a Gaussian state $\hat{\sigma}_V$, which has the same covariance matrix and the mean field. In the spirit of extremality, we then define

$$\delta(\hat{\rho}) = S(\hat{\sigma}_V) - S(\hat{\rho}), \tag{110}$$

where $S(\hat{\rho}) := -\mathrm{tr}[\hat{\rho}\log\hat{\rho}]$. Because von Neumann entropy is constant under unitary transformations, for a Gaussian state it depends only on the symplectic spectrum $\nu_1, \ldots, \nu_m$. In other words, we can calculate $S(\hat{\sigma}_V)$ directly by using the Williamson decomposition (92) on V. We find from Ref. [13] that

$$S(\hat{\sigma}_V) = \sum_{j=1}^{m}\left[\frac{\nu_j+1}{2}\log\frac{\nu_j+1}{2} - \frac{\nu_j-1}{2}\log\frac{\nu_j-1}{2}\right]. \tag{111}$$

However, it should be noted that the entropy of the non-Gaussian states $S(\hat{\rho})$ is generally harder to calculate unless we can accurately approximate the state by a finite density matrix in the Fock basis. Furthermore, if the state $\hat{\rho}$ is pure, we simply find that $\delta(\hat{\rho}) = S(\hat{\sigma}_V)$.

Due to extremality of Gaussian states it directly follows that $\delta(\hat{\rho}) \geqslant 0$, but this does not necessarily mean that $\delta(\hat{\rho})$ is a good measure for non-Gaussianity. References [13,98] establish that

$$\delta(\hat{\rho}) = S(\hat{\rho} \mid |\hat{\sigma}_V), \tag{112}$$

where $S(\hat{\rho} \mid\mid \sigma_V) := \mathrm{tr}[\hat{\rho}(\log\hat{\rho} - \log\hat{\sigma}_V)]$ is the quantum relative entropy between $\hat{\rho}$ and reference state $\hat{\sigma}_V$. The quantum relative entropy allows us to connect $\delta(\hat{\rho})$ to a range of interesting properties, as shown in Ref. [98]. For example, it directly follow that $\delta(\hat{\rho}) = 0$ if and only if $\hat{\rho} = \hat{\sigma}_V$. Furthermore, the measure $\delta(\hat{\rho})$ inherits convexity and monotonicity properties from the relative entropy. These are exactly the properties that made this measure a useful ingredient in the resource theory for quantum non-Gaussianity presented in Ref. [100].

Thus, the connection between Eq. (110) and relative entropy shows that $\delta(\hat{\rho})$ can indeed be used as a measure for non-Gaussianity in the sense that it measures "entropic distance" between $\hat{\rho}$ and $\hat{\sigma}_V$. Yet, there is one important question that remains to be answered: is σ_V indeed the closest Gaussian state to $\hat{\rho}$? An affirmative answer to this question was provided in Ref. [101], where it was shown that

$$\delta(\hat{\rho}) = \min_{\hat{\rho}_G} S(\hat{\rho} \mid |\hat{\rho}_G), \tag{113}$$

where we minimise over all possible Gaussian states $\hat{\rho}_G$. The main idea of the proof is to show that $S(\hat{\rho} \mid\mid \hat{\rho}_G) - S(\hat{\rho} \mid\mid \hat{\sigma}_V) = S(\hat{\sigma}_V \mid\mid \hat{\rho}_G) \geqslant 0$ such that the smallest relative entropy in indeed achieved for $\hat{\sigma}_V$. For the technical details, we refer the interested reader to Ref. [101]. Furthermore, we note that a similar non-Gaussianity measure was introduced by using the Wehrl entropy (based on the Q function) rather than the von Neumann entropy [102].

We have thus shown that Gaussian states are special in the sense that they minimize entanglement and maximize entropy as compared to other states with the same covariance matrix. Another profound distinction can be found when comparing pure Gaussian states to pure non-Gaussian states. In this case, there is a seminal result by Hudson [34] that was extended by Soto and Claverie to multimode systems [35], which states that a pure state can have a non-negative Wigner function if and only if the state is Gaussian. In other words, all non-Gaussian pure states exhibit Wigner negativity.

Here, we follow the approach of Ref. [103] to prove this result. First of all, we introduce the function

$$F^{\star}_{\Psi}(\vec{\alpha}) := \langle\vec{\alpha} \mid \Psi\rangle e^{\frac{1}{8}\|\vec{\alpha}\|^2}, \tag{114}$$

such that the Q function (70) of the state $|\Psi\rangle$ is given by

$$Q(\vec{\alpha}) = \frac{1}{(4\pi)^m}\left|F^{\star}_{\Psi}(\vec{\alpha})\right|^2 e^{-\frac{1}{4}\|\vec{\alpha}\|^2}. \tag{115}$$

From Eq. (114), we directly find that

$$\left|F^{\star}_{\Psi}(\vec{\alpha})\right|^2 \leqslant e^{\frac{1}{4}\|\vec{\alpha}\|^2}. \tag{116}$$

Next, we observe that a Q function that reaches zero implies a negative Wigner function, which can be seen from Eq. (72). Thus, demanding that the state has a positive Wigner function implies demanding that $Q(\vec{\alpha}) > 0$, and thus that $F^{\star}_{\Psi}(\vec{\alpha})$ has no zeros. Using the equivalence between $2m$-dimensional phase space and a complex m-dimensional Hilbert space, we can use the multidimensional but restricted version of the Hadamard theorem [35], which states that any entire function $f : \mathbb{C}^m \mapsto \mathbb{C}$ without any zeros and with order of growth [104] r is an exponential $f(z) = \exp g(z)$, where $g(z)$ is a polynomial of degree $s \leqslant r$. We then note that $F^{\star}_{\Psi}$ is an entire function of maximal growth $r = 2$ as given by Eq. (116).

Hadamard's theorem then tells us that $F^\star_\Psi(\vec{\alpha})$ is Gaussian. In other words, if a pure quantum state $|\Psi\rangle$ has a positive Wigner function $F^\star_\Psi(\vec{\alpha})$ must be Gaussian. The only states for which this is the case are Gaussian states.

In summary, we have seen that Gaussian states inherit a particular extremal behavior from the central limit theorem. This should not come as a surprise given that they are the Gibbs states of a free bosonic field at finite temperature. Thus, a non-Gaussian state can be expected to have more "exotic" features than the Gaussian state with the same covariance matrix. This also formalizes the intuition that Gaussian states are more classical states. This idea is further established by the fact that all pure Gaussian states are the only possible pure states that have a positive Wigner function.

The fact that non-Gaussian states automatically have nonpositive Wigner functions no longer holds when mixed states are considered. A simple example is that state $\hat{\rho} = [|0\rangle\langle 0| + a^\dagger(\vec{f})|0\rangle\langle 0|a(\vec{f})]/2$, which is clearly non-Gaussian but also has a positive Wigner function. There have been considerable efforts to extend Hudson's theorem in some form to mixed states [36]. However, in what follows, we see that there are many ways for a state to be non-Gaussian. This makes it particularly hard to connect a measure such as Eq. (110) to more operational interpretations. In the next section, we start by showing some examples of different non-Gaussian states to make the reader appreciate their variety.

B. Examples of non-Gaussian states

An overview of the different examples discussed in this section is shown in Fig. 2.

A first important class of non-Gaussian states are Fock states, generated by acting with creation operators $a^\dagger(\vec{f})$ on the vacuum state

$$|n_f\rangle := \frac{1}{\sqrt{n!}}[\hat{a}^\dagger(\vec{f})]^n |0\rangle, \tag{117}$$

which is a state of n photons in mode f. These states are inherently single mode, even though they can be embedded in a much larger multimode state space. The Wigner function for such states is commonly found in quantum optics textbooks, but deriving it using Eq. (38) is a good exercise. Here we simply state the result:

$$W_{n_f}(\vec{x}) = \sum_{k=0}^{n} \binom{n}{k} \frac{(-1)^{n+k}\|\vec{x}_\mathbf{f}\|^{2k}}{k!} \frac{e^{-\frac{1}{2}\|\vec{x}\|^2}}{2\pi}. \tag{118}$$

In the most general sense, we write $\vec{x}_\mathbf{f} = (\vec{f}^{\,T}\vec{x})\vec{f} + (\vec{f}^{\,T}\Omega\vec{x})\Omega\vec{f}$ as the projector of $\vec{x}$ on the phase space of mode f, but it is most practical to use the coordinate representation $\vec{x}_\mathbf{f} = (x_f, p_f)^T$ where $x_f = \vec{f}^{\,T}\vec{x}$ and $p_f = \vec{f}^{\,T}\Omega\vec{x}$ to describe the two-dimensional phase space associated with mode f.

Multimode Fock states can be obtained by acting on the vacuum with different creation operators in different modes. Even though these different modes do not necessarily have to be orthogonal [56], we here focus on the case where they are. For example, in the m-mode system, we can chose a basis $\{\vec{e}_1, \Omega\vec{e}_1, \ldots, \vec{e}_m, \Omega\vec{e}_m\}$ of the phase

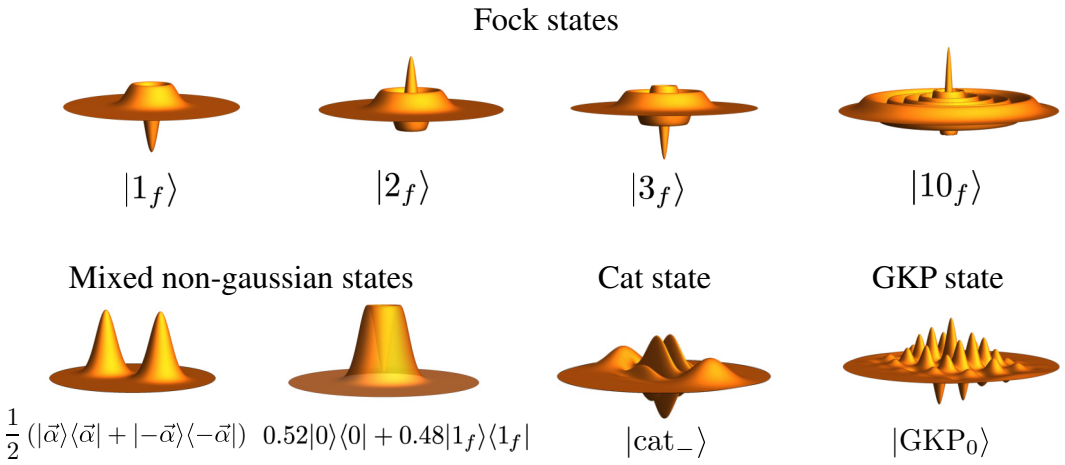

FIG. 2. Several examples of Wigner functions for single-mode non-Gaussian states. The Wigner functions for the Fock states are obtained from Eq. (118) and the mixture of a Fock state and a vacuum is given by Eq. (124). The mixture of coherent states is expressed in Eq. (122) where we set $\|\vec{\alpha}\| = 4$. The Wigner function for the cat state is given by Eq. (126) where we choose $\|\vec{\alpha}\| = 6$. Finally, for the GKP state, Eq. (130), we numerically integrated a wave-function expression of the state with $s = 2$ and $\delta = 0.3$ to obtain the Wigner function.

space $\mathbb{R}^{2m}$ and define multimode Fock states as

$$\begin{aligned}&\left|n_{e_1}\right\rangle \otimes \cdots \otimes \left|n_{e_m}\right\rangle \\ &:= \frac{1}{\sqrt{n_1! \ldots n_m!}} [\hat{a}^\dagger(\vec{e}_1)]^{n_1} \ldots [\hat{a}^\dagger(\vec{e}_m)]^{n_m} |0\rangle , \end{aligned} \quad (119)$$

where the kth mode in the basis contains n_k photons. Note that for $n_k = 0$ we have a vacuum mode. For the Wigner function, this implies

$$W_{n_{e_1},\ldots,n_{e_m}}(\vec{x}) = W_{n_{e_1}}(\vec{x}_{\mathbf{e}_1}) \ldots W_{n_{e_m}}(\vec{x}_{\mathbf{e_f}}). \quad (120)$$

Here we use that the phase-space point $\vec{x}$ can be expressed as $\vec{x} = \vec{x}_{\mathbf{e_1}} \oplus \cdots \oplus \vec{x}_{\mathbf{e_m}}$, where $\vec{x}_{\mathbf{e_k}}$ is the phase-space coordinate within the subspace spanned by $\vec{e}_k$ and $\Omega\vec{e}_k$. We can note $\vec{x}_{\mathbf{e_k}} = (x_k, p_k)^T$, such that we find the coordinate representation $\vec{x} = (x_1, p_1, \ldots, x_m, p_m)^T$ in the chosen basis of phase space.

Non-Gaussian states do not necessarily have to be pure, they can also come in the form of statistical mixtures. The most basic example of such as state is a non-Gaussian mixture of Gaussian states. As a simple example, let is consider a mixture of two coherent states

$$\hat{\rho} = \frac{1}{2} \left(|\vec{\alpha}\rangle \langle\vec{\alpha}| + |-\vec{\alpha}\rangle \langle -\vec{\alpha}| \right). \quad (121)$$

Even though this is a highly classical state, it is still non-Gaussian as clearly seen from its Wigner function

$$W(\vec{x}) = \frac{1}{4\pi} \left(e^{-\frac{1}{2}\|\vec{x}-\vec{\alpha}\|^2} + e^{-\frac{1}{2}\|\vec{x}+\vec{\alpha}\|^2} \right). \quad (122)$$

Another important example of a non-Gaussian mixed state is

$$\hat{\rho}_\lambda = \lambda |0\rangle \langle 0| + (1-\lambda) \left|1_f\right\rangle \left\langle 1_f\right|. \quad (123)$$

In the mode $\vec{f}$, the Wigner function of the state behaves as

$$W_\lambda(\vec{x}) = [(1-\lambda)\|\vec{x}_{\mathbf{f}}\|^2 + 2\lambda - 1] \frac{e^{-\frac{1}{2}\|\vec{x}\|^2}}{2\pi}. \quad (124)$$

This Wigner function reaches negative values as long as $\lambda < 1/2$ and subsequently becomes positive. Nevertheless, it will remain non-Gaussian until $\lambda = 1$. As we see in Sec. C, even when the Wigner function is positive, it is not always possible to describe this state as a mixture of Gaussian states.

Generally speaking, non-Gaussian states can come in a wide variety of shapes, which can be much more exotic than the examples discussed above. A popular class of Gaussian states is obtained by taking coherent superpositions of Gaussian states. As we see in Sec. VII, there has been a strong experimental focus on two specific types of such states: Schrödinger's cat states [105] and GKP states [31]. The former are obtained by coherently superposing two coherent states, and often are split in even $|\text{cat}_+\rangle$ and odd $|\text{cat}_-\rangle$ cat states:

$$|\text{cat}_\pm\rangle := \frac{1}{\mathcal{N}} \left(|\vec{\alpha}\rangle \pm |-\vec{\alpha}\rangle \right), \quad (125)$$

where $\mathcal{N} = \sqrt{2(1 \pm \exp[-\|\vec{\alpha}\|^2])}$ is the normalization coefficient, which depends on the displacement $\vec{\alpha}$. The latter is often referred to as "the size of the cat." The Wigner function of these states resembles that of Eq. (122) but has an additional interference term

$$W_{\text{cat}_\pm}(\vec{x}) = \frac{e^{-\frac{1}{2}\|\vec{x}-\vec{\alpha}\|^2} + e^{-\frac{1}{2}\|\vec{x}+\vec{\alpha}\|^2} \pm \cos(\sqrt{2}\,\vec{\alpha}^T\vec{x}) e^{-\frac{1}{2}\|\vec{x}\|^2}}{4\pi(1 \pm e^{-\|\vec{\alpha}\|^2})}. \quad (126)$$

The appearance of these interference terms creates several regions in phase space where the Wigner function attains negative values. The term "Schrödinger's cat state" has historically grown from the idea that coherent states describe classical electromagnetic fields and can thus be considered "macroscopic," in particular, for large values of $\|\vec{\alpha}\|$. However, one should honestly admit that they fail to capture an important point of Schrödinger's thought experiment [106]: the entanglement with a microscopic quantum system (i.e., the decay event that triggers the smashing of the vial of poison). Nevertheless, the term "cat state" has established itself firmly in the CV jargon, and now also lies at the basis of derived concepts such "cat codes" for error correction [107,108].

Finally, there are the GKP states. In their idealized form, they rely on eigenvectors of the quadrature operators (sometimes also known as infinitely squeezed states). To keep notation simple, we restrict to the single mode with quadrature operators $\hat{x}$ and $\hat{p}$. The eigenvectors of these operators are then formally written as

$$\hat{x}\,|x\rangle = x\,|x\rangle, \quad \text{and} \quad \hat{p}\,|p\rangle = p\,|p\rangle. \quad (127)$$

Furthermore, we have the relations $\langle x' \mid x\rangle = \delta(x'-x)$, $\langle p' \mid p\rangle = \delta(p'-p)$, and $\langle p \mid x\rangle = e^{-ipx}/\sqrt{2\pi}$. GKP states are constructed by considering a grid of such states to create a qubit, by identifying the two following GKP vectors:

$$\left|\text{GKP}_{\tilde{0}}\right\rangle := \sum_{k\in\mathbb{Z}} \left|x = 2k\sqrt{\pi}\right\rangle, \quad (128)$$

$$\left|\text{GKP}_{\tilde{1}}\right\rangle := \sum_{k\in\mathbb{Z}} \left|x = (2k+1)\sqrt{\pi}\right\rangle. \quad (129)$$

Clearly, these states are not normalizable and not physical as they would require infinite energy to be created. Thus, it is common to construct approximate GKP states,

by replacing the states $|x\rangle$ with displaced squeezed states, and by truncating the summation by adding a Gaussian envelope:

$$|\mathrm{GKP}_0\rangle := \mathcal{N}_0 \sum_{k\in\mathbb{Z}} e^{-2\pi[k\delta]^2} \hat{D}[2k\sqrt{\pi}]\,|s\rangle\,, \tag{130}$$

$$|\mathrm{GKP}_1\rangle := \mathcal{N}_1 \sum_{k\in\mathbb{Z}} e^{-2\pi[(k+1/2)\delta]^2} \hat{D}[(2k+1)\sqrt{\pi}]\,|s\rangle\,. \tag{131}$$

Here $\mathcal{N}_{0,1}$ are normalization constants and $|s\rangle$ is a single-mode squeezed vacuum, which implies that its Wigner function is given by Eq. (85) with $\vec{\xi} = (0,0)^T$ and $V = \mathrm{diag}[1/s, s]$ for $s > 1$. To get a good GKP state for quantum error correction, we generally need that $s \gg \sqrt{\pi}$. The Wigner function of an ideal GKP state is a grid of delta functions, which again highlights that it is a nonphysical state. The more realistic states $|\mathrm{GKP}_0\rangle$ and $|\mathrm{GKP}_1\rangle$ have well-defined Wigner functions, even though they are not very insightful to write down explicitly. In Fig. 2, we plot an example that was calculated numerically by taking into account only the first few terms around $k = 0$ in the sum.

GKP states may seem a little artificial at first glance, but they have been developed with a very clear purpose: to encode a qubit in a harmonic oscillator [31]. This encoding implies a notion of fault tolerance as these states are designed to be very efficient at correcting displacement errors. The more realistic incarnations of these states, Eq. (130), are therefore often proposed as candidates for encoding the information in CV quantum computation protocols [29]. Furthermore, it was shown that these states can also be used as the sole non-Gaussian resource to implement a CV quantum computer [32]

Once we progress into the realm of multimode states, the class of non-Gaussian states becomes even more vast. In Sec. 2, we present a dedicated introduction to multimode photon-subtracted states, which is a useful state to illustrate several of the concepts treated in this Tutorial. Furthermore, these states have a particular importance in CV quantum optics experiments. As a final example, we introduce another class of multimode non-Gaussian states, which have been highly relevant for quantum metrology: $N00N$ states [109–111]. Even though these states are very promising for quantum sensing with optical setups, the general idea that underlies these states was first introduced for fermions [112] in an attempt to mimic the advantage that is provided by squeezing in optics.

$N00N$ states are two-mode entangled states defined in a pair of orthogonal modes g_1 and g_2. The state contains exactly N photons, and is a superposition of a state with all photons being mode g_1 and a state with all photons in mode g_2:

$$|N00N\rangle := \frac{1}{\sqrt{2}}\left(\left|N_{g_1}\right\rangle + \left|N_{g_2}\right\rangle\right). \tag{132}$$

Here we recall that the state $\left|N_{g_1}\right\rangle$ can be trivially embedded to the full multimode space by adding vacuum in all other modes. We study these states in more detail for $N = 2$ in our discussion of the Hong-Ou-Mandel effect surrounding Eq. (165). However, here we highlight already that the Wigner function of $|N00N\rangle$ is not simply the sum of Wigner functions of the form Eq. (118). The entanglement will create additional interference terms, just like we saw in Eq. (126). In the present case, these interferences are genuinely multimode, and thus related to quantum correlations.

Experimentally, these states have been created and analyzed using a DV approach [113]. As we highlighted in Sec. C, the distinction between DV and CV is somewhat subtle and mainly depends on what is measured. Because $N00N$ states are built from Fock states and have a well-defined total photon number, they are most natural to analyze using photon-number-resolving detectors.

C. Quantum non-Gaussianity

Non-Gaussian states come in a wide variety, which means also that some of them are more exotic than others. Non-Gaussian states that are of limited interest, are those which are convex combinations of Gaussian states. Gaussian states do not form a convex set, after all, we can immediately see that, e.g., $[W_0(\vec{x} - \vec{\alpha}_1) + W_0(\vec{x} - \vec{\alpha}_2)]/2$ is not a Gaussian function even though it is a convex combination of Gaussian states.

The fact that the set of non-Gaussian states contains mixtures of Gaussian states may lead one to suspect that any mixed state with a positive Wigner function can be written as a well-chosen mixture of Gaussian states. After all, Gaussian states are the only pure states with positive Wigner functions. This intuition turns out to be false [89], which means that the set of states with a nonpositive Wigner function is not the same as the set of states that lie outside of the convex hull of Gaussian states $\mathcal{G}$. More formally, let us define

$$\mathcal{G} := \left\{\hat{\rho} \mid \hat{\rho} = \int \mathrm{d}\gamma\, p(\gamma)\hat{\rho}_G(\gamma)\right\}, \tag{133}$$

where γ is some arbitrary way of labelling Gaussian states $\hat{\rho}_G(\gamma)$ and $p(\gamma)$ is a probability distribution on these labels. Note that Eq. (103) tells us that we can generate all Gaussian states by taking convex combinations of displaced squeezed states, and thus we can limit ourselves to $\hat{\rho}_G(\gamma) = |\Psi_G(\gamma)\rangle\langle\Psi_G(\gamma)|$ in the definition of $\mathcal{G}$.

Any quantum state $\hat{\rho}$ that is not contained in the convex hull of Gaussian states, i.e., $\hat{\rho} \notin \mathcal{G}$, is referred to

as a "quantum non-Gaussian" state. The intuition behind this terminology is that Gaussian pure states are less quantum than non-Gaussian pure states that boast a non-positive Wigner function. A mixed state that is quantum non-Gaussian may have a positive Wigner function, but it cannot be created without adding states with nonpositive Wigner functions into the pure-state decomposition. Hence, these states are more quantum than the states that are in the convex hull of Gaussian states $\mathcal{G}$.

Next, one may wonder how to differentiate between states that are quantum non-Gaussian and states which are in the convex hull $\mathcal{G}$. Throughout the last decade, many methods have been developed to answer this question. We start by introducing the main idea of Ref. [114] because it is based on the Wigner function. The key idea is that Gaussian distributions have tails, which means that we can take an arbitrary pure Gaussian state $W_0[S^{-1}(\vec{x}-\vec{\alpha})]$, and evaluate the Wigner function at to origin of phase space:

$$W_0(S^{-1}\vec{\alpha}) = \frac{e^{-\frac{1}{2}\|S^{-1}\vec{\alpha}\|^2}}{(2\pi)^2}. \tag{134}$$

Clearly, when $\|S^{-1}\vec{\alpha}\|^2 \to \infty$, we do find that $W_0(S^1\vec{\alpha}) \to 0$. This limit essentially corresponds to a system with infinite energy. It is thus natural to try to bound the value of the Wigner function in the origin by a function that depends on the energy $\overline{N}$ of the state. For an arbitrary pure Gaussian state, we find that

$$\overline{N} = \sum_{j=1}^{m} \mathrm{tr}[\hat{\rho}\hat{a}^\dagger(\vec{e}_j)\hat{a}(\vec{e}_j)] = \frac{1}{4}\left(\mathrm{tr}[S^T S - \mathbb{1}] + \|\vec{\alpha}\|^2\right). \tag{135}$$

Using the properties of the operator norm, we write $\|S^{-1}\vec{\alpha}\|^2 = \|S^{-1}O\vec{\alpha}\|^2$, where O is a symplectic orthogonal transformation. Furthermore, we note that $\|\vec{\alpha}\|^2 = \|O\vec{\alpha}\|^2$. It is then useful to explicitly write the coordinate representation of the vector $O\vec{\alpha} = (\alpha_1^{(x)}, \alpha_1^{(p)}, \ldots, \alpha_m^{(x)}, \alpha_m^{(p)})^T$, such that

$$\overline{N} = \sum_{j=1}^{m} \frac{1}{4}\left(s_j + \frac{1}{s_j} + (\alpha_j^{(x)})^2 + (\alpha_j^{(p)})^2 - 2\right) \tag{136}$$

$$= \sum_{j=1}^{m} \overline{n}_j. \tag{137}$$

At the same time, we expand

$$\|S^{-1}\vec{\alpha}\|^2 = \sum_{j=1}^{m} s_j(\alpha_j^{(x)})^2 + \frac{(\alpha_j^{(p)})^2}{s_j}, \tag{138}$$

and with a little algebra we can show that

$$\frac{1}{2}\|S^{-1}\vec{\alpha}\|^2 \leqslant \sum_{j=1}^{m} 4\overline{n}_j(2\overline{n}_j + 1) \leqslant 4\overline{N}(2\overline{N}+1), \tag{139}$$

such that

$$W_0(S^{-1}\vec{\alpha}) \geqslant \frac{1}{(2\pi)^m} e^{-4\overline{N}(2\overline{N}+1)}. \tag{140}$$

This means that the value of the Wigner function of a pure Gaussian state in the origin of phase space is bounded from below by a function of the average number of particles $\overline{N}$. However, it is not obvious that we can extend this bound to arbitrary mixtures of Gaussian states. Let us assume that $\hat{\rho} \in \mathcal{G}$, then the Wigner function in the origin is given by

$$W(\vec{0}) = \int \mathrm{d}\gamma\, p(\gamma) W_0(S_\gamma^{-1}\vec{\alpha}_\gamma), \tag{141}$$

where we saw that $W_0(S_\gamma^{-1}\vec{\alpha}_\gamma)$ is the value of a pure Gaussian state's Wigner function in the origin and γ is some arbitrary label for the Gaussian states in the mixture. Therefore we can bound the states in the convex combination

$$W(\vec{0}) \geqslant \frac{1}{(2\pi)^m} \int_0^\infty \mathrm{d}\overline{N}_\gamma \tilde{p}(\overline{N}_\gamma) e^{-4\overline{N}_\gamma(2\overline{N}_\gamma+1)}, \tag{142}$$

where we introduce a probability distribution $\tilde{p}$ on the average particle numbers of the pure Gaussian states in the mixture. The overall average number of particles in the state $\hat{\rho}$ is then given by $\overline{N} = \int_0^\infty \mathrm{d}\overline{N}_\gamma\, \tilde{p}(\overline{N}_\gamma)\overline{N}_\gamma$. The final element that we require is the fact that $\exp[-4\overline{N}_\gamma(2\overline{N}_\gamma + 1)]$ is a convex function, such that we can apply Jensen's inequality to find that

$$\hat{\rho} \in \mathcal{G} \implies W(\vec{0}) \geqslant \frac{1}{(2\pi)^m} e^{-4\overline{N}(2\overline{N}+1)}. \tag{143}$$

This means that we can simply use the total energy of the state to construct a witness for quantum non-Gaussianity. This clearly shows that there are quantum non-Gaussian states with positive Wigner functions. An explicit example can be constructed by tuning the γ in the state $[(1-\gamma)|0\rangle\langle 0| + \gamma|1_{\vec{f}}\rangle\langle 1_{\vec{f}}|]$ to $1/2 > \gamma > 1/2 - e^{-4\gamma(2\gamma+1)}$ (where we use that γ is also the average particle number in this particular state).

Nevertheless, there are many quantum non-Gaussian states that do not violate inequality (143). After all, why would the origin of phase space be the most interesting point? A first solution is provided in Ref. [114], where it is

argued that

$$\hat{\rho} \in \mathcal{G} \implies W_\Gamma(\vec{0}) \geqslant \frac{1}{(2\pi)^m} e^{-4\overline{N}_\Gamma(2\overline{N}_\Gamma+1)}, \tag{144}$$

for all Gaussian channels Γ that act on the Wigner function of $\hat{\rho}$ as $W(\vec{x}) \overset{\Gamma}{\mapsto} W_\Gamma(\vec{x})$. Recall that the action of a Gaussian channel was defined in Eq. (102). The quantity $\overline{N}_\Gamma$ then denotes the average number of particles in the state $\Gamma(\hat{\rho})$. Further generalizations of this scheme have been worked out in Ref. [115]. Moreover, Ref. [116] has considered combinations of the value of the Wigner function in several points to reach better witnesses for quantum non-Gaussianity. Further progress has been made by identifying observables that are more easily measurable with typical CV techniques [117]. Others have considered other phase-space representations of the state to identify witnesses of quantum non-Gaussianity [118,119].

Quantum non-Gaussianity has been investigated with a wide range of tools. In this Tutorial we have limited ourselves to phase-space methods in the spirit of the CV approach. However, there is also a significant body of work on quantum non-Gaussianity using techniques that are more typical in DV quantum optics. The earliest works on the subject used photon statistics to distinguish quantum non-Gaussian states from convex mixtures of Gaussian states [89]. This research line has been continued in recent years to uncover new aspects of quantum non-Gaussian states, such as the "non-Gaussian depth" [120] and techniques to differentiate different types of multiphoton states [121]. Ultimately, these photon-counting techniques were extended to develop a whole hierarchy of quantum non-Gaussian states [37]. These ideas have been further formalized and generalized through the notion of the "stellar rank" of a quantum state.

D. Stellar rank

An interesting starting point to introduce the stellar representation is the method [103] to prove Hudson's theorem. We recall the definition

$$F^\star_\Psi(\vec{\alpha}) := \langle \vec{\alpha} \mid \Psi \rangle e^{\frac{1}{8}\|\vec{\alpha}\|^2},$$

of what we henceforth refer to as the stellar function. To avoid technical complications, let us now restrict ourselves to single-mode systems such that the optical phase space is $\mathbb{R}^2$. Note that, in a single-mode system, there is only one creation operator $\hat{a}^\dagger$ with associated annihilation operators $\hat{a}$. We can then follow Ref. [38] to introduce the stellar representation of single-mode quantum states. Note that some similar ideas are also present in other works [122].

First, we develop the stellar representation for pure states, which will then be used to generalize the framework to mixed states in Eq. (148). We use the definition of the displacement operator to show that

$$|\Psi\rangle = F^\star_\Psi(\hat{a}^\dagger)\,|0\rangle\,, \tag{145}$$

which immediately implies that the stellar representation is unique. In other words, if $F^\star_\Psi = F^\star_\Phi$ it follows that $|\Psi\rangle = |\Phi\rangle$ up to a phase. In our proof of Hudson's theorem, we have already highlighted that $F^\star_\Psi$ satisfies the property (116), which means that it is an entire function with growth order $r = 2$. Because in this single-mode setting $F^\star_\Psi$ can be interpreted as a function of a single complex variable, the Hadamard-Weierstrass theorem implies that $F^\star_\Psi$ can be fully represented by its zeros (one can consider this as a generalization of the fundamental theorem of calculus). This thus implies that a single-mode state is completely determined by the zeros of the $F^\star_\Psi$, and thus by the zeros of the Q function in Eq. (115).

It is thus natural to use these zeros in order to classify pure single-mode quantum states and thus the stellar rank is introduced. Ultimately, the stellar rank is simply given by the number of zeros of $F^\star_\Psi$ or alternatively the number of zeros of the Q function. Because in practice zeros may coincide, one should also consider the multiplicity of the zeros. We thus define the stellar rank $r^\star(\Psi)$ of $|\Psi\rangle$ as the number of zeros of $F^\star_\Psi : \mathbb{C} \mapsto \mathbb{C}$ counted with multiplicity. Alternatively one may count the zeros of the Q function with multiplicity and divide by two.

The fact that a state is fully characterized by its stellar representation $F^\star_\Psi$ can be made more explicit by considering the roots $\{\vec{\alpha}_1, \ldots, \vec{\alpha}_{r^\star(\Psi)}\}$ of the Q function, which represents $|\Psi\rangle$ [note that we use (44) to interchange between phase-space representation and complex Hilbert-space representation]. The single-mode state $|\Psi\rangle$ can then be used to express

$$|\Psi\rangle = \frac{1}{\mathcal{N}} \prod_{j=1}^{r^\star(\Psi)} \hat{D}^\dagger(\vec{\alpha}_j)\hat{a}^\dagger\hat{D}(\vec{\alpha}_j)\,|\Psi_G\rangle\,, \tag{146}$$

where $|\Psi_G\rangle$ is a pure Gaussian state and $\mathcal{N}$ a normalization constant. We can then use the stellar rank to induce some further structure in the set of states by defining

$$\mathcal{R}_N := \{|\Psi\rangle \mid r^\star(\Psi) = N\}. \tag{147}$$

that groups all states of stellar rank N. Note that the Hadamard-Weierstrass theorem also considers functions with an infinite amount of zeros and the case $N = \infty$ is thus mathematically well defined. It turns out that this case is not just a pathological limit. An evaluation of the Q function shows that Gottesman-Kitaev-Preskill states and Schrödinger cat states inhabit the set $\mathcal{R}_\infty$.

Clearly, all that was introduced so far only works for pure states. We can naturally extend this result via a convex

roof construction, by defining

$$r^\star(\hat{\rho}) := \inf_{\{p(\gamma),|\Psi(\gamma)\rangle\}} \sup_{\gamma}\{r^\star[\Psi(\gamma)]\}, \tag{148}$$

where the infimum is considered of all probability distributions on the set of pure states that lead to $\hat{\rho} = \int \mathrm{d}\gamma p(\gamma)\,|\Psi(\gamma)\rangle\langle\Psi(\gamma)|$. In words, there are many ways to decompose the state $\hat{\rho}$ in pure states and we consider all of them. For each decomposition, we define the stellar rank as the highest rank of the states in the decomposition. Then we minimize these values over all possible decompositions to arrive at the stellar rank of $\hat{\rho}$.

Convex roof constructions are commonly used to treat mixed states as they are easy and natural to formally define. However, they are often much harder to calculate in practice. This is where the stellar representation unveils its most remarkable property: stellar robustness. To formalize this idea, Ref. [38] introduced the robustness as the trace distance between the state and the nearest possible state of lower stellar rank.

$$R^\star(\Psi) := \inf_{r^\star(\hat{\rho})<r^\star(\Psi)} \frac{1}{2}\mathrm{tr}\sqrt{(|\Psi\rangle\langle\Psi| - \hat{\rho})^2}. \tag{149}$$

And it can be shown that

$$R^\star(\Psi) = \sqrt{1 - \sup_{r^\star(\hat{\rho})<r^\star(\Psi)} \langle\Psi|\,\hat{\rho}\,|\Psi\rangle}, \tag{150}$$

where $\langle\Psi|\,\hat{\rho}\,|\Psi\rangle$ is the fidelity of $\hat{\rho}$ with target state $|\Psi\rangle$. Remarkably, it can be shown that $R^\star(\Psi) > 0$ when $r^\star(\Psi) < \infty$ [38]. This means that any state that is sufficiently close to a pure state of $r^\star(\Psi)$ is also of rank $r^\star(\Psi)$ or higher.

The idea of stellar robustness can be generalized [123] by introducing k robustness. For any $k < r^\star(\Psi)$, we define

$$R^\star_k(\Psi) := \inf_{r^\star(\Phi)\leqslant k} \sqrt{1 - |\langle\Phi \mid \Psi\rangle|^2}, \tag{151}$$

and show subsequently that

$$R^\star_k(\Psi) = \sqrt{1 - \sup_{r^\star(\hat{\rho})\leqslant k} \langle\Psi|\,\hat{\rho}\,|\Psi\rangle}. \tag{152}$$

The k robustness can thus be interpreted as the nearest distance from a state $|\Psi\rangle$ at which we can find any state of stellar rank k, provided $k < r^\star(\Psi)$. Beyond showing that $R^\star_k(\Psi)$ is nonzero when $|\Psi\rangle$ is of finite stellar rank, Ref. [123] also provides an explicit method to calculate $R^\star_k(\Psi)$. Thus, for whichever state $\hat{\rho}$ is available in an experiment one can attempt to find a pure target state $|\Psi\rangle$ for which $\langle\Psi|\,\hat{\rho}\,|\Psi\rangle > 1 - R^\star_k(\Psi)^2$ to prove that $\hat{\rho}$ is at least of stellar rank k.

We note that for pure states $r^\star(\Psi) = 0$ implies that the state is Gaussian (this is essentially what is proven in Hudson's theorem). From the definition (148) we can then deduce that

$$r^\star(\hat{\rho}) = 0 \iff \hat{\rho} \in \mathcal{G}, \tag{153}$$

where $\mathcal{G}$ denotes, again, the convex hull of Gaussian states. On the other hand, states for which $r^\star(\hat{\rho}) > 0$ cannot be written as a mixture of Gaussian states and are thus quantum non-Gaussian.

This idea can be extended by using the stellar k robustness, Eq. (152), as a witness of quantum non-Gaussianity. When we want to check whether $\hat{\rho}$ is quantum non-Gaussian, it suffices to find a pure target state $|\Psi\rangle$ such that $\hat{\rho}$ is closer to $|\Psi\rangle$ than the 1 robustness $R^\star_k(\Psi)$. More formally written, whenever a pure state $|\Psi\rangle$ exists with the following property:

$$\langle\Psi|\,\hat{\rho}\,|\Psi\rangle > 1 - R^\star_1(\Psi)^2 \implies \hat{\rho} \notin \mathcal{G}, \tag{154}$$

and $\hat{\rho}$ is quantum non-Gaussian. This may seem like a complicated challenge, but for a single-photon state $|\Psi\rangle = |1\rangle$ we find that $1 - R^\star_1(\Psi)^2 \approx 0.478$ [123]. This means that any state that has a fidelity of more than 0.478 with respect to a Fock state is quantum non-Gaussian. This idea can be extended to higher stellar ranks: whenever we find a target state $|\Psi\rangle$ such that $\langle\Psi|\,\hat{\rho}\,|\Psi\rangle > 1 - R^\star_k(\Psi)^2$, the state $\hat{\rho}$ is at least of stellar rank k. Note that the fidelity of an experimentally generated state $\hat{\rho}$ with any pure target state $|\Psi\rangle$ can be calculated from double homodyne measurements on $\hat{\rho}$ [123]. There is no need to experimentally create the pure state $|\Psi\rangle$, the latter is just theoretical input needed to analyze the data.

Obviously, the stellar rank imposes a lot of additional structure on the state space. It rigorously orders all states that can be achieved by combining a finite number of creation operators and Gaussian transformations. The creation operator serves as a tool to increase the stellar rank by one and the stellar rank actually corresponds to the minimal number of times the creation operator must be applied to obtain the state, together with Gaussian operations. The stellar rank remains unchanged under Gaussian unitary transformations, which makes sense for a measure of the non-Gaussian character of the state, and thus it falls within the set of intrinsic properties of a state as discussed in Sec. C. Furthermore, the class of states with infinite stellar rank can be understood as the set that contains the most exotic states. However, it is lonely at the top as it can be shown that $R^\star_\infty(\Psi) = 0$ for states of infinite rank. This means that we can find states of finite stellar rank arbitrarily close to a state of infinite stellar rank. As stressed in Ref. [38], this implies that finite-rank states are dense in the full Fock space and any state of infinite rank can be arbitrarily well approximated by finite-rank states. Whereas a finite

stellar rank k of any experimental state can be certified by achieving a sufficiently high fidelity to a target state $|\Psi\rangle$ to fall within the range given by its k robustness $R_k^\star(\Psi)$, a similar procedure is impossible for infinite stellar ranks. This means, in practice, that genuinely infinite-rank states are impossible to certify in experiments because one never achieves perfect fidelity. Nevertheless, different states of infinite rank may differ significantly in the values $R_k^\star(\Psi)$ for $k < \infty$.

Many of the results on stellar rank rely on the Hadamard-Weierstrass theorem that allows one to uniquely factorize $F_\Psi^\star(\vec{\alpha})$ as a Gaussian and a polynomial, where the roots of the polynomial are the roots of $F_\Psi^\star$ and thus also the roots of the Q function. Sadly, this theorem cannot be straightforwardly generalized to a multimode setting, which is known in mathematics as Cousin's second problem [124]. Notable progress was made in Ref. [125] where one studies multimode stellar functions, which are polynomials and it was shown that there is no straightforward generalization of Eq. (146).

E. Wigner negativity

Hudson's theorem shows us that all pure non-Gaussian states have nonpositive Wigner functions, which sets them apart from normal probability distributions of phase space. For mixed states, this no longer holds and thus we spent the previous two sections developing methods to characterize the non-Gaussian features of these states. Whether it is through quantum non-Gaussianity or the more refined stellar rank, these methods focus on characterizing the non-Gaussian resources that are required to generate a certain state. In this subsection, we change the perspective and focus rather on negative values of the Wigner function ("Wigner negativity" in short) as a resource of interest.

Wigner negativity has the advantage of being a clear quantum feature, it reflects that different quadratures in the same mode cannot be jointly measured and thus goes hand in hand with the principle of complementarity. More formally, it has even been connected to the principle of quantum contextuality [64]. Indeed, in Sec. VI we elaborate on the fact that Wigner negativity is a necessary resource for reaching a quantum advantage, i.e., performing a task that cannot be efficiently simulated by a classical computer. However, the idea of using Wigner negativity as a signature of nonclassicality was already around before it was connected to a quantum computational advantage. An important step to formalize this idea was the introduction of a measure for Wigner negativity [90], which lies at the basis of recent resource theories of Wigner negativity [100,126].

A priori, there are several natural measures than can be used for Wigner negativity. It is therefore useful to consider some desirable properties that are required for a measure of Wigner negativity. First of all, we want the measure to be zero if and only if the Wigner function is positive. It seems natural to demand that, furthermore, Wigner negativity remains unchanged under Gaussian unitary transformations. It is then tempting to simply consider the absolute value of the lowest possible value of the Wigner function, but this would have some unnatural outcomes. It would mean that a single-photon state would have more Wigner negativity than a two-photon state. We thus need to look for a different measure.

The starting point of Ref. [90] is that the normalization of the Wigner function implies that

$$\int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\, |W(\vec{x})| \geqslant 1, \tag{155}$$

and that the inequality is strict whenever there is Wigner negativity. Furthermore, Liouville's theorem implies that integrals over phase space are unchanged by Gaussian transformations. A first possible way of measuring Wigner negativity is through the negativity volume

$$\mathcal{N}(\hat{\rho}) := \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\, |W(\vec{x})| - 1. \tag{156}$$

This measure has the major advantage of being convex due to the triangle inequality, which means that for $\hat{\rho} = \int \mathrm{d}\gamma p(\gamma)\hat{\rho}(\gamma)$ we find that

$$\mathcal{N}(\hat{\rho}) \leqslant \int \mathrm{d}\gamma p(\gamma)\mathcal{N}[\hat{\rho}(\gamma)]. \tag{157}$$

However, this measure is not additive, i.e., $\mathcal{N}(\hat{\rho}_1 \otimes \hat{\rho}_2) \neq \mathcal{N}(\hat{\rho}_1) + \mathcal{N}(\hat{\rho}_2)$. To circumvent this shortcoming, another measure for Wigner negativity has been introduced [100, 126,127]:

$$\mathfrak{N}(\hat{\rho}) := \log \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\, |W(\vec{x})|\,. \tag{158}$$

Clearly, $\mathfrak{N}(\hat{\rho}_1 \otimes \hat{\rho}_2) = \mathfrak{N}(\hat{\rho}_1) + \mathfrak{N}(\hat{\rho}_2)$ making this measure additive. However, the introduction of the logarithm destroys the convexity of the measure. Note that the two measures are closely related by $\mathfrak{N}(\hat{\rho}) = \log[\mathcal{N}(\hat{\rho}) + 1]$. Thus when $\mathcal{N}(\hat{\rho}_1) > \mathcal{N}(\hat{\rho}_2)$, we also find that $\mathfrak{N}(\hat{\rho}_1) > \mathfrak{N}(\hat{\rho}_2)$.

The single-mode examples that are considered in Ref. [90] lead to some interesting observations. First of all, they show that for Fock states Wigner negativity increases with the photon number. Furthermore, they show that for Schrödinger cat states the integral is bounded from above by a value smaller than the Wigner negativity of a two-photon state. Even though Fock states of increasing stellar rank have increasing Wigner negativity, there is no clear relation between stellar rank and Wigner negativity for more general classes of states. For example, Schrödinger

cat states are of infinite stellar rank, suggesting that they are in this regard the most exotic states, but they manifest only a limited amount of Wigner negativity.

As a case study, let us briefly concentrate on the Wigner negativity of Fock states, Eq. (117). One can now evaluate the Wigner negativity of such states to find that

$$\mathcal{N}(|1\rangle) \approx 0.426\,12 \quad \text{and} \quad \mathfrak{N}(|1\rangle) \approx 0.354\,959, \tag{159}$$

$$\mathcal{N}(|2\rangle) \approx 0.728\,99 \quad \text{and} \quad \mathfrak{N}(|2\rangle) \approx 0.547\,537, \tag{160}$$

$$\mathcal{N}(|3\rangle) \approx 0.976\,67 \quad \text{and} \quad \mathfrak{N}(|3\rangle) \approx 0.681\,415, \tag{161}$$

which shows that the negativity does not simply increase linearly with the number of photons even for the additive measure $\mathfrak{N}$. However, let us now look at a multimode n-photon state where each photon occupies a different mode, i.e., a Fock state generated by creation operators in $\vec{f}_1, \ldots, \vec{f}_n$ with $\text{span}\{\vec{f}_j, \Omega\vec{f}_j\} \neq \text{span}\{\vec{f}_k, \Omega\vec{f}_k\}$ for all $j \neq k$,

$$\hat{a}^\dagger(\vec{f}_1)\ldots\hat{a}^\dagger(\vec{f}_n)\,|0\rangle = \left|1_{f_1}\right\rangle \otimes \cdots \otimes \left|1_{f_n}\right\rangle. \tag{162}$$

The Wigner function for this state can be shown to be [showing this based on (38) is again a good exercise]

$$W_{1_{f_1},\ldots,1_{f_n}}(\vec{x}_{\mathbf{f_1}} \oplus \cdots \oplus \vec{x}_{\mathbf{f_n}}) = \prod_{k=1}^{n} W_{1_{f_k}}(\vec{x}_{\mathbf{f_k}}). \tag{163}$$

Either by explicitly using the expression of the Wigner function, or by using the additivity property, we find that

$$\mathfrak{N}\left(\left|1_{f_1}\right\rangle \otimes \cdots \otimes \left|1_{f_n}\right\rangle\right) = n\mathfrak{N}(|1\rangle). \tag{164}$$

Numerically, we can show that $n\mathfrak{N}(|1\rangle) > \mathfrak{N}(|n\rangle)$ and thus we can generally conclude that n photons in different modes hold more Wigner negativity than n photons in the same mode.

Let us now concentrate on the case where $n = 2$. We showed that two photons in different modes are more Wigner negative than two photons in the same mode, and now we combine this finding with the idea that Wigner negativity remains unchanged under Gaussian transformations. A particularly simple Gaussian transformation is a balanced beam splitter, which ultimately just implements a change in mode basis that we describe by an orthonormal transformation O_{BS}. When we mix two photons, prepared in orthogonal modes f_1 and f_2 by such a balanced beam splitter, we see the Hong-Ou-Mandel effect in action (more details can be found in Ref. [56] where a similar notation is used):

$$\left|1_{f_1}\right\rangle \otimes \left|1_{f_2}\right\rangle \overset{O_{\text{BS}}}{\mapsto} \frac{1}{\sqrt{2}}\left(\left|2_{g_1}\right\rangle - \left|2_{g_2}\right\rangle\right) := |\text{HOM}\rangle\,, \tag{165}$$

where f_1, f_2 and g_1, g_2 are the input and output modes of the beam splitter, respectively. The Hong-Ou-Mandel output state $|\text{HOM}\rangle$ is thus a superposition of two photons in mode g_1 and two photons in mode g_2. We can now use the simple fact that Wigner negativity is unchanged under Gaussian unitary transformations to show that

$$\mathfrak{N}(|\text{HOM}\rangle) = \mathfrak{N}\left(\left|1_{f_1}\right\rangle \otimes \left|1_{f_2}\right\rangle\right) = 2\mathfrak{N}(|1\rangle) > \mathfrak{N}(|2\rangle), \tag{166}$$

this then also implies that $\mathcal{N}(|\text{HOM}\rangle) > \mathcal{N}(|2\rangle)$. At first sight, this is somewhat of a peculiar finding: by taking a superposition of two states with the same Wigner negativity one finds a state with a higher Wigner negativity.

An explicit look at the Wigner function of the Hong-Ou-Mandel state $|\text{HOM}\rangle$ provides some insight. We find that this Wigner function can be written as (yet again a good exercise to show this explicitly)

$$W_{\text{HOM}}(\vec{x}_{\mathbf{g_1}} \oplus \vec{x}_{\mathbf{g_2}}) = \frac{1}{2}[W_{2_{g_1}}(\vec{x}_{\mathbf{g_1}}) + W_{2_{g_2}}(\vec{x}_{\mathbf{g_2}})] + W_{\text{int}}(\vec{x}_{\mathbf{g_1}} \oplus \vec{x}_{\mathbf{g_2}}), \tag{167}$$

where W_{int} is the contribution to the Wigner function that contains all the interference terms that are induced by the superposition. We can calculate that

$$\mathcal{N}\left(\frac{1}{2}[W_{2_{g_1}}(\vec{x}_{\mathbf{g_1}}) + W_{2_{g_2}}(\vec{x}_{\mathbf{g_2}})]\right) \leqslant \mathcal{N}(|2\rangle), \tag{168}$$

and thus, by additionally applying the triangle inequality, we can understand that the additional negativity in the Hong-Ou-Mandel state is due to the term W_{int}.

In the Hong-Ou-Mandel effect, it is common to talk about interference between particles, but in a more general CV language this interference will be equivalent to some form of entanglement, which is exactly described by the Wigner-function contribution W_{int}. In other words, the superposition between $\left|2_{g_1}\right\rangle$ and $\left|2_{g_2}\right\rangle$ has more Wigner negativity than each of its two constituents because it creates entanglement between the modes g_1 and g_2. This is a first indication that there is a connection between quantum correlations and non-Gaussian features of the Wigner function. We explore this connection in further detail in Sec. V.

Even though Wigner negativity is an important non-Gaussian feature, it is often hard to witness [128]. The most common experimental technique is homodyne tomography [129] to fully reconstruct the quantum state. These methods come with the inconvenience that it is hard to set good error bars. Techniques to circumvent the need for a full tomography have been developed based on homodyne [130] and double-homodyne (or heterodyne) measurements [123,128]. These methods come with the advantage of permitting to put a degree of confidence on the proclaimed Wigner negativity.

IV. CREATING NON-GAUSSIAN STATES

In Sec. III we have discussed the many ways of characterizing non-Gaussian quantum states and their properties. In this section, we explore the different theoretical frameworks for creating these states. An overview of some important experimental advances to put these theoretical techniques into practice is left for Sec. VII.

Gaussian quantum states can in some sense be understood as naturally occurring states. The foundational work of Planck that lies at the basis of all of quantum mechanics provides a first description of the thermal states of light that describe black-body radiation. In a more modern language, we refer to this as the thermal states of an ensemble of quantum harmonic oscillators or a free bosonic field. It has long been understood that these states are Gaussian [2–4]. Creating this kind of Gaussian states of light is thus literally as simple as switching on a light bulb.

When we turn towards more sophisticated light sources such as lasers, we can encounter coherent light that is described by coherent states [5,6]. Generating squeezed light becomes much harder and typically requires nonlinear optics [131]. Nevertheless, pumping a nonlinear crystal with a coherent pump generally suffices to deterministically create a squeezed state [132]. Recall from the end of Sec. D that from a theoretical point of view all these pure Gaussian states can be created by applying Gaussian unitary transformation to the vacuum state.

From an experimental point of view, the creation of non-Gaussian states is much harder than the creation of their Gaussian counterparts. Nevertheless, we start by introducing an ideal theoretical approach that is not too different from Gaussian states. In essence, it suffices to apply a non-Gaussian unitary operation to the state to create a non-Gaussian state. In Sec. A we dig deeper into the desired structure of such non-Gaussian unitary transformations that would in principle allow for the deterministic generation of non-Gaussian quantum states. In experiments (in particular, those in optics) such non-Gaussian unitary transformations are hard to come by, which is why one very often uses different preparation schemes. In Sec. B, we provide a general introduction into the conditional preparation of non-Gaussian quantum states, where one measures part of the system and conditions on a certain measurement outcome. This process projects the remainder of the system into a new non-Gaussian state.

A. Deterministic methods

To introduce some further structure in the sets of Gaussian and non-Gaussian unitary transformations, it is useful to take a quantum computation approach that is inspired by Ref. [26,133]. The central idea of this work is that Gaussian unitary transformations are always generated by "Hamiltonians" that are at most quadratic in the quadrature operators (or equivalently in the creation and annihilation operators). Let us denote that as

$$\hat{U}_G = \exp\{i\mathcal{P}_2(\hat{q})\} \tag{169}$$

where the polynomials $\mathcal{P}_2(\hat{q})$ are generated by combining terms of the types $\mathbb{1}$, $\hat{q}(\vec{f})$, and $\hat{q}(\vec{f}_1)\hat{q}(\vec{f}_2)$. A remarkable property of these three types of observables is that they are closed under the action of a commutator. Indeed, using the canonical commutation relation (29) and the general properties of commutators, we can show that

$$[\hat{q}(\vec{f}_1), \hat{q}(\vec{f}_2)] \sim \mathbb{1}, \tag{170}$$

$$[\hat{q}(\vec{f}_1), \hat{q}(\vec{f}_2)\hat{q}(\vec{f}_3)] \sim \hat{q}(\vec{f}\,'), \tag{171}$$

$$[\hat{q}(\vec{f}_1)\hat{q}(\vec{f}_2), \hat{q}(\vec{f}_3)\hat{q}(\vec{f}_4)] \sim \sum \hat{q}(\vec{f}_1{}')\hat{q}(\vec{f}_2{}'). \tag{172}$$

Thus, we can use the Baker-Campbell-Hausdorff formula to show that the combination of two Gaussian unitaries $\hat{U}_G\hat{U}'_G$ is again a Gaussian unitary.

This notion lies at the basis of universal gate sets in the CV approach. Using typical techniques from Lie groups, we can look for a minimal set of Gaussian unitaries than can be combined to generate all possible Gaussian unitary transformations. Generally, such a set is clearly not unique, but there are some natural choices. For example, we previously saw that a Gaussian unitary transformation is a combination of displacement operations and symplectic transformations. Furthermore, the Bloch-Messiah decomposition (95) shows us that any symplectic transformation can be decomposed into a combination of multimode interferometers and single-mode squeezing. In turn, interferometers can be decomposed as a combination of beam splitters and phase shifters [134]. Indeed, we can choose the set of Gaussian gates to be

$$\hat{U}_D(\vec{\lambda}) := \hat{D}(\vec{\lambda}), \tag{173}$$

$$\hat{U}_S(\vec{\lambda}) := \exp i[\hat{q}(\vec{\lambda})\hat{q}(\Omega\vec{\lambda}) + \hat{q}(\Omega\vec{\lambda})\hat{q}(\vec{\lambda})], \tag{174}$$

$$\hat{U}_P(\vec{\lambda}) := \exp i[\hat{q}(\vec{\lambda})^2 + \hat{q}(\Omega\vec{\lambda})^2], \tag{175}$$

$$\hat{U}_{\mathrm{BS}}(\vec{\lambda}_1, \vec{\lambda}_2) := \exp i[\hat{q}(\vec{\lambda}_1)\hat{q}(\vec{\lambda}_2) + \hat{q}(\Omega\vec{\lambda}_1)\hat{q}(\Omega\vec{\lambda}_2)], \tag{176}$$

where we note that $\vec{\lambda}, \vec{\lambda}_1, \vec{\lambda}_2 \in \mathbb{R}^{2m}$ are not normalized and $\vec{\lambda}_1 \perp \vec{\lambda}_2$. These unitary operators describe a displacement, a squeezer, a phase shifter, and a beam splitter, respectively. We note that all these operations act on a single mode, except for the beam splitter, which connects a pair of modes. These transformations are referred to as a Gaussian gate set; when we can implement all these gates in all the modes of some mode basis, we can generate any multimode Gaussian transformation, and thus any Gaussian state.

To generate non-Gaussian unitary transformations and thus non-Gaussian states, we need to add more unitary

gates to the Gaussian gate set. The relevant question is thus how many gates one should add and which gates are the best choices. The answer to the first question is surprising: one needs to add just one single gate [27]. The argument is simple, when we consider the operators $\hat{q}(\vec{\lambda})^3$, we find that

$$\begin{aligned}[\hat{q}(\vec{\lambda}_1)^3, \hat{q}(\vec{\lambda}_2)^n] =& \hat{q}(\vec{\lambda}_1)^2[\hat{q}(\vec{\lambda}_1), \hat{q}(\vec{\lambda}_2)^n] \\ &+ [\hat{q}(\vec{\lambda}_1), \hat{q}(\vec{\lambda}_2)^n]\hat{q}(\vec{\lambda}_1)^2 \\ &+ \hat{q}(\vec{\lambda}_1)[\hat{q}(\vec{\lambda}_1), \hat{q}(\vec{\lambda}_2)^n]\hat{q}(\vec{\lambda}_1). \quad (177)\end{aligned}$$

With the canonical commutation relations we can show that $[\hat{q}(\vec{\lambda}_1), \hat{q}(\vec{\lambda}_2)^n] \sim \hat{q}(\vec{\lambda}_2)^{n-1}$, which can be inserted into Eq. (177) to obtain

$$\begin{aligned}[\hat{q}(\vec{\lambda}_1)^3, \hat{q}(\vec{\lambda}_2)^n] \sim& \hat{q}(\vec{\lambda}_1)^2\hat{q}(\vec{\lambda}_2)^{n-1} + \hat{q}(\vec{\lambda}_2)^{n-1}\hat{q}(\vec{\lambda}_1)^2 \\ &+ \hat{q}(\vec{\lambda}_1)\hat{q}(\vec{\lambda}_2)^{n-1}\hat{q}(\vec{\lambda}_1). \quad (178)\end{aligned}$$

This calculation thus shows that commutation with the operators $\hat{q}(\vec{\lambda})^3$ increases the order of the quadrature operators. Thus, with the quadratic Hamiltonians to generate all operations that conserve the order of polynomials of quadrature operators and $\hat{q}(\vec{\lambda})^3$ to increase the order of the polynomial by one, we can ultimately generate the full algebra of observables. On the level of unitary gates, this implies that a full universal gate set is given by

$$\mathcal{U} = \{\hat{U}_D(\vec{\lambda}), \hat{U}_S(\vec{\lambda}), \hat{U}_P(\vec{\lambda}), \hat{U}_{\mathrm{BS}}(\vec{\lambda}_1, \vec{\lambda}_2), \hat{U}_C(\vec{\lambda})\}, \quad (179)$$

$$\text{with} \quad \hat{U}_C := \exp i\hat{q}(\vec{\lambda})^3. \quad (180)$$

In other words, combining sufficiently many of these gates allows us to built any arbitrary unitary transformation generated by a Hamiltonian, which is polynomial in the quadrature operators.

The non-Gaussian gate $\hat{U}_C$ is known as the cubic phase gate. The argument above shows that any experiment that can implement Gaussian transformations and a cubic phase gate can in principle generate any arbitrary non-Gaussian state. Even though many protocols have been proposed to experimentally realize a cubic phase gate [28,135–138], any convincing implementations have yet to be demonstrated. One of the key problems is that experimental imperfections and finite squeezing are detrimental for the most commonly proposed methods [30].

In principle, there is no particular reason to limit our attention to cubic phase gates. Already in the very first work on the subject it is argued that essentially any Hamiltonian of a higher than quadratic order can be used as a generator [26]. Thus, optical processes that perform photon triplet generation can also be used as a non-Gaussian gate, which can even be converted into the cubic phase gate [139]. This requires well-controlled high $\chi^{(3)}$ nonlinearities, which are generally only achieved by using exotic nonlinear crystals or well-controlled individual atoms. Handling such setups with a sufficient degree of control to actually implement a quantum gate is extremely challenging.

Other CV systems are more appropriate for the implementation of non-Gaussian unitary transformations. In particular, the systems used in circuit QED have such non-Gaussian contributions in their Hamiltonians [140,141], which suggests that they may be more capable of deterministically generating non-Gaussian states than their optical counterparts. Still, the characterization, detection, and control of such states is expected to be challenging. Recently, some important progress was made by demonstrating triplet generation in these systems [142].

B. Conditional methods

The experimental difficulties that are encountered when trying to implement non-Gaussian unitary transformations can be circumvented by abandoning the demand of unitarity. This implies that we no longer consider operations that can be implemented deterministically, but rather resort to what can be broadly referred to as conditional operations. This idea was formalized by Kraus when characterizing the most general ways of manipulating quantum states [143].

In the most general sense, we can implement a conditional operation by taking a set of linear operators on Fock space $\hat{X}_1, \hat{X}_2, \ldots$ and acting on the state in the following way:

$$\hat{\rho} \mapsto \frac{\sum_j \hat{X}_j \hat{\rho} \hat{X}_j^\dagger}{\mathrm{tr}[\hat{\rho} \sum_j \hat{X}_j^\dagger \hat{X}_j]}. \quad (181)$$

This formalism is typically implemented by performing some form of generalized measurement on the state $\hat{\rho}$ [144]. When $\hat{\rho}$ is a deterministically generated Gaussian state, the action of a well-chosen set of operators $\hat{X}_1, \hat{X}_2, \ldots$ can turn it into a non-Gaussian state. In optics, two of the most well-known examples of this technique are single-photon addition and subtraction. In both cases, there is only a single operator $\hat{X}_1$. For photon addition, we implement $\hat{X}_1 = \hat{a}^\dagger(\vec{f})$, whereas photon subtraction requires the realization of a case where $\hat{X}_1 = \hat{a}(\vec{f})$.

In many physical setups, and, in particular, in optics, the problem is that measurements are destructive and a measurement effectively removes the measured mode from the system. Therefore, it is common to prepare large multimode Gaussian states of which a subset of modes is measured in order to conditionally prepare a non-Gaussian state in the remaining modes. We now introduce a general framework to describe the non-Gaussian Wigner functions that are created accordingly [39].

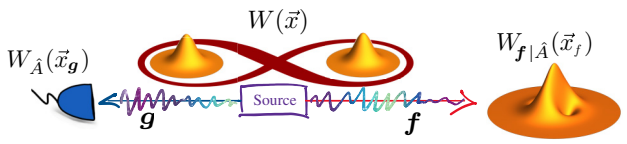

FIG. 3. Sketch representation of the conditional preparation scheme for creating the non-Gaussian states described by Eq. (190). Note that both **f** and **g** can be highly multimode. The Wigner function shown on the right is obtained by a conditional protocol shown in Ref. [39].

1. General framework

First of all, let us consider a general multimode phase space and separate it into two subsystems, i.e., $\mathbb{R}^{2m} = \mathbb{R}^{2l} \oplus \mathbb{R}^{2l'}$, where we perform some generalized measurement on the l' modes and leave the remaining l modes untouched. This introduces a general structure in the points of phase space $\vec{x} \in \mathbb{R}^{2m}$, which can now be written as $\vec{x} = \vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}$ with $\vec{x}_{\mathbf{f}} \in \mathbb{R}^{2l}$ and $\vec{x}_{\mathbf{g}} \in \mathbb{R}^{2l'}$. The general procedure is schematically outlined in Fig. 3 and we present the details step by step.

Any state $\hat{\rho}$ on this system then comes with a Wigner function $W(\vec{x}) = W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$ that is defined on the global phase space. This state can be reduced to one of the two subsystems by tracing out the other subsystem, which can be described on the level of the Wigner function by the following integrals:

$$W_{\mathbf{f}}(\vec{x}_{\mathbf{f}}) := \int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}} W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}), \tag{182}$$

$$W_{\mathbf{g}}(\vec{x}_{\mathbf{g}}) := \int_{\mathbb{R}^{2l}} \mathrm{d}\vec{x}_{\mathbf{f}} W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}). \tag{183}$$

When the state is Gaussian and the Wigner function is given by Eq. (85), the structure of the phase space is reflected in the mean field vector $\vec{\xi}$ and in the covariance matrix V:

$$\vec{\xi} = \vec{\xi}_{\mathbf{f}} \oplus \vec{\xi}_{\mathbf{g}}, \tag{184}$$

$$V = \begin{pmatrix} V_{\mathbf{f}} & V_{\mathbf{fg}} \\ V_{\mathbf{gf}} & V_{\mathbf{g}} \end{pmatrix}, \tag{185}$$

with $V_{\mathbf{fg}} = V_{\mathbf{gf}}^T$. The matrices $V_{\mathbf{f}}$ and $V_{\mathbf{g}}$ describe all the variances and correlations of the modes within $\mathbb{R}^{2l}$ and $\mathbb{R}^{2l'}$, respectively. In addition, the submatrix $V_{\mathbf{gf}}$ contains all the correlations between the modes in the different subspaces, which will be important for conditional state preparation. One can show that for such Gaussian states, the reduced states are also Gaussian, for the modes in $\mathbb{R}^{2l}$ given by

$$W_{\mathbf{f}}(\vec{x}_{\mathbf{f}}) = \frac{e^{-\frac{1}{2}(\vec{x}_{\mathbf{f}} - \vec{\xi}_{\mathbf{f}})^T V_{\mathbf{f}}^{-1} (\vec{x}_{\mathbf{f}} - \vec{\xi}_{\mathbf{f}})}}{(2\pi)^m \sqrt{\det V_{\mathbf{f}}}}, \tag{186}$$

and analogously for the modes in $\mathbb{R}^{2l'}$. As Gaussian states are the states that are least challenging to produce, they form the starting point of the conditional state preparation scheme.

As a next step, we must implement some form of operation on the modes that correspond to the phase space $\mathbb{R}^{2l'}$. To do so, we consider the action of a general positive operator-valued measure (POVM) element $\hat{A} \geqslant 0$ that corresponds to a specific measurement outcome. We can then obtain a conditional state via

$$\hat{\rho}_{\mathbf{f}|\mathbf{A}} := \frac{\mathrm{tr}_{\mathbf{g}}[\hat{A}\hat{\rho}]}{\mathrm{tr}[\hat{A}\hat{\rho}]}. \tag{187}$$

The partial trace $\mathrm{tr}_{\mathbf{g}}[\hat{A}\hat{\rho}]$ runs only over the modes in $\mathbb{R}^{2l'}$ because the other modes are left untouched. The denominator $\mathrm{tr}[\hat{A}\hat{\rho}]$ renormalizes the state and gives the probability of actually obtaining the measurement result that corresponds to $\hat{A}$. In an actual experiment, this operation is implemented by many repeated measurements of the modes in $\mathbb{R}^{2l'}$ and $\hat{A}$ corresponds to a specific detector output of these measurements. The nonmeasured part of the state is only used when the detector indicates this specific output, otherwise it is simply discarded. This conditional selection of the state significantly changes the properties of the state in a way that is strongly influenced by $\hat{A}$.

As we described in Eq. (42), the operator $\hat{A}$ comes with an associated phase-space representation $W_A(\vec{x}_{\mathbf{g}})$, which can be used to formally describe the phase-space representation of $\hat{\rho}_{\mathbf{f}|\mathbf{A}}$:

$$W_{\mathbf{f}|\mathbf{A}}(\vec{x}_{\mathbf{f}}) = \frac{\int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}} W_A(\vec{x}_{\mathbf{g}}) W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})}{\int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}} W_A(\vec{x}_{\mathbf{g}}) W_{\mathbf{g}}(\vec{x}_{\mathbf{g}})}. \tag{188}$$

There is a more practical way of expressing this Wigner function by exploiting the fact that the initial multimode Wigner function $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$ is positive and therefore describes a well-defined probability distribution on phase space. This implies that the conditional probability distribution

$$W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}}) := \frac{W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})}{W_{\mathbf{f}}(\vec{x}_{\mathbf{f}})}, \tag{189}$$

is also a well-defined probability distribution, which is obtained when we fix one point in phase space $\vec{x}_{\mathbf{f}} \in \mathbb{R}^{2l}$ and look at the probability distribution for the remaining modes in $\mathbb{R}^{2l'}$. We can then use this conditional probability distribution to write $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}) = W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}}) W_{\mathbf{f}}(\vec{x}_{\mathbf{f}})$, which can be inserted in Eq. (188) to find

$$W_{\mathbf{f}|\mathbf{A}}(\vec{x}_{\mathbf{f}}) = \frac{\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}}}{\langle \hat{A} \rangle} W_{\mathbf{f}}(\vec{x}_{\mathbf{f}}), \tag{190}$$

where we define

$$\langle \hat{A} \rangle := (4\pi)^{l'} \int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}} W_A(\vec{x}_{\mathbf{g}}) W_{\mathbf{g}}(\vec{x}_{\mathbf{g}}), \tag{191}$$

$$\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}} := (4\pi)^{l'} \int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}} W_A(\vec{x}_{\mathbf{g}}) W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}}). \tag{192}$$

The quantity $\langle \hat{A} \rangle$ is simply the expectation value of the observable $\hat{A}$ in the state $\hat{\rho}$. $\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$, on the other hand, is the expectation value of the function $W_A(\vec{x}_{\mathbf{g}})$ where $\vec{x}_{\mathbf{g}}$ is distributed according to the distribution $W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}})$, which makes $\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$ a function of the selected phase-space point $\vec{x}_{\mathbf{f}}$. However, even though $W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}})$ is a well-defined probability distribution on phase space, it does not necessarily correspond to a quantum state. Indeed, $W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}})$ may violate the Heisenberg inequality, which will be of vital importance in Sec. V as it is narrowly connected to quantum steering.

In the specific case where $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$ is Gaussian, we find that $W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}})$ is also a Gaussian probability distribution, given by

$$W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}}) = \frac{\exp\left[-\frac{1}{2}(\vec{x}_{\mathbf{g}} - \vec{\xi}_{\mathbf{g}|\vec{x}_{\mathbf{f}}})^T V_{\mathbf{g}|\vec{x}_{\mathbf{f}}}^{-1} (\vec{x}_{\mathbf{g}} - \vec{\xi}_{\mathbf{g}|\vec{x}_{\mathbf{f}}})\right]}{(2\pi)^{l'} \sqrt{\det V_{\mathbf{g}|\vec{x}_{\mathbf{f}}}}}. \tag{193}$$

Using the notation of Eq. (185), we express its covariance matrix

$$V_{\mathbf{g}|\vec{x}_{\mathbf{f}}} = V_{\mathbf{g}} - V_{\mathbf{gf}} V_{\mathbf{f}}^{-1} V_{\mathbf{gf}}^T, \tag{194}$$

and mean field vector

$$\vec{\xi}_{\mathbf{g}|\vec{x}_{\mathbf{f}}} = \vec{\xi}_{\mathbf{g}} + V_{\mathbf{gf}} V_{\mathbf{f}}^{-1} (\vec{x}_{\mathbf{f}} - \vec{\xi}_{\mathbf{f}}). \tag{195}$$

The covariance matrix $V_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$ is known in the mathematics literature [145] as the Schur complement of V. The Schur complement has interesting properties, for example, V is a positive matrix if and only if the same holds for the Schur complement $V_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$. This immediately implies that the Gaussian probability distribution in Eq. (193) is well defined. Furthermore, the Schur complement also plays an important role in the theory of Gaussian quantum correlations [146]. It should be noted that $V_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$ does not actually depend on the chosen value for $\vec{x}_{\mathbf{f}}$. Thus, the conditional expectation value $\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$ depends only on the phase-space point $\vec{x}_{\mathbf{f}}$ through the displacement $\vec{\xi}_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$. This is a particular feature of Gaussian states.

Finally, remark that the derivation of Eq. (190) holds true for all initial states with a positive Wigner function. Whenever the initial multimode Wigner function $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$ is positive, it follows that $W_{\mathbf{f}}(\vec{x}_{\mathbf{f}})$ is also positive. Furthermore, given that $\langle \hat{A} \rangle$ is the quantum expectation value of a positive semidefinite operator it clearly also is a positive quantity. Hence, Wigner negativity is entirely contained with $\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$. The fact that $\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$ can take negative values is exactly due to $W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}})$ not being the Wigner function of a quantum state. Furthermore, Eq. (192) teaches us that the conditionally generated Wigner function $W_{\mathbf{f}|\mathbf{A}}(\vec{x}_{\mathbf{f}})$ can only achieve negative value when $W_A(\vec{x}_{\mathbf{g}})$ is nonpositive.

Thus, in order to conditionally prepare a state with Wigner negativity, one faces strict requirements, on both the POVM element $\hat{A}$ that is conditioned upon and on the conditional probability distribution $W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}})$ that is obtained from the initial multimode state. We discuss this point in greater detail in Sec. B. For a more experimentally inclined perspective on the production of non-Gaussian states, we refer to Ref. [81].

Before we move on to consider photon subtraction as an example of conditional creation of non-Gaussian states, let us take for a moment the opposite process: Gaussification. The authors of Ref. [147] consider several copies of an initial non-Gaussian state, which are mixed through linear optics and subsequently some output modes are measured with on-off detectors. The conditioning is done of the events where no photons are detected, and such that we can interpret $\hat{A}$ as a projector in vacuum. By repeating several iterations of this scheme (assuming many successful conditioning events), the initial non-Gaussian state is converted into a Gaussian state. The Gaussification process thus relies on starting from a non-Gaussian state and conditioning by projecting on a Gaussian state: the vacuum. This point of view nicely complements our approach to create non-Gaussian states.

2. An example: photon subtraction

Single-photon subtracted states are theoretically obtained by acting with an annihilation operator on the state. Their density matrices are given by

$$\hat{\rho}^- = \frac{\hat{a}(\vec{b})\hat{\rho}\hat{a}^\dagger(\vec{b})}{\mathrm{tr}[\hat{a}^\dagger(\vec{b})\hat{a}(\vec{b})\hat{\rho}]}, \tag{196}$$

if the photon is subtracted in one specific mode **b**. In practice [129,148,149], we can implement this operation on the state $\hat{\rho}$ through a mode-selective beam splitter $\hat{U} = \exp\{\theta[\hat{a}^\dagger(\vec{g})\hat{a}(\vec{b}) - \hat{a}^\dagger(\vec{b})\hat{a}(\vec{g})]\}$, that couples the mode **b** to an auxiliary mode **g**, which is prepared in a vacuum state. We thus describe the action of the beam splitter on the system of interest and the auxiliary mode as $\hat{U}(\hat{\rho} \otimes |0\rangle\langle 0|)\hat{U}^\dagger$. As a next step, we mount a photon detector on one of the output modes of the beam splitter. This detector is crucial to make sure that no information is lost, without it we would effectively trace out the mode and the beam splitter would simply induce losses. In contrast, we condition on the specific events where the detector counts a single

photon, we generate the state

$$\hat{\rho}_\theta^- = \frac{\mathrm{tr}_{\mathbf{g}}[\hat{U}(\hat{\rho}\otimes|0\rangle\langle 0|)\hat{U}^\dagger(\mathbb{1}\otimes|1\rangle\langle 1|)]}{\mathrm{tr}[\hat{U}(\hat{\rho}\otimes|0\rangle\langle 0|)\hat{U}^\dagger(\mathbb{1}\otimes|1\rangle\langle 1|)]}. \tag{197}$$

The reader can now recognize Eq. (187). As a next step, we assume that the beam splitter is transmitting nearly all the incoming light, such that $\theta \to 0$. We can then approximate $\hat{U} \approx \mathbb{1} + \theta[\hat{a}^\dagger(\vec{g})\hat{a}(\vec{b}) - \hat{a}^\dagger(\vec{b})\hat{a}(\vec{g})]$. Then, when we insert this approximation in the expression for $\hat{\rho}_\theta^-$, we find that only the terms proportional to θ^2 survive such that

$$\hat{\rho}^- = \lim_{\theta\to 0}\hat{\rho}_\theta^- = \frac{\hat{a}(\vec{b})\hat{\rho}\hat{a}^\dagger(\vec{b})}{\mathrm{tr}[\hat{a}^\dagger(\vec{b})\hat{a}(\vec{b})\hat{\rho}]}. \tag{198}$$

A much more detailed analysis of multimode photon subtraction with imperfect mode selectivity can be found in Ref. [150]. We note that through this approach, photon subtraction can be understood as a weak measurement of the number of photons [151].

We can now derive the Wigner function of a single-photon-subtracted state through Eq. (193) by following the idea of Eq. (197). We initially start from a Gaussian state with covariance matrix $V_{\mathbf{f}}$ and one auxiliary mode that is prepared in the vacuum

$$V_{\mathrm{ini}} = \begin{pmatrix} V_{\mathbf{f}} & 0 \\ 0 & \mathbb{1} \end{pmatrix}. \tag{199}$$

We then implement a mode-selective beam splitter that mixes one specific mode **b** with the auxiliary vacuum mode, following the scheme outlined in Fig. 4. An effective way to describe such a transformation is by designing a new mode basis $\mathcal{B}$, which has **b** as one of the modes in the mode basis. We complete the basis with complementary modes $b_1^c, \ldots b_{m-1}^c$, such that the modes basis of phase space is given by $\mathcal{B} = \{\vec{b}_1^c, \Omega\vec{b}_1^c, \ldots, \vec{b}_{m-1}^c, \Omega\vec{b}_{m-1}^c, \vec{b}, \Omega\vec{b}\}$. Thus, we can perform such a basis change as

$$\begin{pmatrix} V_{\mathbf{f}} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \mapsto \begin{pmatrix} O_{\mathcal{B}}^T & 0 \\ 0 & \mathbb{1} \end{pmatrix}\begin{pmatrix} V_{\mathbf{f}} & 0 \\ 0 & \mathbb{1} \end{pmatrix}\begin{pmatrix} O_{\mathcal{B}} & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \tag{200}$$

where the matrix of basis change is given by

$$O_{\mathcal{B}} = \begin{pmatrix} | & | & & | & | & | & | \\ \vec{b}_1^c & \Omega\vec{b}_1^c & \ldots & \vec{b}_{m-1}^c & \Omega\vec{b}_{m-1}^c & \vec{b} & \Omega\vec{b} \\ | & | & & | & | & | & | \end{pmatrix}. \tag{201}$$

It is now instructive to explicitly write the rows and columns corresponding to mode b:

$$O_{\mathcal{B}}^T V_{\mathbf{f}} O_{\mathcal{B}} = \begin{pmatrix} V_{\mathbf{f}}^c & V_{\mathbf{f}}^{cb} \\ V_{\mathbf{f}}^{bc} & V_{\mathbf{f}}^b \end{pmatrix}. \tag{202}$$

Note that $V_{\mathbf{f}}^b$ is the 2×2 matrix that describes the initial state covariances of mode b, while $V_{\mathbf{f}}^c$ is the $(m-1)\times(m-1)$ that describes all the covariances in the complementary modes. The rectangular matrices $V_{\mathbf{f}}^{cb}$ and $V_{\mathbf{f}}^{bc}$ contain all the correlations between the mode b and the complementary modes in the basis. Now, we mix the mode b and the auxiliary vacuum mode on a beam splitter. This beam splitter is implemented by the transformation

$$V_{\mathrm{BS}}^{(\mathcal{B})} = O_{\mathrm{BS}}^{(\mathcal{B})}\begin{pmatrix} V_{\mathbf{f}}^c & V_{\mathbf{f}}^{cb} & 0 \\ V_{\mathbf{f}}^{bc} & V_{\mathbf{f}}^b & 0 \\ 0 & 0 & \mathbb{1} \end{pmatrix} O_{\mathrm{BS}}^{(\mathcal{B})\,T}, \tag{203}$$

where $O_{\mathrm{BS}}^{(\mathcal{B})}$ is given by

$$O_{\mathrm{BS}}^{(\mathcal{B})} = \begin{pmatrix} \mathbb{1} & 0 & 0 \\ 0 & \cos\theta\mathbb{1} & -\sin\theta\mathbb{1} \\ 0 & \sin\theta\mathbb{1} & \cos\theta\mathbb{1} \end{pmatrix}. \tag{204}$$

As a final step, we change the basis back to the original basis, such that the final state's covariance matrix becomes

$$V = \begin{pmatrix} O_{\mathcal{B}} & 0 \\ 0 & \mathbb{1} \end{pmatrix} V_{\mathrm{BS}}^{(\mathcal{B})} \begin{pmatrix} O_{\mathcal{B}}^T & 0 \\ 0 & \mathbb{1} \end{pmatrix}. \tag{205}$$

We can now rewrite this entire transformation such that the matrix V in Eq. (185) is given by

$$V = O_{\mathrm{BS}} V_{\mathrm{ini}} O_{\mathrm{BS}}^T, \tag{206}$$

with

$$\begin{aligned} O_{\mathrm{BS}} &= \begin{pmatrix} O_{\mathcal{B}} & 0 \\ 0 & \mathbb{1} \end{pmatrix} O_{\mathrm{BS}}^{(\mathcal{B})} \begin{pmatrix} O_{\mathcal{B}}^T & 0 \\ 0 & \mathbb{1} \end{pmatrix} \\ &= \begin{pmatrix} (\cos\theta - 1)\,BB^T + \mathbb{1} & \sin\theta\, B \\ -\sin\theta\, B^T & \cos\theta\mathbb{1} \end{pmatrix}. \end{aligned} \tag{207}$$

We introduce the $2m\times 2$ matrix B, which implements the mode selectivity of the beam splitter in mode b and is defined as

$$B = \begin{pmatrix} | & | \\ \vec{b} & \Omega\vec{b} \\ | & | \end{pmatrix}. \tag{208}$$

Hence, we can simply use O_{BS} as a mode-selective beam splitter that mixes one specific mode of a multimode state with the auxiliary mode. We should highlight that O_{BS} ultimately turns out to be independent of the complementary modes $b_1^c, \ldots b_{m-1}^c$. This means that the finer details of the interferometer $O_{\mathcal{B}}$ are not important for the final O_{BS}, the key point is that $O_{\mathcal{B}}$ changes towards a mode basis in which $\vec{b}$ and $\Omega\vec{b}$ are basis vectors of the phase space.

For the particular case of photon subtraction, we consider a very weak beam splitter, such that we consider the limit $\theta \to 0$. In this case, we can express the conditional

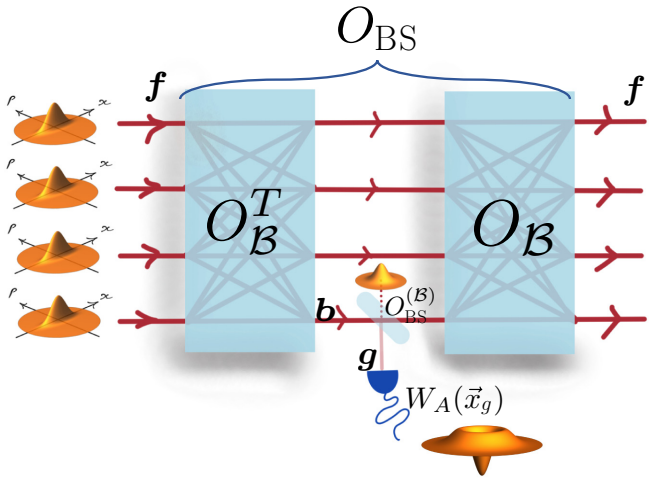

FIG. 4. Schematic representation of an implementation of mode-selective photon subtraction. See main text for details. For illustration, the initial state on the left is a product of single-mode squeezed vacuum states, but the protocol can in principle be applied to any Gaussian state.

mean field, Eq. (195), and covariance matrix, Eq. (194), of the auxiliary mode **g** by

$$\begin{aligned}\vec{\xi}_{\mathbf{g}|\vec{x}_\mathbf{f}} &\approx \theta(V_{\mathbf{bf}} - B^T)V_\mathbf{f}^{-1}(\vec{x}_\mathbf{f} - \vec{\xi}_\mathbf{f}) + \theta\vec{\xi}_\mathbf{b} \\ &= \theta B^T(\mathbb{1} - V_\mathbf{f}^{-1})(\vec{x}_\mathbf{f} - \vec{\xi}_\mathbf{f}) + \theta\vec{\xi}_\mathbf{b} \qquad (209)\\ V_{\mathbf{g}|\vec{x}_\mathbf{f}} &\approx \mathbb{1} + \theta^2\left[V_\mathbf{b} - \mathbb{1} - (V_{\mathbf{bf}} - B^T)V_\mathbf{f}^{-1}(V_{\mathbf{fb}} - B)\right], \\ &= \mathbb{1} + \theta^2\left(\mathbb{1} - B^T V_\mathbf{f}^{-1}B\right), \qquad (210)\end{aligned}$$

where we introduce the matrices $V_\mathbf{b} = B^T V_\mathbf{f} B$, $V_{\mathbf{b}\,\mathbf{f}} = B^T V_\mathbf{f}$, and $V_{\mathbf{fb}} = V_\mathbf{f} B$ as well as the vector $\vec{\xi}_\mathbf{b} = B^T\vec{\xi}_\mathbf{f}$. We can then use these quantities to evaluate that

$$\begin{aligned}W(\vec{x}_\mathbf{g} \mid \vec{x}_\mathbf{f}) \approx \frac{e^{-\frac{1}{2}\|\vec{x}_\mathbf{g}\|}}{(2\pi)^m}\Big(1 &+ \theta\,\vec{x}_\mathbf{g}^T\cdot B^T(\mathbb{1} - V_\mathbf{f}^{-1})(\vec{x}_\mathbf{f} - \vec{\xi}_\mathbf{f}) \\ &+ \theta\vec{x}_\mathbf{g}^T\cdot\vec{\xi}_\mathbf{b} + \frac{\theta^2}{2}\big[(\vec{x}_\mathbf{g}^T\cdot B^T(\mathbb{1} - V_\mathbf{f}^{-1})(\vec{x}_\mathbf{f} - \vec{\xi}_\mathbf{f}) \\ &+ \vec{x}_\mathbf{g}^T\cdot\vec{\xi}_\mathbf{b})^2 - \|B^T(\mathbb{1} - V_\mathbf{f}^{-1})(\vec{x}_\mathbf{f} - \vec{\xi}_\mathbf{f}) + \vec{\xi}_\mathbf{b}\|^2 \\ &+ \vec{x}_\mathbf{g}^T\left(\mathbb{1} - B^T V_\mathbf{f}^{-1}B\right)\vec{x}_\mathbf{g}\big] + \mathcal{O}(\theta^3)\Big). \qquad (211)\end{aligned}$$

As a next step, we must choose a POVM element $\hat{A}$ to measure. In the case of photon subtraction, we mount a photon counter on the auxiliary mode and for single-photon subtraction we condition on the event where this detector detects exactly one photon. Because we use a very weakly reflective beam splitter, the probability of obtaining such an event is small but when it occurs, we have created a photon subtracted state on the remaining modes.

On a theoretical level, mounting a photon counter and conditioning on a single photon is translated to choosing $\hat{A} = |1_g\rangle\langle 1_g|$. We already encountered the corresponding Wigner function in Eq. (118), and thus we can combine this with Eq. (211) to obtain

$$\begin{aligned}\langle\hat{A}\rangle_{\mathbf{g}|\vec{x}_\mathbf{f}} &= 4\pi\int_{\mathbb{R}^2} \mathrm{d}\vec{x}_\mathbf{g} W_{1_{\vec{g}}}(\vec{x}_\mathbf{g}) W(\vec{x}_\mathbf{g} \mid \vec{x}_\mathbf{f}) \\ &\approx \frac{\theta^2}{2}\Big(\|B^T(\mathbb{1} - V_\mathbf{f}^{-1})(\vec{x}_\mathbf{f} - \vec{\xi}_\mathbf{f}) + \vec{\xi}_\mathbf{b}\|^2 \\ &\quad + \mathrm{tr}\left[\mathbb{1} - B^T V_\mathbf{f}^{-1}B\right]\Big) + \mathcal{O}(\theta^3), \qquad (212)\end{aligned}$$

and in a similar fashion we find that

$$\begin{aligned}0\langle\hat{A}\rangle &= 4\pi\int_{\mathbb{R}^2} \mathrm{d}\vec{x}_\mathbf{g} W_{1_{\vec{g}}}(\vec{x}_\mathbf{g}) W_\mathbf{g}(\vec{x}_\mathbf{g}) \\ &\approx \frac{\theta^2}{2}\left(\mathrm{tr}\left[V_\mathbf{b} - \mathbb{1}\right] + \|\vec{\xi}_\mathbf{b}\|^2\right) + \mathcal{O}(\theta^3). \qquad (213)\end{aligned}$$

The actual evaluation of these integrals is not completely straightforward. As a key idea, we use that the integral takes the form of a polynomial multiplied by a Gaussian. We can thus evaluate the expectation value of the polynomial with respect to this Gaussian distribution. In essence, this boils down to calculating a set of moments of a Gaussian probability distribution. We see rather quickly that the lowest orders in θ vanish, such that the leading order is θ^2. Putting everything together, we find that

$$\begin{aligned}\lim_{\theta\to 0} W_{\mathbf{f}|\mathbf{A}}(\vec{x}_\mathbf{f}) &= \lim_{\theta\to 0}\frac{\langle\hat{A}\rangle_{\mathbf{g}|\vec{x}_\mathbf{f}}}{\langle\hat{A}\rangle} W_\mathbf{f}(\vec{x}_\mathbf{f}) \\ &= \frac{\|B^T(\mathbb{1} - V_\mathbf{f}^{-1})(\vec{x}_\mathbf{f} - \vec{\xi}_\mathbf{f}) + \vec{\xi}_\mathbf{b}\|^2 + \mathrm{tr}\left[\mathbb{1} - B^T V_\mathbf{f}^{-1}B\right]}{\mathrm{tr}\,(V_\mathbf{b} - \mathbb{1}) + \|\vec{\xi}_\mathbf{b}\|^2} \\ &\quad W_\mathbf{f}(\vec{x}_\mathbf{f}). \qquad (214)\end{aligned}$$

As such, we obtain the Wigner function for a multimode photon-subtracted state. This Wigner function can be obtained using several different methods, ranging from algebraic [152,153] to analytical [154]. The difference between those approaches and our method here is that we do not directly use the properties of the annihilation operator, but rather model the exact experimental setup, while relying entirely on phase-space representations.

The methods presented here for treating photon-subtracted states can straightforwardly be extended to the subtraction of multiple photons in different modes and we can easily replace photon-number-resolving detection with an on-off detector by setting $\hat{A} = \mathbb{1} - |0\rangle\langle 0|$. The techniques used in the calculations remain essentially the same and it yields the same result in the $\theta \to 0$ limit (doing this calculation may prove to be a good exercise for the motivated reader). However, any real implementation of a photon-subtraction experiment will use a beam splitter with finite reflectivity, such that there will be a

difference between on-off detectors and photon-number-resolving detectors due to the small contributions of higher order terms in θ. In practice, one chooses the reflectivity of the beam splitter with respect to the energy content of the initial state to effectively suppress all higher-order terms in Eq. (211). In the single-mode case, an early thorough analysis of the implementation of photon subtraction can be found in Ref. [155]. There are also proposals in the literature to use a photon-subtraction setup with larger values of θ to gain an additional advantage in quantum state preparation [156–158].

As a final note, we point out that a similar treatment can be used to describe photon-added states, which are also relevant in experiments [159,160]. It is perhaps surprising that such a state can be obtained by performing a measurement on a part of a Gaussian state, but it suffices to replace the beam splitter in Eq. (197) with a two-mode squeezer. In other words, we set $\hat{U} = \exp\{\theta[\hat{a}^\dagger(\vec{g})\hat{a}^\dagger(\vec{b}) - \hat{a}(\vec{b})\hat{a}(\vec{g})]\}$, and consider again the limit where the parameter θ is small, i.e., weak squeezing. Even though this is a simple step in theory, it is much harder in an actual experimental setting. Photon subtraction can be implemented with a passive linear optics element, while photon addition always requires squeezing and thus a nonlinear optics implementation.

V. NON-GAUSSIAN STATES AND QUANTUM CORRELATIONS

In this section, we explore the interplay between non-Gaussian effects and quantum correlations. First, in Sec. A, we provide a crash course to introduce the unfamiliar reader to the most important types of quantum correlations: entanglement, steering, and Bell nonlocality. In Sec. B we subsequently highlight how certain types of quantum correlations can be used to create certain types of non-Gaussian states via the methods of Sec. B. In Sec. C, we then explore how non-Gaussian operations can create or enhance quantum correlations by focusing on photon-subtracted states. Finally, we explore the role that is played by non-Gaussian states in Bell inequalities in Sec. D.

A. Quantum correlations: a crash course

We start by giving a quick introduction to the different kinds of common quantum correlations. Readers who want to get a more thorough overview on these subjects are referred to Refs. [161–163] as natural starting points.

In this Tutorial, we solely consider bipartite quantum correlations. This implies that we structure the system in a similar way as in Sec. B and divide the m-mode system in two parts, each with their own phase space, i.e., $\mathbb{R}^{2m} = \mathbb{R}^{2l} \oplus \mathbb{R}^{2l'}$. It is noteworthy that the corresponding Fock space takes the structure $\Gamma(\mathcal{H}_m) = \Gamma(\mathcal{H}_l) \otimes \Gamma(\mathcal{H}_{l'})$, where we again use the mapping, Eq. (44), between the phase space $\mathbb{R}^{2k}$ and the k-dimension Hilbert space $\mathcal{H}_k$. These structures are crucial to understand quantum correlations.

1. Correlations

To better understand quantum correlations, it is useful to start by generally defining what a correlations is. In a statistical sense, two stochastic variables X and Y are correlated when the expectation values have the following property:

$$\mathbb{E}(XY) \neq \mathbb{E}(X)\mathbb{E}(Y). \tag{215}$$

This can be translated to the level of probability distributions by stating that the joint probability distribution for outcomes $X = x$ and $Y = y$ is not the product of the marginals

$$P(x,y) \neq P(x)P(y), \tag{216}$$

where

$$P(x) = \int_{\mathcal{Y}} \mathrm{d}y\, P(x,y), \text{ and } P(y) = \int_{\mathcal{X}} \mathrm{d}x\, P(x,y). \tag{217}$$

Here, $\mathcal{X}$ and $\mathcal{Y}$ denote the possible outcomes of the stochastic variables X and Y, respectively [164].

When we talk about quantum systems, there are generally many observables that can be considered. When we consider a global multimode system with phase space $\mathbb{R}^{2m}$ and two subsystems with phase spaces $\mathbb{R}^{2l}$ and $\mathbb{R}^{2l'}$, there is a whole algebra of observables involved. The role of the stochastic observables X and Y will be taken up by local observables $\hat{X}$ and $\hat{Y}$ that are contained in the observable algebra generated by, respectively, $\hat{q}(\vec{f})$ and $\hat{q}(\vec{g})$, with $\vec{f} \in \mathbb{R}^{2l}$ and $\vec{g} \in \mathbb{R}^{2l'}$. These local observables are correlated when

$$\mathrm{tr}(\hat{X} \otimes \hat{Y}\hat{\rho}) \neq \mathrm{tr}(\hat{X}\hat{\rho}_{\mathbf{f}})\mathrm{tr}(\hat{Y}\hat{\rho}_{\mathbf{g}}), \tag{218}$$

where $\hat{\rho}_{\mathbf{f}}$ and $\hat{\rho}_{\mathbf{g}}$ are the marginals (or reduced states) of $\hat{\rho}$ for the subsystems $\mathbb{R}^{2l}$ and $\mathbb{R}^{2l'}$.

When we talk about correlated systems rather than correlated observables, we consider that there exists a pair of local observables such that Eq. (218) holds. Thus, if two systems are not correlated, it follows that for all possible observables $\mathrm{tr}(\hat{X} \otimes \hat{Y}\hat{\rho}) = \mathrm{tr}(\hat{X}\hat{\rho}_{\mathbf{f}})\mathrm{tr}(\hat{Y}\hat{\rho}_{\mathbf{g}})$. This lack of correlations can be expressed on the level of the quantum state by the identity $\hat{\rho} = \hat{\rho}_{\mathbf{f}} \otimes \hat{\rho}_{\mathbf{g}}$. On the level of Wigner functions, we can therefore say that a state contains correlations if the Wigner function satisfies

$$W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}) \neq W_{\mathbf{f}}(\vec{x}_{\mathbf{f}})W_{\mathbf{g}}(\vec{x}_{\mathbf{g}}), \tag{219}$$

where the marginal Wigner functions are defined as in Eqs. (182), (183).

It is clear that correlations between systems can occur, both, in the context of classical probability theory and in quantum theory. However, we already established that quantum physics imposes additional constraints on the statistics of observables, which ultimately make it impossible to describe CV quantum systems in terms of probability distributions on phase space. Similarly, quantum physics leads to new features for the correlations of subsystems. Thus, in our study of quantum correlations we explore correlated systems, in the sense of Eq. (219), and we seek to differentiate between correlations that are of classical origin and those that can be attributed to a quantum origin.

2. Quantum entanglement

Quantum entanglement is probably the most well-known type of quantum correlation. The notion of entanglement derives directly from the structure of the quantum state space and is related to the contrast between pure states in classical and quantum physics.

To understand this contrast, we loosely follow the idea of Ref. [165]. Let us be a bit more precise as to what is meant with pure states in classical physics in the context of CV systems. Classically, in a context of statistical mechanics, any CV system can be described by a probability distribution on phase space. From a mathematical point of view, this means that the space containing all the possible classical states is a convex set because any convex combination of two probability distributions is again a probability distribution. Pure states are formally defined as the extreme points of the convex set, i.e., the states that cannot be decomposed as being a convex combination of two other states. In a classical theory, where states can unambiguously be represented by probability distributions on phase space, the pure states are delta functions centered on the different points of phase space. From a physical point of view, this corresponds to the intuition that pure states are "the least noisy" states, which simply corresponds to a single point in phase space.

For our phase space $\mathbb{R}^{2m} = \mathbb{R}^{2l} \oplus \mathbb{R}^{2l'}$ these delta functions factorize with respect to the subsystems, i.e., $\delta(\vec{x} - \vec{x}') = \delta(\vec{x}_{\mathbf{f}} - \vec{x}'_{\mathbf{f}})\delta(\vec{x}_{\mathbf{g}} - \vec{x}'_{\mathbf{g}})$, with $\vec{x}, \vec{x}' \in \mathbb{R}^{2m}$, $\vec{x}_{\mathbf{f}}, \vec{x}'_{\mathbf{f}} \in \mathbb{R}^{2l}$, and $\vec{x}_{\mathbf{g}}, \vec{x}'_{\mathbf{g}} \in \mathbb{R}^{2l'}$. In the light of Eq. (216) we thus conclude that pure states of classical systems are always uncorrelated [166]. Any correlations that are present in classical states are thus obtained by taking a convex combination of uncorrelated pure states.

In quantum systems, pure states are represented by state vectors in a Hilbert space (in our case Fock space). They also can be seen as the extreme points of a convex set of states that contains all density matrices $\hat{\rho}$. As we saw in the example where we discussed the Hong-Ou-Mandel state $|\mathrm{HOM}\rangle$ in Eq. (165), pure quantum states can actually be correlated in the sense of Eq. (219). This crucial difference between classical and quantum pure states lies at the basis of quantum entanglement.

The notion of entanglement derives directly from the structure of the quantum state and is defined as the opposite of a separable state. For pure states, separable states $|\Psi\rangle \in \Gamma(\mathcal{H}_m)$ are the pure states that are uncorrelated and can thus be written as $|\Psi\rangle = |\Psi_l\rangle \otimes |\Psi_{l'}\rangle$ with $|\Psi_l\rangle \in \Gamma(\mathcal{H}_l)$ and $|\Psi_{l'}\rangle \in \Gamma(\mathcal{H}_{l'})$. All other pure states are said to be entangled. They possess correlations that are not due to some type of convex combination of uncorrelated states, something which is impossible for classical pure states.

The situation is more subtle when considering mixed states, i.e., convex combinations of pure states. Convex mixtures of classical pure states can also show correlations, and it is therefore crucial to make a distinction between this type of classical correlations and quantum correlations. Due to the structure of classical pure states, we find that any classical joint probability distribution on phase space can be written as a convex combination of local probability distributions

$$P(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}) = \int \mathrm{d}\gamma p(\gamma) P^{(\gamma)}(\vec{x}_{\mathbf{f}}) P^{(\gamma)}(\vec{x}_{\mathbf{g}}), \tag{220}$$

where γ is some arbitrary way of labeling states, distributed according to distribution $p(\gamma)$. This notion of classical correlations can directly be generalized to quantum states [167], and thus a mixed state is said to be separable when all of its correlations are classical, i.e., when it is a convex mixture of product states

$$\hat{\rho} = \int \mathrm{d}\gamma \; p(\gamma) \hat{\rho}_{\mathbf{f}}^{(\gamma)} \otimes \hat{\rho}_{\mathbf{g}}^{(\gamma)}. \tag{221}$$

In the language of Wigner functions, the separability condition translates to

$$W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}) = \int \mathrm{d}\gamma p(\gamma) W_{\mathbf{f}}^{(\gamma)}(\vec{x}_{\mathbf{f}}) W_{\mathbf{g}}^{(\gamma)}(\vec{x}_{\mathbf{g}}), \tag{222}$$

where we again use the definitions of Eqs. (182), (183). Quantum states that cannot be described by a Wigner function of the form Eq. (222) are not separable and are said to be entangled.

Hence, quantum entanglement describes the origin of the quantum correlations rather than their properties. Nevertheless, the set of separable states is a convex set and thus the Hahn-Banach separation theorem [168,169] teaches us that it is in principle possible to use observables to distinguish between separable and entangled states. In this sense the difference between entangled and separable states is measurable. For the sake of uniformity, we highlight that separable states lead to the following measurement

statistics of local observables $\hat{X}$ and $\hat{Y}$:

$$P(x,y) = \int \mathrm{d}\gamma p(\gamma) P^{(\gamma)}_{\hat{\rho}}(x) P^{(\gamma)}_{\hat{\rho}}(y). \tag{223}$$

It is crucial to emphasize that the distributions of measurement outcomes $P^{(\gamma)}_{\hat{\rho}}(x)$ and $P^{(\gamma)}_{\hat{\rho}}(y)$ are governed by the laws of quantum physics. Formally, we can use the spectral theorem to write

$$\hat{X} = \int_{\mathcal{X}} \mathrm{d}x\, x \hat{E}_x, \quad \text{and} \quad \hat{Y} = \int_{\mathcal{Y}} \mathrm{d}y\, y \hat{E}_y, \tag{224}$$

such that $\hat{E}_x$ and $\hat{E}_y$ are the POVM elements that correspond to the measurement outcomes x and y for the measurement of the (generalized) observables $\hat{X}$ and $\hat{Y}$, respectively. The probability distribution $P^{(\gamma)}_{\hat{\rho}}(x)$ is then given by

$$P^{(\gamma)}_{\hat{\rho}}(x) = \mathrm{tr}[\hat{E}_x \hat{\rho}^{(\gamma)}_{\mathbf{f}}] = (4\pi)^l \int_{\mathbb{R}^{2l}} \mathrm{d}\vec{x}_{\mathbf{f}}\, W_{E_x}(\vec{x}_{\mathbf{f}}) W_{\mathbf{f}}(\vec{x}_{\mathbf{f}}), \tag{225}$$

and analogously for $P^{(\gamma)}_{\hat{\rho}}(y)$.

For separable states, Eq. (223) with local probability distribution given by Eq. (225) holds for any arbitrary pair of local observables. The model that is described by these equations is known as a local hidden variable model for quantum entanglement, where γ is the hidden variable. We may not necessarily know the origins and behavior of γ, but the model generally captures two important elements. First, all correlations are induced by the common variable γ that governs the convex mixture. Second, the local probability distributions $P^{(\gamma)}_{\hat{\rho}}(x)$ and $P^{(\gamma)}_{\hat{\rho}}(y)$ have a quantum origin. For CV systems the latter point, for example, implies that these local probability distributions must satisfy the Heisenberg inequality. These quantum constraints on the local probability distributions $P^{(\gamma)}_{\hat{\rho}}(x)$ and $P^{(\gamma)}_{\hat{\rho}}(y)$ are typically useful for the falsification of the local hidden variable model, Eq. (223), and thus prove the presence of quantum entanglement [170,171].

3. Quantum steering

In a formal sense, quantum steering is a rather recent addition to the family of quantum correlations. Nevertheless, it is exactly this phenomenon that lies at the basis of the Einstein-Podolsky-Rosen (EPR) paradox [172]. Schrödinger's response to the EPR paper [173,174] lies at the basis of what we now call quantum steering, but the broader implications of these results were only sporadically discovered and formalized [175,176].

Just like for quantum entanglement, a system is said to be steerable if the measurement statistics cannot be explained in terms of a local hidden variable model. A peculiarity of quantum steering is that it involves a certain directionality, where one of the subsystems is said to "steer" the other subsystem. This asymmetry is represented in the local hidden variable model, which takes the following form:

$$P(x,y) = \int \mathrm{d}\gamma p(\gamma) P^{(\gamma)}(x) P^{(\gamma)}_{\hat{\rho}}(y), \tag{226}$$

where we emphasize the striking resemblance to Eq. (223). Note that, contrary to the case of quantum entanglement, we now allow the probability distribution $P^{(\gamma)}(x)$ for the first subsystem to be arbitrary and thus do not impose any constraints of quantum theory on it. If there exist observables $\hat{X}$ and $\hat{Y}$ for which the probability distribution is not consistent with the model, Eq. (226), the subsystems with phase space $\mathbb{R}^{2l}$ is able to steer the subsystem with phase space $\mathbb{R}^{2l'}$.

Quantum steering is perhaps most logically explained in terms of conditional states and probability distributions. For nonsteerable states, the local hidden variable model, Eq. (226), must hold for all observables, which in turn imposes conditions on the level of states. Here these conditions manifest on the level of conditional states of the type Eq. (187). To see this, we consider the conditional probability distribution associated with Eq. (226):

$$P(y \mid x) = \frac{\int \mathrm{d}\gamma p(\gamma) P^{(\gamma)}(x) P^{(\gamma)}_{\hat{\rho}}(y)}{P(x)}, \tag{227}$$

where the probability to obtain a certain outcome $\hat{X} = x$ is given by

$$P(x) = \int \mathrm{d}\gamma p(\gamma) P^{(\gamma)}(x). \tag{228}$$

Note that for any x the function

$$\tilde{P}(\gamma \mid x) := \frac{p(\gamma) P^{(\gamma)}(x)}{P(x)} \tag{229}$$

is a well-defined probability distribution. Furthermore, if we demand that Eq. (227) holds for all observables $\hat{Y}$, we find the following condition for the conditional state:

$$\hat{\rho}_{\mathbf{g}|\hat{X}=x} = \int \mathrm{d}\gamma\, \tilde{P}(\gamma \mid x) \hat{\rho}^{(\gamma)}_{\mathbf{g}}. \tag{230}$$

Because quantum steering is a property of the state, we again require Eq. (230) to hold for all observables $\hat{X}$ for a state to not be steerable.

The local hidden variable model, Eq. (226), and the consequence for the conditional state, Eq. (230), may seem stringent, but it is often intricate to formally prove that such a model cannot explain observed data. It turns out

that computational methods based on semidefinite programming [177] are well suited to prove that the set of all possible conditional states is inconsistent with Eq. (230). A more physical point of view is based on developing steering inequalities [178]. As a notable example, one can derive a type of conditional Heisenberg inequality for states of the form Eq. (230). The local hidden variable model, Eq. (226), assumes that the laws of quantum physics constrain the measurement statistics in the second subsystem. We can then define the conditional variance of an arbitrary observable $\hat{Y}$

$$\Delta^2(\hat{Y} \mid \hat{X} = x) := \mathrm{tr}[\hat{Y}^2 \hat{\rho}_{\mathbf{g}|\hat{X}=x}] - \mathrm{tr}[\hat{Y}\hat{\rho}_{\mathbf{g}|\hat{X}=x}]^2, \quad (231)$$

which leads to the "average inference variance"

$$\Delta^2_{\mathrm{inf}}(\hat{Y}) := \int_{\mathcal{X}} \mathrm{d}x P(x) \Delta^2(\hat{Y} \mid \hat{X} = x), \quad (232)$$

that characterizes the precision with which we can infer the measurement outcome of $\hat{Y}$, given a measurement outcome of $\hat{X}$. Under the assumption that Eq. (230) holds, we can than prove the inference Heisenberg inequality [178]

$$\Delta^2_{\mathrm{inf}}(\hat{Y}_1)\Delta^2_{\mathrm{inf}}(\hat{Y}_2) \geqslant \frac{1}{2}\int_{\mathcal{X}_3} \mathrm{d}x P(x) \left|\mathrm{tr}\left([\hat{Y}_1, \hat{Y}_2]\hat{\rho}_{\mathbf{g}|\hat{X}_3=x}\right)\right|^2, \quad (233)$$

where $\Delta^2_{\mathrm{inf}}(\hat{Y}_1)$ and $\Delta^2_{\mathrm{inf}}(\hat{Y}_2)$ can be conditioned on any observables $\hat{X}_1$ and $\hat{X}_2$, respectively.

Thus, whenever one performs a series of conditional measurements that violate the inference Heisenberg inequality (233), the assumption (230) cannot hold and thus the measurements in the subsystem with phase space $\mathbb{R}^{2l}$ have steered those in the subsystem with phase space $\mathbb{R}^{2l'}$. In more colloquial terms, the inequality (233) sets a limit on the precision with which classical correlations between observables can be used to infer measurement outcomes of one quantum subsystem, based on measurement outcome of the other subsystem (regardless of whether it is quantum or not). Quantum correlations allow us to outperform these bounds and provide better inference than classically possible, and this phenomenon is the essence of quantum steering.

Now let us now express Eq. (233) for quadrature operators:

$$\Delta^2_{\mathrm{inf}}[\hat{q}(\vec{g}_1)]\Delta^2_{\mathrm{inf}}[\hat{q}(\vec{g}_2)] \geqslant \left|\vec{g}_1^T \Omega \vec{g}_2\right|^2, \quad (234)$$

where $\vec{g}_1, \vec{g}_2 \in \mathbb{R}^{2l'}$. As a next step, we must understand the properties of the average inference variance $\Delta^2_{\mathrm{inf}}[\hat{q}(\vec{g}_1)]$, which we obtain by conditioning on a quadrature observable in the other subsystem's phase space $\mathbb{R}^{2l}$. More specifically let us assume that we condition on measurements of $\hat{q}(\vec{f}_1)$, such that we must evaluate the conditional variance $\Delta^2[\hat{q}(\vec{g}_1) \mid \hat{q}(\vec{f}_1) = x]$. The conditional variance $\Delta^2[\hat{q}(\vec{g}_1) \mid \hat{q}(\vec{f}_1) = x]$ is then given by the matrix element of the covariance matrix that describes $W(\vec{x}_{\mathbf{g}} \mid x\vec{f}_1)$ as defined in Eq. (193):

$$\Delta^2[\hat{q}(\vec{g}_1) \mid \hat{q}(\vec{f}_1) = x] = \vec{g}_1^T \left[V_{\mathbf{g}} - \frac{V_{\mathbf{gf}}\vec{f}_1\vec{f}_1^T V_{\mathbf{fg}}}{\vec{f}_1^T V_{\mathbf{f}}\vec{f}_1}\right]\vec{g}_1, \quad (235)$$

because the quantity does not depend on the actual outcome that is postselected upon, we find that

$$\Delta^2_{\mathrm{inf}}[\hat{q}(\vec{g}_1)] = \vec{g}_1^T \left[V_{\mathbf{g}} - \frac{V_{\mathbf{gf}}\vec{f}_1\vec{f}_1^T V_{\mathbf{fg}}}{\vec{f}_1^T V_{\mathbf{f}}\vec{f}_1}\right]\vec{g}_1, \quad (236)$$

$$\Delta^2_{\mathrm{inf}}[\hat{q}(\vec{g}_2)] = \vec{g}_2^T \left[V_{\mathbf{g}} - \frac{V_{\mathbf{gf}}\vec{f}_2\vec{f}_2^T V_{\mathbf{fg}}}{\vec{f}_2^T V_{\mathbf{f}}\vec{f}_2}\right]\vec{g}_2. \quad (237)$$

From Eqs. (236) and (237) we can deduce that $\Delta^2_{\mathrm{inf}}[\hat{q}(\vec{g})] \geqslant \vec{g}^T V_{\mathbf{g}|\vec{x}_{\mathbf{f}}}\vec{g}$ for all $\vec{g} \in \mathbb{R}^{2l'}$ regardless of the $\hat{q}(\vec{f})$ that is conditioned upon. Thus, if $V_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$ satisfies the Heisenberg inequality the inference Heisenberg inequality (234) is also satisfied.

The setting with homodyne measurements, or more general Gaussian measurements, is close to the system that is discussed in Ref. [172]. For this reason, we refer to quantum steering with Gaussian measurements as EPR steering in contrast to more general quantum steering. This type of steering has been studied extensively in the literature, e.g., [176,179–181] and will be a key element in Sec. 2.

Note that both subsystems clearly play a very different role in this setting. The first subsystem simply produces measurement results of different observables. The information of these measurements in the first subsystem is then used to infer measurement results in the second subsystem, which is assumed to be a quantum system. In a quantum communication context, this asymmetry corresponds to a level of trust: we position ourselves in the steered system and trust that our system is a well-behaved quantum system, but we do not trust the party that controls the other subsystem (up to a point where we do not even want to assume that the data that are communicated to us come from an actual quantum system). The violation of a steering inequality practically allows verification in such a setting that there is indeed a quantum correlation between the two subsystems [182].

The inference Heisenberg inequality (233) shows that quantum steering describes certain properties of the quantum correlations. States that can perform quantum steering thus possess correlations that can be used to infer measurement outcomes better than any classical correlations

could. These correlations cannot be described by a hidden variable model of the form Eq. (226), which is more general than the model, Eq. (223). Thus, all states that produce statistics consistent with Eq. (223) are also consistent with Eq. (226) such that states that can perform steering must be entangled. However, there are states that produce statistics that is consistent with Eq. (226), but inconsistent with Eq. (223). In other words, not all entangled states can be used to perform quantum steering. In this sense, quantum steering can be said to be "stronger" than quantum entanglement.

4. Bell nonlocality

To date, the seminal work of John S. Bell on the Einstein-Podolsky-Rosen paradox [183] is probably one of the most remarkable findings on the foundations of quantum physics. What most had long taken for granted, the existence of local hidden variables to explain the probabilistic nature of quantum physics, turned out to be inconsistent with the theoretical quantum formalism. It is here that we find the real historical origin of the concept of quantum correlations as something fundamentally different from classical ones.

As for quantum entanglement and steering, the story of Bell nonlocality starts from a local hidden variable model that bears a strong resemblance to Eqs. (223) and (226). In this case, the model attempts to describe the joint measurement statistics of $\hat{X}$ and $\hat{Y}$ as

$$P(x,y) = \int \mathrm{d}\gamma p(\gamma) P^{(\gamma)}(x) P^{(\gamma)}(y). \tag{238}$$

The key observation is that now all the quantum constraints on the probability distributions have been dropped and both the $P^{(\gamma)}(x)$ and $P^{(\gamma)}(y)$ can be any mathematically well-defined probability distributions. Even though the difference between Eqs. (223) and (226) on the one hand, and Eq. (238) on the other hand, may appear small, the impact of dropping the constraints on the local distributions is enormous. Think, for example, of the Hahn-Banach separation theorem that is invoked to define entanglement witnesses, this crucially relies on the Hilbert-space structure of the state space. Think for example of (233) which crucially depends on the fact that quantum probabilities are constrained by the Heisenberg inequality. Abandoning all connections that tie probabilities to operator algebras on Hilbert spaces deprives quantum mechanics of their toolbox. Nevertheless, it turns out that some quantum states induce statistics that is inconsistent with Eq. (238).

Again, we note that states that can be described by the models, Eqs. (223) or (226), can also be described by the model, Eq. (238). Bell's local hidden variable model, Eq. (238), is thus the most general one and the class of states that lead to measurement statistics that cannot be described by it is the smallest. Therefore, we say that the correlations that lead to a violation of the mode, Eq. (238), also known as Bell nonlocality, are the strongest types of quantum correlations.

The key insight of Bell's work [183,184] is that Eq. (238) puts constraints on the correlations of different combinations of observables in the subsystems. These constraints, cast in the form of Bell inequalities can be violated by certain quantum states. The inconsistency of quantum physics with the model, Eq. (238), can in itself be seen as a special case of contextuality [185]. Over the decades, many different kinds of Bell inequalities have been derived (see, for example, Ref. [186]). Here we restrict to presenting one of the most commonly used incarnations: the Clauser-Horne-Shimony-Holt (CHSH) inequality [187]. This inequality relies on the measurement of four observables: $\hat{X}$ and $\hat{X}'$ on the first subsystem and $\hat{Y}$ and $\hat{Y}'$ on the second subsystem. Furthermore, we consider that the observables can take two possible values: -1 or 1. Assuming the model in Eq. (238) it is then possible to derive

$$\left|\langle \hat{X}\hat{Y}\rangle - \langle \hat{X}\hat{Y}'\rangle + \langle \hat{X}'\hat{Y}\rangle + \langle \hat{X}'\hat{Y}'\rangle\right| \leqslant 2, \tag{239}$$

where $\langle . \rangle$ denotes the expectation value. In this Tutorial we skip the derivation of this result, but the interested reader is referred to Ref. [188] for a detailed discussion. Remarkably, certain highly entangled states can violate this inequality.

The experimental violation of Bell's inequalities formally shows that quantum correlations are profoundly different than classical correlations [189–192]. However, one needs clever combinations of several observables in both subsystems to actually observe the difference. With most experimental loopholes now closed [193–197], Bell inequalities can now in principle be used to impose a device-independent level of security on various quantum protocols [198].

As a concluding remark, it is interesting to highlight the existence of a semidevice-independent framework for testing quantum correlations [199,200]. The key idea is that nothing is assumed about the measurement devices nor about the states, much like in the scenario of Bell inequalities. Yet, in the framework of Refs. [199,200] one does add an additional level of trust in the sense that one assumes that the inputs of the measurement device can be controlled and trusted. In a way, this additional intermediate level of trust is somewhat reminiscent of quantum steering. This framework was very recently extended to the CV setting [201].

B. Non-Gaussianity through quantum correlations

In Sec. B, we explained how conditional operations can be used to create non-Gaussian quantum states. The presence of correlations plays an essential role in this

framework. Indeed, in the absence of correlations the combination of Eqs. (222) and (189) implies that $W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}}) = W_{\mathbf{g}}(\vec{x}_{\mathbf{g}})$. As a consequence, we see from Eq. (192) for the conditional expectation value $\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}} = \langle \hat{A} \rangle$, and thus from Eq. (190) that $W_{\mathbf{f}|\hat{A}}(\vec{x}_{\mathbf{f}}) = W_{\mathbf{f}}(\vec{x}_{\mathbf{f}})$. In other words, the conditional operation has no effect whatsoever and gives the same result as tracing out the modes in $\mathbb{R}^{2l'}$.

A closer look at the explicit expressions

$$W_{\mathbf{f}|\mathbf{A}}(\vec{x}_{\mathbf{f}}) = \frac{\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}}}{\langle \hat{A} \rangle} W_{\mathbf{f}}(\vec{x}_{\mathbf{f}}),$$

and

$$\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}} := (4\pi)^{l'} \int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}} W_A(\vec{x}_{\mathbf{g}}) W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}}),$$

shows that whenever there are correlations, and thus $\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}} \neq \langle \hat{A} \rangle$, the conditional Wigner function is *a priori* non-Gaussian. When we use explicitly that the initial state is Gaussian and thus that $W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}})$ is given by Eq. (193), this condition can be translated to the existence of nonzero components in $V_{\mathbf{gf}}$ in Eq. (185). The precise properties of the resulting non-Gaussian quantum state depend on the conditional expectation value $\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$.

In the literature, some attention has been devoted to proposing different types of measurements for such heralding procedures. One may think of using on-off detectors [202], photon-number-resolving detectors [156], parity detectors [203], and more exotic multimode setups [204,205]. However, these works usually assume that the initial quantum state is a pure Gaussian state obtained by an idealized source of multimode squeezed vacuum states. As we saw in Sec. 2, for pure-state correlations automatically imply entanglement, and it even turns out that all correlated pure states violate a Bell inequality [206]. In other words, for pure states all correlations are quantum correlations and all these quantum correlations are of the strongest type. When we no longer make such assumptions on the initial multimode Gaussian state, we see that $\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$ will not only depend on the chosen POVM $\hat{A}$, but also on the properties of $W(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}})$. In Secs. 1 and 2, we explain that certain types of non-Gaussian features can only be achieved through certain types of quantum correlations in the initial Gaussian state. An overview of the results of this section is provided in Fig. 5.

1. Quantum non-Gaussianity and entanglement

To understand the role of quantum entanglement in a conditional preparation scheme, we contrast it to a system with only classical correlations. In that regard, let us suppose that the initial quantum state is separable such that its Wigner function can be cast in the form Eq. (222). By inserting this form in Eq. (188), we find that

$$W_{\mathbf{f}|\mathbf{A}}(\vec{x}_{\mathbf{f}}) = \int \mathrm{d}\gamma p(\gamma) \frac{\int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}} W_A(\vec{x}_{\mathbf{g}}) W_{\mathbf{g}}^{(\gamma)}(\vec{x}_{\mathbf{g}})}{\int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}} W_A(\vec{x}_{\mathbf{g}}) W_{\mathbf{g}}(\vec{x}_{\mathbf{g}})} W_{\mathbf{f}}^{(\gamma)}(\vec{x}_{\mathbf{f}}). \tag{240}$$

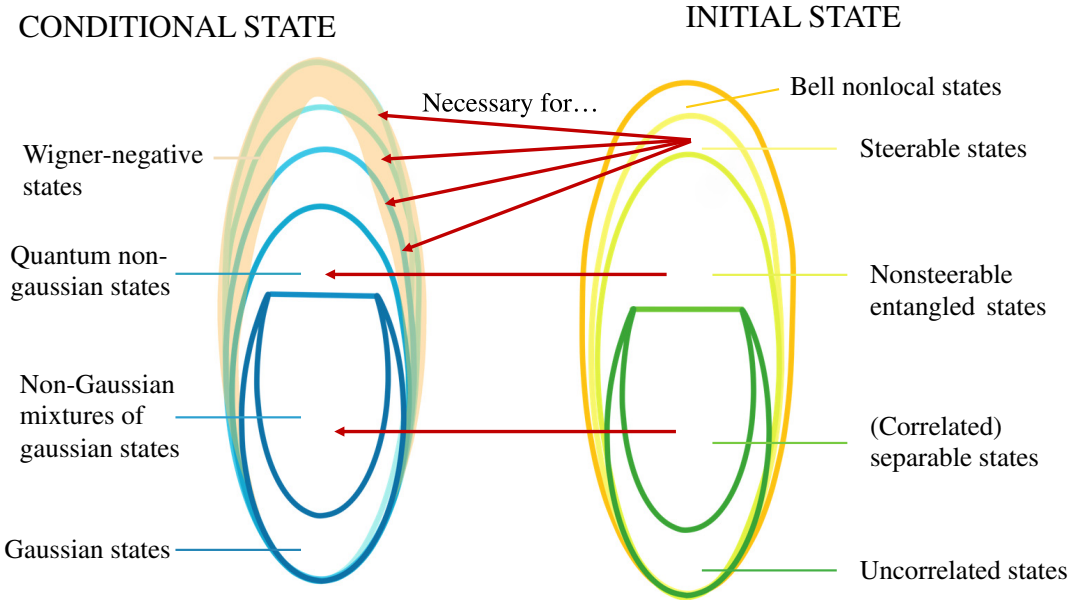

FIG. 5. Different types of quantum correlations are required to be present in the initial Gaussian state $W(\vec{x})$ to create conditional states $W_{\mathbf{f}|\mathbf{A}}(\vec{x}_{\mathbf{f}})$, as described in Eq. (190), that belong to a certain class. We thus show how the typical hierarchy of quantum correlations (right) can be connected to the structure of the CV state space that was introduced previously in Fig. 1. Throughout Sec. B, we prove that these different types of quantum correlations are necessary resources to achieve different types of states.

As a next step, we define

$$\tilde{p}_A(\gamma) := p(\gamma)\frac{\int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}} W_A(\vec{x}_{\mathbf{g}}) W_{\mathbf{g}}^{(\gamma)}(\vec{x}_{\mathbf{g}})}{\int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}} W_A(\vec{x}_{\mathbf{g}}) W_{\mathbf{g}}(\vec{x}_{\mathbf{g}})}, \tag{241}$$

and show that $\tilde{p}_A(\gamma)$ is a well-defined probability distribution. First, we use the definition of the reduced state

$$W_{\mathbf{g}}(\vec{x}_{\mathbf{g}}) = \int_{\mathbb{R}^{2l}} \mathrm{d}\vec{x}_{\mathbf{f}} W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}) \tag{242}$$

$$= \int \mathrm{d}\gamma p(\gamma) W_{\mathbf{g}}^{(\gamma)}(\vec{x}_{\mathbf{g}}), \tag{243}$$

and thus we immediately find that $\int \mathrm{d}\gamma \tilde{p}_A(\gamma) = 1$. Furthermore, we note that $W_{\mathbf{g}}^{(\gamma)}(\vec{x}_{\mathbf{g}})$ is the Wigner function of a well-defined quantum state $\hat{\rho}_{\mathbf{g}}^{(\gamma)}$ and thus

$$\int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}} W_A(\vec{x}_{\mathbf{g}}) W_{\mathbf{g}}^{(\gamma)}(\vec{x}_{\mathbf{g}}) = \mathrm{tr}[\hat{\rho}_{\mathbf{g}}^{(\gamma)} \hat{A}] \geqslant 0. \tag{244}$$

The final inequality follows from the fact that $\hat{A}$ is a positive semidefinite operator. As a consequence, we find that $\tilde{p}_A(\gamma) \geqslant 0$ for every possible γ. Thus, we find that for a separable initial state

$$W_{\mathbf{f}|\mathbf{A}}(\vec{x}_{\mathbf{f}}) = \int \mathrm{d}\gamma \tilde{p}_A(\gamma) W_{\mathbf{f}}^{(\gamma)}(\vec{x}_{\mathbf{f}}). \tag{245}$$

Up to this point, we assumed only that the initial state is separable. As we saw in Sec. C, a mixed quantum state with a positive Wigner function cannot necessarily be decomposed in states with positive Wigner functions. Therefore, we can generally not infer much about the properties of the Wigner function $W_{\mathbf{f}}^{(\gamma)}(\vec{x}_{\mathbf{f}})$ in Eq. (245).

As a next step, we use the fact that the initial state is also a Gaussian state. Recall from Eq. (103) that any mixed Gaussian state can be decomposed as a mixture of pure Gaussian states. *A priori*, however, it is not trivial that this decomposition is consistent with decomposition in separable states, Eq. (222). Thus, it remains to show that for Gaussian separable states the Wigner functions $W_{\mathbf{f}}^{(\gamma)}(\vec{x}_{\mathbf{f}})$ and $W_{\mathbf{g}}^{(\gamma)}(\vec{x}_{\mathbf{g}})$ in Eq. (222) are also Gaussian.

We start from a crucial observation on covariance matrices that was made in Ref. [207]: whenever an m-mode state with covariance matrix V is separable, there are covariance matrices $V'_{\mathbf{f}}$ and $V'_{\mathbf{g}}$ such that

$$V \geqslant \begin{pmatrix} V'_{\mathbf{f}} & 0 \\ 0 & V'_{\mathbf{g}} \end{pmatrix} = V'_{\mathbf{f}} \oplus V'_{\mathbf{g}}. \tag{246}$$

Note that $V'_{\mathbf{f}}$ and $V'_{\mathbf{g}}$ are covariance matrices on the phase spaces $\mathbb{R}^{2l}$ and $\mathbb{R}^{2l'}$, respectively. Nevertheless, $V'_{\mathbf{f}}$ and $V'_{\mathbf{g}}$ are generally not the same as the covariance matrices $V_{\mathbf{f}}$ and $V_{\mathbf{g}}$ of Eq. (185) that describe the marginal distributions. We should emphasize that the Williamson, Eq. (92), and Bloch-Messiah, Eq. (95), decompositions offer the necessary tools to explicitly construct $V'_{\mathbf{f}}$ and $V'_{\mathbf{g}}$ (we come back to this point in Sec. C). This allows us to use similar techniques as in Eq. (103). Let us first define

$$W'_{\mathbf{f}}(\vec{x}_{\mathbf{f}}) := \frac{e^{-\frac{1}{2}\vec{x}_{\mathbf{f}}^T V_{\mathbf{f}}'^{-1} \vec{x}_{\mathbf{f}}}}{(2\pi)^m \sqrt{\det V'_{\mathbf{f}}}}, \tag{247}$$

$$W'_{\mathbf{g}}(\vec{x}_{\mathbf{g}}) := \frac{e^{-\frac{1}{2}\vec{x}_{\mathbf{g}}^T V_{\mathbf{g}}'^{-1} \vec{x}_{\mathbf{g}}}}{(2\pi)^m \sqrt{\det V'_{\mathbf{g}}}}. \tag{248}$$

We can then use Eq. (246) to define a positive definite matrix $V_c := V - V'_{\mathbf{f}} \oplus V'_{\mathbf{g}}$, such that a decomposition of the type Eq. (103) gives us

$$W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}) = \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{y}\ W'_{\mathbf{f}}(\vec{x}_{\mathbf{f}} - \vec{y}_{\mathbf{f}}) W'_{\mathbf{g}}(\vec{x}_{\mathbf{g}} - \vec{y}_{\mathbf{g}}) \times \frac{e^{-\frac{1}{2}(\vec{y}-\vec{\xi})^T V_c^{-1}(\vec{y}-\vec{\xi})}}{(2\pi)^m \sqrt{\det V_c}}, \tag{249}$$

where we again impose the structure of the bipartition on $\vec{y} = \vec{y}_{\mathbf{f}} \oplus \vec{y}_{\mathbf{g}}$, with $\vec{y}_{\mathbf{f}} \in \mathbb{R}^{2l}$ and $\vec{y}_{\mathbf{g}} \in \mathbb{R}^{2l'}$. Furthermore, recall that $\vec{\xi}$ is the mean field of the initial Gaussian state $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$. The structure we obtain in Eq. (249) exactly corresponds to Eq. (222), where $\vec{y}$ now labels the states and thus plays the role of the abstract variable γ.

We can then use the structure Eq. (249) in the derivation Eq. (245) and then we find that

$$W_{\mathbf{f}|\mathbf{A}}(\vec{x}_{\mathbf{f}}) = \int_{\mathbb{R}^{2l}} \mathrm{d}\vec{y}_{\mathbf{f}} W'_{\mathbf{f}}(\vec{x}_{\mathbf{f}} - \vec{y}_{\mathbf{f}}) \tilde{p}_A(\vec{y}_{\mathbf{f}}). \tag{250}$$

In any concrete choice of $\hat{A}$, one can use Eq. (241) to derive an explicit expression for $\tilde{p}_A(\vec{y}_{\mathbf{f}})$, which will generally be a non-Gaussian probability distribution, such that $W_{\mathbf{f}|\mathbf{A}}(\vec{x}_{\mathbf{f}})$ describes a non-Gaussian state. However, the resulting conditional state, Eq. (250), is clearly a statistical mixture of Gaussian states and thus lies in the convex hull of Gaussian states. In the language of Sec. C this means that the conditional state is non-Gaussian but not quantum non-Gaussian and has a stellar rank 0.

In summary, we have assumed that our initial state with Wigner function $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$ is a separable Gaussian state. Without making any assumptions on the POVM element $\hat{A}$ of the conditional operation, we retrieve that the conditional state always is a convex combination of Gaussian states, given by Eq. (250). Thus, when the initial state is Gaussian, entanglement is a necessary resource to produce quantum non-Gaussian states via conditional operations.

2. Wigner negativity and Einstein-Podolsky-Rosen steering

In Sec. E, we explained that Wigner negativity is a "stronger" non-Gaussian feature than quantum non-Gaussianity. Here, we show that also stronger types of quantum correlations are required to conditionally create Wigner negativity. To understand how Wigner negativity can be achieved through a conditional preparation scheme, it suffices to understand when the conditional expectation value $\langle\hat{A}\rangle_{\mathbf{g}|\vec{x}_\mathbf{f}}$ in Eq. (192) reaches negative values.

Regardless of the chosen POVM, $W_A(\vec{x}_\mathbf{g})$ is the Wigner function of a positive semidefinite operator $\hat{A}$ as defined by Eq. (42). Thus, whenever there is a quantum state $\hat{\rho}'$ that has $W(\vec{x}_\mathbf{g} \mid \vec{x}_\mathbf{f})$ as associated Wigner function, Eq. (43) implies that $\langle\hat{A}\rangle_{\mathbf{g}|\vec{x}_\mathbf{f}} = \mathrm{tr}[\hat{\rho}'\hat{A}] \geqslant 0$. Hence, to conditionally create a nonpositive Wigner function (190) the conditional probability distribution $W(\vec{x}_\mathbf{g} \mid \vec{x}_\mathbf{f})$ cannot be a well-defined Wigner function. This observation holds whenever the initial state has a positive Wigner function.

When in addition we assume that the initial state is Gaussian, we find that $W(\vec{x}_\mathbf{g} \mid \vec{x}_\mathbf{f})$ is a Gaussian distribution (193). Whether the conditional probability distribution Eq. $W(\vec{x}_\mathbf{g} \mid \vec{x}_\mathbf{f})$ describes a Gaussian quantum state depends entirely in the properties of its covariance matrix, i.e., the Schur complement $V_{\mathbf{g}|\vec{x}_\mathbf{f}}$. Indeed, $W(\vec{x}_\mathbf{g} \mid \vec{x}_\mathbf{f})$ describes a quantum state if and only if $V_{\mathbf{g}|\vec{x}_\mathbf{f}}$ satisfies the Heisenberg inequality. Because $V_{\mathbf{g}|\vec{x}_\mathbf{f}}$ does not depend on the choice $\vec{x}_\mathbf{f} \in \mathbb{R}^{2l}$, it follows that $W(\vec{x}_\mathbf{g} \mid \vec{x}_\mathbf{f})$ corresponds to a quantum state either for all $\vec{x}_\mathbf{f} \in \mathbb{R}^{2l}$ [if the Schur complement Eq. (194) satisfies the Heisenberg inequality] or for none of the $\vec{x}_\mathbf{f} \in \mathbb{R}^{2l}$ [if the Schur complement Eq. (194) violates the Heisenberg inequality].

If $V_{\mathbf{g}|\vec{x}_\mathbf{f}}$ satisfies the Heisenberg inequality, the conditional state's Wigner function $W_{\mathbf{f}|\mathbf{A}}(\vec{x}_\mathbf{f})$ must thus be positive. To better understand the physical resources required to conditionally create Wigner negativity, one must comprehend what it means for $V_{\mathbf{g}|\vec{x}_\mathbf{f}}$ to violate Heisenberg's inequality in terms of quantum correlations. It turns out that this condition is closely related to the original argument of the EPR paper [172]. The violation of Heisenberg's inequality by the Schur complement $V_{\mathbf{g}|\vec{x}_\mathbf{f}}$ corresponds to Gaussian quantum steering in the state $W(\vec{x}_\mathbf{g} \oplus \vec{x}_\mathbf{f})$.

To understand the connection between the conditional covariance matrix $V_{\mathbf{g}|\vec{x}_\mathbf{f}}$ and quantum steering, we first express the Wigner function obtained by conditioning on a Gaussian measurement, such that the associated POVM element has a Wigner function $W_G(\vec{x}_\mathbf{f})$:

$$W_{\mathbf{g}|G}(\vec{x}_\mathbf{g}) = \frac{\int_{\mathbb{R}^{2l}} \mathrm{d}\vec{x}_\mathbf{f} W_G(\vec{x}_\mathbf{f}) W(\vec{x}_\mathbf{f} \mid \vec{x}_\mathbf{g})}{\int_{\mathbb{R}^{2l}} \mathrm{d}\vec{x}_\mathbf{f} W_G(\vec{x}_\mathbf{f}) W_\mathbf{f}(\vec{x}_\mathbf{f})} W_\mathbf{g}(\vec{x}_\mathbf{g}). \tag{251}$$

In a very similar way, we can also show that

$$W_{\mathbf{g}|G}(\vec{x}_\mathbf{g}) = \int_{\mathbb{R}^{2l}} \mathrm{d}\vec{x}_\mathbf{f} \frac{W_G(\vec{x}_\mathbf{f}) W_\mathbf{f}(\vec{x}_\mathbf{f})}{\int_{\mathbb{R}^{2l}} \mathrm{d}\vec{x}_\mathbf{f} W_G(\vec{x}_\mathbf{f}) W_\mathbf{f}(\vec{x}_\mathbf{f})} W(\vec{x}_\mathbf{g} \mid \vec{x}_\mathbf{f}). \tag{252}$$

Hence, when $W(\vec{x}_\mathbf{g} \mid \vec{x}_\mathbf{f})$ is a bona fide Wigner function for every $\vec{x}_\mathbf{f}$ this expression is an explicit manifestation of the local hidden variable model Eq. (230). In other words, whenever $W(\vec{x}_\mathbf{g} \mid \vec{x}_\mathbf{f})$ describes a quantum state, the modes in $\mathbf{g}$ cannot be steered by Gaussian measurements on the modes $\mathbf{f}$. Note that we can generalize Gaussian measurements to any measurement with a positive Wigner function.

The remarkable feature of EPR steering is that the inverse statement also holds: when $W(\vec{x}_\mathbf{g} \mid \vec{x}_\mathbf{f})$ is not a bona fide Wigner function Gaussian measurements can steer the state. Let us assume that Eq. (230) holds for Gaussian measurements. It then follows that a well-defined covariance matrix U exists such that the covariance matrix $V_{\mathbf{g}|G}$ of the conditional state $W_{\mathbf{g}|G}(\vec{x}_\mathbf{g})$ satisfies $V_{\mathbf{g}|G} \geqslant U$ for all Gaussian measurements. Furthermore, U is physical and satisfies the Heisenberg inequality. Reference [176] shows that the existence of such a U implies that the full covariance matrix of the system satisfies $V + 0_\mathbf{f} \oplus i\Omega_\mathbf{g} \geqslant 0$, which in turn implies that $V_{\mathbf{g}|\vec{x}_\mathbf{f}}$, the Schur complement of V, satisfies the Heisenberg inequality.

This shows that we can only generate Wigner negativity through Eq. (190) if the initial state can be steered by Gaussian measurements on the subsystem associated with phase space $\mathbb{R}^{2l}$. Note that the creation of Wigner negativity occurs in the opposite direction to the steering: We can produce Wigner negativity in the modes $\mathbf{f}$ by performing a measurement on the modes $\mathbf{g}$ if the modes $\mathbf{g}$ can be steered by performing Gaussian measurements on the modes $\mathbf{f}$. Somewhat counterintuitively, it turns out that the created Wigner negativity volume, Eq. (156), is not directly proportional to the strength of EPR steering [208].

As a final remark, we note that, in a multimode context, EPR steering is constrained by monogamy relations [209–211]. Notably, this implies that when a single mode g can be steered by a single other mode f, it is impossible for any other mode to also steer g. This naturally has profound consequences for the conditional generation of Wigner negativity that we discussed in this section. The monogamy relations for quantum steering can be used to derive similar monogamy relations [208] for the created Wigner negativity volume (156).

C. Quantum correlations through non-Gaussianity

In Sec. B, we extensively considered the use of quantum correlations as a resource to create non-Gaussian effects. In this subsection, we focus on the opposite idea where non-Gaussian operations increase or even create quantum

correlations. The subject of entanglement in non-Gaussian states is generally difficult to study, for some states it may be sufficient to evaluate lower-order moments [212] and when the density matrix in the Fock representation is available one can apply DV approaches to characterize entanglement [213]. However, these methods cannot always be applied and there are no universally applicable entanglement criteria that are practical to evaluate for arbitrary CV quantum states.

1. Entanglement measures on phase space

In Sec. 2, we argued that any pure state that manifests correlations between subsystems contains entanglement. Measuring entanglement in this case becomes equivalent to measuring the amount of correlation within the pure state. In particular, for pure states, one finds a wide range of entanglement measures in the literature [161]. In the case of CV systems, some measures are more appropriate than others, and here we focus on one particularly intuitive measure that is based on purity.

When we consider an arbitrary bipartite pure quantum state with Wigner function $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$ (with $\vec{x}_{\mathbf{f}} \in \mathbb{R}^{2l}$ and $\vec{x}_{\mathbf{g}} \in \mathbb{R}^{2l'}$), we find that its purity is $\mu = 1$ by definition. However, this is not necessarily true for the subsystems **f** and **g**. We can use Eq. (94) to evaluate the purity of any state based on its Wigner function, and we define

$$\mu_{\mathbf{f}} = \int_{\mathbb{R}^{2l}} \mathrm{d}\vec{x}_{\mathbf{f}} [W_{\mathbf{f}}(\vec{x}_{\mathbf{f}})]^2, \text{ and } \quad \mu_{\mathbf{g}} = \int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}} [W_{\mathbf{g}}(\vec{x}_{\mathbf{g}})]^2, \tag{253}$$

where we again use the definitions (182), (183). Because the global state with Wigner function $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$ is pure, we always find that $\mu_{\mathbf{f}} = \mu_{\mathbf{g}}$ (this is a general consequence of the existence of a Schmidt decomposition for pure states). Furthermore, if the pure state is separable, we find $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}) = W_{\mathbf{f}}(\vec{x}_{\mathbf{f}}) W_{\mathbf{g}}(\vec{x}_{\mathbf{g}})$ and as a consequence we obtain that $\mu_{\mathbf{f}} = \mu_{\mathbf{g}} = 1$. However, when $\mu_{\mathbf{f}} = \mu_{\mathbf{g}} < 1$ there must be correlations between the subsystems **f** and **g** and the smaller the purity of the subsystems, the stronger these correlations are. Without delving into the details, we stress that the opposite notion also holds: when there is a correlation between the subsystems, the purity of the subsystems is smaller than one.

To convert this quantity into an entanglement measure [214], it is useful to define the Rényi-2 entropy for subsystem **f**

$$S_R := -\log \mu_{\mathbf{f}}. \tag{254}$$

We then find that $S_R \geqslant 0$ and $S_R = 0$ if and only if the state is separable. Furthermore, it should be clear that S_R cannot be increased by local unitary operations on the subsystems **f** and **g**. We can thus define an entanglement measure for the pure state $|\Psi\rangle$ with Wigner function $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$ by setting

$$\mathcal{E}_R(|\Psi\rangle) := S_R. \tag{255}$$

This constitutes a well-defined entanglement measure for any chosen bipartition and any pure state on the phase space.

To extend this measure to mixed states, we follow Ref. [214] and construct a convex roof. Any mixed state $\hat{\rho}$ can be decomposed in pure states as $\hat{\rho} = \int \mathrm{d}\gamma\, p(\gamma) \left|\Psi^{(\gamma)}\right\rangle\!\left\langle\Psi^{(\gamma)}\right|$, we abbreviate this decomposition as the ensemble $\{p(\gamma), \left|\Psi^{(\gamma)}\right\rangle\}$. For each pure state in this ensemble, we can evaluate the entanglement $\mathcal{E}_R(\left|\Psi^{(\gamma)}\right\rangle)$ and subsequently average all of these values according to $p(\gamma)$. However, the decomposition of $\hat{\rho}$ in pure states is far from unique and different ensembles $\{p(\gamma), \left|\Psi^{(\gamma)}\right\rangle\}$ generally lead to a different value of entanglement even though they are all constrained to produce the same state $\hat{\rho}$. Therefore, it is common to define

$$\mathcal{E}_R(\hat{\rho}) := \inf_{\{p(\gamma), \left|\Psi^{(\gamma)}\right\rangle\}} \int \mathrm{d}\gamma\, p(\gamma) \mathcal{E}_R\left(\left|\Psi^{(\gamma)}\right\rangle\right) \tag{256}$$

as the general "Rényi-2 entanglement" of the state $\hat{\rho}$.

Formally, this is an elegant definition that can in principle be calculated directly from the Wigner function. However, in practice it is nearly impossible to actually identify all possible decompositions $\{p(\gamma), \left|\Psi^{(\gamma)}\right\rangle\}$, which makes this measure notoriously hard to evaluate for mixed states. This has sparked some alternative definitions of entanglement measures for Gaussian states, where any Gaussian state can be decomposed in an ensemble of Gaussian states, Eq. (103). Thus, one can define "Gaussian Rényi-2 entanglement" by restricting Eq. (256) to only Gaussian decompositions [215]. In this sense, Gaussian Rényi-2 entanglement is by construction an upper bound to the general Rényi-2 entanglement.

As an alternative to entanglement measures, it is common to use entanglement witnesses. These have been particularly successful for Gaussian states [146,216–220], where one commonly applies methods based on the covariance matrix of the state. Due to the extremality of Gaussian states [12] these results also provide witnesses for entanglement if the state is non-Gaussian. However, there are several examples of non-Gaussian entangled states for which no entanglement can be detected from the covariance matrix. Notable progress was made by developing entanglement witnesses for non-Gaussian states with specific structure in their Wigner function [221].

It is noteworthy to emphasize that the positive-partial transpose (PPT) criterion of Ref. [216] can in principle be implemented on the level of Wigner functions. To make this apparent, let us first define the transposition operator T that implements $\hat{\rho} \mapsto \hat{\rho}^T$. When $W(\vec{x})$ with $\vec{x} \in \mathbb{R}^{2m}$

denotes the Wigner function of the state $\hat{\rho}$, we can write the Wigner function of $\hat{\rho}^T$ as $W(T\vec{x})$. The matrix T can be written as

$$T = \bigoplus^{m} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{257}$$

which can be derived from the definition of the Wigner function [222]. The concept of partial transposition in entanglement theory relies on the simple idea that one can apply a transpose only on one of the two subsystems in the bipartition. In our context, this means that the Wigner function changes as $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}) \mapsto W(\vec{x}_{\mathbf{f}} \oplus T\vec{x}_{\mathbf{g}})$ (where T is now taken only on the l' modes of subsystem **g**). The PPT criterion is based on the idea that, in absence of entanglement, the function $W(\vec{x}_{\mathbf{f}} \oplus T\vec{x}_{\mathbf{g}})$ still gives a well-defined Wigner function of a quantum state. However, there are entangled states for which this is no longer true and $W(\vec{x}_{\mathbf{f}} \oplus T\vec{x}_{\mathbf{g}})$ becomes unphysical. This lack of physicality is expressed by the fact there exist positive semidefinite operators $\hat{A}$ for which

$$(4\pi)^m \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\, W_A(\vec{x}) W(\vec{x}_{\mathbf{f}} \oplus T\vec{x}_{\mathbf{g}}) < 0. \tag{258}$$

Finding such observables $\hat{A} \geqslant 0$ for a non-Gaussian state $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$ is generally a very hard task. For Gaussian states, on the other hand, the physicality of $W(\vec{x}_{\mathbf{f}} \oplus T\vec{x}_{\mathbf{g}})$ is simply checked through Heisenberg's inequality. For more general non-Gaussian states, this is insufficient and one should check a full hierarchy of inequalities instead [212]. Nevertheless, one may yet uncover more direct methods to check the properties of $W(\vec{x}_{\mathbf{f}} \oplus T\vec{x}_{\mathbf{g}})$.

2. Entanglement increase

One of the most well-known protocols for increasing entanglement is entanglement distillation. In this protocol, one acts with local operations on a large number of mixed entangled states that are shared by two parties and concentrates the entanglement in a smaller number of maximally entangled pairs [223]. When the initial states are pure and the local operation serves only to increase the entanglement and not the purity, we speak of entanglement concentration [224]. Conditional operations play an important role in these protocols, and we can alternatively think of entanglement distillation as the idea that a conditional operation can increase the entanglement of a state. For Gaussian quantum states, there is a notorious no-go theorem that states that Gaussian measurements (or Gaussian operations in general) cannot increase bipartite entanglement [225–227]. It was quickly realized that these no-go results can be circumvented by even the most basics non-Gaussian states: those created through a non-Gaussian noise process [228,229]. On the other hand, if one wants to distill entanglement in a CV system starting from initial Gaussian states one really requires non-Gaussian operations. One such example is given in Refs. [230,231], where the authors propose to use a Kerr nonlinearity to distill entanglement for mixed Gaussian states. In contrast, conditional schemes have also been proposed [147,232,233], avoiding the need for optical nonlinearities. In those protocols, one first uses conditional operations to create non-Gaussian states and subsequently uses Gaussification to obtain states with higher entanglement. A narrowly related protocol [234] relies on the implementation of noiseless linear amplification [235], where the non-Gaussian element is injected in the form of auxiliary Fock states.

The realization that photon subtraction and addition can be used to increase the entanglement of a Gaussian input state was developed reasonably early [236–238] and was further formalized in works such as Refs. [239–241]. Remarkably, all of these works explicitly assume that the initial state under consideration is a two-mode squeezed state and the approach strongly relies on the structure of this type of state in the Fock basis. Beyond the two mode setting, the class of CV graph states has also been studied in the context of entanglement increase [242,243]. Here we provide an alternative approach, based on phase-space representations to understand entanglement increase due to the subtraction of a single photon.

Our approach relies on the fact that we can easily apply the entanglement measure (256) when the global state is pure. This means that we are focusing on a context of entanglement concentration. Furthermore, when we perform photon subtraction on a pure Gaussian state, the resulting photon-subtracted state is also pure, as we saw in Sec. 2. The starting point is the Wigner function of the photon-subtracted state, Eq. (214), which we rewrite as

$$W^-(\vec{x}) = \frac{W(\vec{x})}{\mathrm{tr}\,(V_{\mathbf{b}} - \mathbb{1}) + \|\vec{\xi}_{\mathbf{b}}\|^2} \Big(\|B^T(\mathbb{1} - V^{-1})(\vec{x} - \vec{\xi}) + \vec{\xi}_{\mathbf{b}}\|^2 + \mathrm{tr}\left[\mathbb{1} - B^T V^{-1} B\right] \Big). \tag{259}$$

The state $W^-(\vec{x})$ is thus obtained by subtracting a photon from the Gaussian state $W(\vec{x})$. As we consider a pure two-mode state we assume that the state has a 4×4 covariance matrix of the form $V = S^T S$, where S is a symplectic matrix. We assume that the photon is locally subtracted in one of the modes of the mode basis, such that

$$B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{260}$$

However, to assess the entanglement in the system, we must obtain the Wigner function for the reduced state associated to either of the two modes. When we focus on mode

b where the photon is subtracted, we can simply obtain the reduced photon subtracted state $W_{\mathbf{b}}^{-}(\vec{x}_{\mathbf{b}})$ by subtracting a photon from the reduced Gaussian state $W_{\mathbf{b}}(\vec{x}_{\mathbf{b}})$. As such, we obtain

$$W_{\mathbf{b}}^{-}(\vec{x}_{\mathbf{b}}) = \frac{\|(\mathbb{1} - V_{\mathbf{b}}^{-1})\vec{x}_{\mathbf{b}} + V_{\mathbf{b}}^{-1}\vec{\xi}_{\mathbf{b}}\|^2 + \mathrm{tr}\left[\mathbb{1} - V_{\mathbf{b}}^{-1}\right]}{\mathrm{tr}\,(V_{\mathbf{b}} - \mathbb{1}) + \|\vec{\xi}_{\mathbf{b}}\|^2} W_{\mathbf{b}}(\vec{x}_{\mathbf{b}}). \tag{261}$$

This is now a single-mode photon subtracted state, but it is no longer pure. This lack of purity is notably reflected by $V_{\mathbf{b}}$, which is no longer symplectic. Nevertheless, we can use the Williamson decomposition, Eq. (92), and write

$$V_{\mathbf{b}} = \nu \begin{pmatrix} r & 0 \\ 0 & r^{-1} \end{pmatrix}, \tag{262}$$

where we set the phase such that the squeezing coincides with one of the axes of phase space. What remains is for us to calculate the purity

$$\mu_{\mathbf{b}}^{-} = 4\pi \int_{\mathbb{R}^2} \mathrm{d}\vec{x}_{\mathbf{b}} [W_{\mathbf{b}}^{-}(\vec{x}_{\mathbf{b}})]^2. \tag{263}$$

The final expression for the purity is not very insightful. When on top we use that the purity $\mu_{\mathbf{b}}$ of the Gaussian state $W_{\mathbf{b}}(\vec{x}_{\mathbf{b}})$ is given by $\mu_{\mathbf{b}} = 1/\nu$, an explicit calculation of $\mu_{\mathbf{b}}^{-}$ makes it possible to prove (the motivated reader can use a combination of patience and software for symbolic algebra to do so) that

$$\frac{\mu_{\mathbf{b}}^{-}}{\mu_{\mathbf{b}}} \leqslant \frac{1}{2}. \tag{264}$$

In other words, photon subtraction reduces the purity at most by a factor of 2.

When we use Eq. (256) to define the entanglement of the two-mode photon-subtracted state, Eq. (259), we find that it is given by

$$\mathcal{E}_R(|\Psi^{-}\rangle) = -\log \mu_{\mathbf{b}}^{-}, \tag{265}$$

because the two-mode state is pure. The entanglement of the initial Gaussian state is given by $\mathcal{E}_R(|\Psi_G\rangle) = -\log \mu_{\mathbf{b}}$, such that we can use Eq. (264) to find that

$$\Delta\mathcal{E}_R := \mathcal{E}_R(|\Psi^{-}\rangle) - \mathcal{E}_R(|\Psi_G\rangle) \leqslant \log 2. \tag{266}$$

In other words, photon subtraction can increase the Rényi-2 entanglement of an arbitrary Gaussian state, but at most by an amount $\log 2$. It turns out that this result can be generalized to all bipartitions of Gaussian pure states of an arbitrary number of modes [244]. Furthermore, the same work shows that when the entanglement measure $\mathcal{E}_R(|\Psi_G\rangle)$ is replaced with the Gaussian Rényi-2 entropy of Ref. [215], the result holds for all bipartitions of all Gaussian states (including mixed ones).

For the particular case of a two-mode pure Gaussian state, we can directly evaluate $\Delta\mathcal{E}_R$ for some important examples. Say, for example, that we consider the EPR state that is obtained by mixing two oppositely squeezed vacuum states on a balanced beam splitter. In this case $\vec{\xi} = 0$ and V is given by

$$\begin{aligned} V &= \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}^T \begin{pmatrix} s & & & \\ & s^{-1} & & \\ & & s^{-1} & \\ & & & s \end{pmatrix} \\ &\quad \times \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \\ &= \frac{1}{2s}\begin{pmatrix} s^2+1 & 0 & s^2-1 & 0 \\ 0 & s^2+1 & 0 & 1-s^2 \\ s^2-1 & 0 & s^2+1 & 0 \\ 0 & 1-s^2 & 0 & s^2+1 \end{pmatrix}. \end{aligned} \tag{267}$$

We then extract directly that

$$V_{\mathbf{b}} = \frac{s^2+1}{2s}\,\mathbb{1}, \tag{268}$$

such that we find that the parameters in Eq. (262) are set to $r = 1$ and $\nu = (s^2+1)/(2s)$. And thus we directly obtain

$$\Delta\mathcal{E}_R = \log(2) - \log\left(\frac{s^4 + 6s^2 + 1}{(s^2+1)^2}\right). \tag{269}$$

We clearly see that the entanglement increase vanishes in absence of squeezing, whereas we achieve the $\log(2)$ limit for $s \to \infty$. Adding a mean field with $\vec{\xi}_{\mathbf{b}} \neq 0$ immediately complicates the problem. As can be seen in Fig. 6, where we plot the case $\vec{\xi}_{\mathbf{b}} = (0,1)^T$, the presence of a mean field in the mode of photon subtraction lowers the entanglement increase $\Delta\mathcal{E}_R$. Nevertheless, in the limit $s \to \infty$ we reach the limit $\log(2)$ regardless of the displacement.

This example clearly shows that photon subtraction can be used as a tool to increase entanglement. The setting corresponds to the case that is typically studied in most works on CV entanglement distillation such as Refs. [239–241]. It turns out that one can further increase entanglement in such systems by subtracting more photons. Furthermore, photon addition and the combination of addition and subtraction on both modes have also been considered. The methods we use in this Tutorial are not easily generalized to the subtraction and addition of many photons, but in return they can be applied to a much wider class of initial Gaussian states.

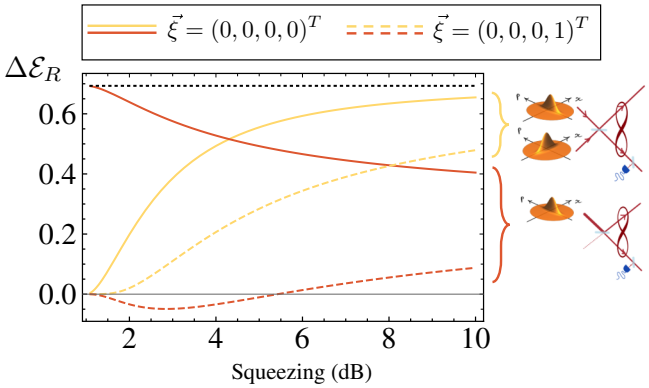

FIG. 6. Entanglement increase, Eq. (266), through photon subtraction in one mode of a pair of entangled modes. The initial Gaussian states are obtained by mixing either two equally squeezed modes (yellow curves) or one squeezed mode and one vacuum mode (red curves) on a beam splitter (see also sketches on the right). We show how a variation of squeezing (in dB compared to shot noise level) in these initial squeezed vacuum states influences the entanglement increase due to photon subtraction. We consider cases without mean field (solid curves) and with a mean field $\vec{\xi} = (0,0,0,1)^T$ (dashed curves).

As a second example, we consider a single-mode squeezed state that is split in two on a balanced beam splitter. This means that the Gaussian state is given by

$$V = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}^T \begin{pmatrix} s & & & \\ & s^{-1} & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} s+1 & 0 & s-1 & 0 \\ 0 & \frac{s+1}{s} & 0 & \frac{1}{s}-1 \\ s-1 & 0 & s+1 & 0 \\ 0 & \frac{1}{s}-1 & 0 & \frac{s+1}{s} \end{pmatrix}, \tag{270}$$

such that we get

$$V_{\mathbf{b}} = \frac{1}{2}\begin{pmatrix} s+1 & 0 \\ 0 & \frac{s+1}{s} \end{pmatrix}, \tag{271}$$

such that we find that we identify the parameters of Eq. (262) as $\nu = \sqrt{2+s+s^{-1}}/2$ and $r = (1+s)/\sqrt{2+s+s^{-1}}$. In absence of any mean field, i.e., with $\vec{\xi} = 0$, we then find an entanglement gain given by

$$\Delta\mathcal{E}_R = \log(2) - \log\left(\frac{3+2s+3s^2}{2(s+1)^2}\right). \tag{272}$$

Interestingly, in this case we reach the maximal entanglement gain for vanishing squeezing $s \to 1$, where we reach $\Delta\mathcal{E}_R \to \log(2)$. This case may seem somewhat counterintuitive, but it should be emphasized that the success probability of photon subtraction also vanishes in this case. Yet, our conditional approach assumes that we are in the scenario where a photon was subtracted and the negligible fraction of the state that is not in vacuum is enhanced. In the limit of vanishing squeezing, the photon subtracted state converges to the Bell state $(|1,0\rangle + |0,1\rangle)/\sqrt{2}$. On the other hand, in the limit where squeezing is high we still find a finite entanglement increase as $\Delta\mathcal{E}_R \to \log(4/3)$.

When we add a mean field given by $\vec{\xi}_{\mathbf{b}} \neq 0$, there is an importance of the phase because our state locally has some remaining asymmetry (which can be seen from $r \neq 1$). In Fig. 6 we particularly show the case where $\vec{\xi}_{\mathbf{b}} = (0,1)^T$ such that the direction of the displacement coincides with the quadrature where the noise is minimal. In this case we observe that for some values of initial squeezing, the entanglement decreases due to photon subtraction. Note that this quite remarkably implies that in some cases photon subtraction can actually be used to increase the purity of a state.

We thus showed that photon subtraction is a useful non-Gaussian operation to increase entanglement. However, in the presence of a mean field in the subtraction mode, it is also possible to decrease entanglement. Even though this subject has been studied for nearly two decades, for arbitrary Gaussian input states, there are still many open questions. Notably, there has not been much work on the effect of photon subtraction on multipartite entanglement, nor on stronger types of quantum correlations. Our discussion in Sec. 2 suggests an important interplay between EPR steering and Wigner negativity, and thus it is intriguing to wonder whether well-chosen non-Gaussian operations can increase quantum steering. Since all steerable states are also entangled, it is a reasonable conjecture that some of the protocols that can increase quantum entanglement should also increase quantum steering.

We have followed the terminology found in the literature and referred to this process as entanglement distillation, because our conditional operation has only a finite success probability. This implies that we can use a large number of Gaussian entangled states and use photon subtraction to obtain a much smaller number of more entangled states. Yet, it must be stressed that there is a more subtle process happening: the entanglement is increased by adding non-Gaussian entanglement on top of the existing Gaussian entanglement. To get a better grasp of this non-Gaussian entanglement, it is useful to go to a setting where no other type of entanglement is present as we do in Sec. 3.

3. Purely non-Gaussian quantum entanglement

In this subsection, we explore an idea that is in many ways complementary to the previous subsection: rather

than using a local non-Gaussian operation to increase already existing entanglement, we now use a nonlocal non-Gaussian operation to create entanglement between unentangled modes.

Let us again assume that our state is initially Gaussian as described by Eq. (85), and we induce the non-Gaussian features through the conditional methods of Sec. B. The mean field of the initial state is given by $\vec{\xi} = \vec{\xi}_{\mathbf{f}} \oplus \vec{\xi}_{\mathbf{g}}$, and

$$V = \begin{pmatrix} V_{\mathbf{f}} & V_{\mathbf{fg}} \\ V_{\mathbf{gf}} & V_{\mathbf{g}} \end{pmatrix}, \quad \text{with} \quad V_{\mathbf{f}} = V_{\mathbf{f_1}} \oplus V_{\mathbf{f_2}}. \tag{273}$$

Here, we have introduced the modes of interest, labeled by $\mathbf{f}$ and a set of auxiliary modes $\mathbf{g}$ upon which a measurement will be performed to induce non-Gaussian features in the modes $\mathbf{f}$. In the initial state, we consider a bipartition in the modes $\mathbf{f}$ without any direct correlations, hence $V_{\mathbf{f}} = V_{\mathbf{f_1}} \oplus V_{\mathbf{f_2}}$. In other words, the modes in $\mathbf{f_1}$ are completely uncorrelated to the modes in $\mathbf{f_2}$.

To induce non-Gaussian effects, we resort to the conditional framework by acting with a POVM element $\hat{A}$ upon the auxiliary modes $\mathbf{g}$, and we rewrite Eq. (190) as

$$W_{\mathbf{f}|\mathbf{A}}(\vec{x}_{\mathbf{f_1}} \oplus \vec{x}_{\mathbf{f_2}}) = \frac{\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f_1}} \oplus \vec{x}_{\mathbf{f_2}}}}{\langle \hat{A} \rangle} W_{\mathbf{f_1}}(\vec{x}_{\mathbf{f_1}}) W_{\mathbf{f_2}}(\vec{x}_{\mathbf{f_2}}), \tag{274}$$

and the conditional expectation value $\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f_1}} \oplus \vec{x}_{\mathbf{f_2}}}$ is again given by Eq. (192). The entanglement in the resulting state thus crucially depends on the exact properties of $\langle \hat{A} \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f_1}} \oplus \vec{x}_{\mathbf{f_2}}}$.

First of all, note that $W_{\mathbf{f_1}}(\vec{x}_{\mathbf{f_1}})$ and $W_{\mathbf{f_2}}(\vec{x}_{\mathbf{f_2}})$ are generally not pure states and as a consequence $W_{\mathbf{f}|\mathbf{A}}(\vec{x}_{\mathbf{f_1}} \oplus \vec{x}_{\mathbf{f_2}})$ is not a pure state either. Even though the specific structure of the Wigner function makes it a suitable case to apply the methods of Ref. [221], we follow a different route in this Tutorial by focusing on a particular example for which we can assume that $W_{\mathbf{f_1}}(\vec{x}_{\mathbf{f_1}})$ and $W_{\mathbf{f_2}}(\vec{x}_{\mathbf{f_2}})$ are pure.

Just as in Sec. 2, we concentrate on photon subtraction. To get a conceptual idea of such a setup in this specific scenario, we present two equivalent schemes in (a) and (b) of Fig. 7. Note that the equivalence stems from the fact that the beam splitters that subtract the light from the signal beams to send it to the photodetector are of extremely low reflectivity. In this limit, we can be sure that there is at most one photon in the path and when it is detected, we herald a single-photon-subtracted state. In Fig. 7(a), the combination of this heralding process and the presence of at most one photon avoids that the unmeasured output causes any losses or impurities. Nevertheless, the unmeasured output will practically change the success probability of the heralding process, such that for practical implementations Fig. 7(b) may be the preferential setup. Recall that the Wigner function for a state with a photon subtracted in a particular mode $\mathbf{b}$ was given by Eq. (259), which here becomes

$$\begin{aligned} W^{-}(\vec{x}_{\mathbf{f_1}} \oplus \vec{x}_{\mathbf{f_2}}) = & \frac{W_{\mathbf{f_1}}(\vec{x}_{\mathbf{f_1}}) W_{\mathbf{f_2}}(\vec{x}_{\mathbf{f_2}})}{\mathrm{tr}\,(V_{\mathbf{b}} - \mathbb{1}) + \|\vec{\xi}_{\mathbf{b}}\|^2} \\ & \times \Big(\|B^T(\mathbb{1} - V_{\mathbf{f_1}}^{-1} \oplus V_{\mathbf{f_2}}^{-1})(\vec{x}_{\mathbf{f_1}} \oplus \vec{x}_{\mathbf{f_2}} \\ & - \vec{\xi}_1 \oplus \vec{\xi}_2) + \vec{\xi}_{\mathbf{b}}\|^2 \\ & + \mathrm{tr}\left[\mathbb{1} - B^T(V_{\mathbf{f_1}}^{-1} \oplus V_{\mathbf{f_2}}^{-1})B\right] \Big). \end{aligned} \tag{275}$$

Because we consider a limit where the state is completely transmitted by the beam splitter and only a negligible amount is sent to the photon counter to subtract the photon, we can indeed assume that the state is pure. For simplicity, we also assume that $\mathbf{f_1}$ and $\mathbf{f_2}$ are single modes. As we did before, we now calculate the reduced state

$$W_1^{-}(\vec{x}_{\mathbf{f_1}}) = \int_{\mathbb{R}^2} \mathrm{d}\vec{x}_{\mathbf{f_2}} W^{-}(\vec{x}_{\mathbf{f_1}} \oplus \vec{x}_{\mathbf{f_2}}). \tag{276}$$

The integral is rather tedious to evaluate, therefore we immediately jump to the result (see Ref. [153] for an alternative method that circumvents the explicit calculation of integrals):

$$\begin{aligned} W_1^{-}(\vec{x}_{\mathbf{f_1}}) = & \frac{W_{\mathbf{f_1}}(\vec{x}_{\mathbf{f_1}})}{\mathrm{tr}\,(V_{\mathbf{b}} - \mathbb{1}) + \|\vec{\xi}_{\mathbf{b}}\|^2} \Big(\|B^T F_1(\mathbb{1} - V_{\mathbf{f_1}}^{-1})(\vec{x}_{\mathbf{f_1}} - \vec{\xi}_1) \\ & + \vec{\xi}_{\mathbf{b}}\|^2 + \mathrm{tr}\left[B^T F_1(\mathbb{1} - V_{\mathbf{f_1}}^{-1}) F_1^T B\right] \\ & + \mathrm{tr}\left[B^T F_2(V_{\mathbf{f_2}} - \mathbb{1}) F_2^T B\right] \Big), \end{aligned} \tag{277}$$

where we introduce the matrices F_k, given by

$$F_k = \begin{pmatrix} | & | \\ \vec{f}_k & \Omega\vec{f}_k \\ | & | \end{pmatrix}, \tag{278}$$

such that we can use the properties of the symplectic form Ω to obtain

$$B^T F_k = \begin{pmatrix} \vec{b}^T\vec{f}_k & \vec{b}^T\Omega\vec{f}_k \\ -\vec{b}^T\Omega\vec{f}_k & \vec{b}^T\vec{f}_k \end{pmatrix}. \tag{279}$$

If mode $\mathbf{b}$ is orthogonal to mode $\mathbf{f}_1$, we find that $B^T F_1 = 0$ such that $W_1^{-}(\vec{x}_{\mathbf{f_1}}) = W_{\mathbf{f_1}}(\vec{x}_{\mathbf{f_1}})$. On the other hand, when mode $\mathbf{b}$ is exactly the same as $\mathbf{f}_1$ we find that $B^T F_1 = \mathbb{1}$ such that the photon is only subtracted there. In this case $W_1^{-}(\vec{x}_{\mathbf{f_1}})$ is a pure state and no entanglement is created. In this case, one can check that $W_2^{-}(\vec{x}_{\mathbf{f_2}}) = W_2(\vec{x}_{\mathbf{f_2}})$.

To create entanglement, we are thus interested in the case where $\mathbf{b}$ is a superposition of the two modes $\mathbf{f}_1$ and $\mathbf{f}_2$. To keep things simple, let us assume that $\vec{b} = \cos\theta\, \vec{f}_1 +$

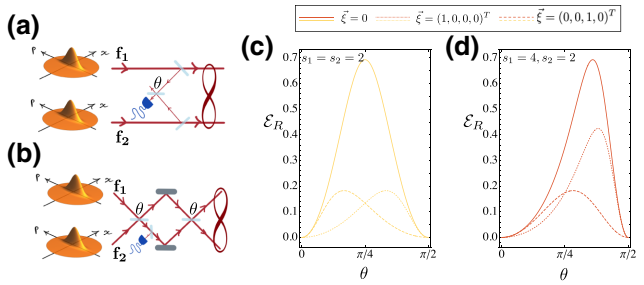

FIG. 7. Entanglement creation through photon subtraction in a superposition of uncorrelated modes $\mathbf{f_1}$ and $\mathbf{f_2}$. (a),(b) Sketches of two equivalent setups to implement a photon subtraction in the mode $\mathbf{b}$, with $\vec{b}=\cos\theta\vec{f}_1+\sin\theta\vec{f}_2$. In (c),(d), we show the created entanglement, as measure through the Rényi entropy, Eq. (256), for varying values of θ. The initial Gaussian states are pure, with covariance matrices $V_{\mathbf{f_1}}=\mathrm{diag}[s_1,1/s_1]$ and $V_{\mathbf{f_2}}=\mathrm{diag}[s_2,1/s_2]$ for modes $\mathbf{f_1}$ and $\mathbf{f_2}$, respectively. The global mean field, i.e., displacement, is varied $\vec{\xi}=0$ (solid curves), $\vec{\xi}=(1,0,0,0)^T$ (dotted curves), and $\vec{\xi}=(0,0,1,0)^T$ (dotted curves). (c) The particular case where the squeezing is balanced, i.e., $s_1=s_2=2$. (d) An unbalanced example where $s_1=4$ and $s_2=2$. All squeezing values s_1 and s_2 are measured in units of vacuum noise.

$\sin\theta\vec{f}_2$. Because the modes $\mathbf{f}_1$ and $\mathbf{f}_2$ are orthogonal, we can use that $\vec{f}_1{}^T\vec{f}_2=0$ and thus we find that $B^TF_1=\cos\theta\,\mathbb{1}$ and $B^TF_2=\sin\theta\,\mathbb{1}$. Nevertheless, the general expression for $W_1^-(\vec{x}_{\mathbf{f_1}})$ does not simplify much.

To acquire additional insight, let us now assume that both modes $\mathbf{f_1}$ and $\mathbf{f_2}$ have exactly the same squeezing in the same quadrature:

$$V_{\mathbf{f_1}}=V_{\mathbf{f_2}}=\begin{pmatrix}s & 0\\ 0 & \frac{1}{s}\end{pmatrix}. \tag{280}$$

Furthermore, let us assume that there is no mean field, such that $\vec{\xi}=0$. In this particular case, we find the expression

$$W_1^-(x_{\mathbf{f_1}},p_{\mathbf{f_1}})=W_1(x_{\mathbf{f_1}},p_{\mathbf{f_1}})\times\left[p_{\mathbf{f_1}}^2s+\frac{x_{\mathbf{f_1}}^2}{s}+\cos(2\theta)\left(p_{\mathbf{f_1}}^2s+\frac{x_{\mathbf{f_1}}^2}{s}-2\right)\right]. \tag{281}$$

In particular, it turns out that the purity takes a simple form, such that we can quantify the entanglement for this state as

$$\mathcal{E}_R=\log(2)-\log\left(\frac{\cos(4\theta)+3}{2}\right). \tag{282}$$

This shows that the maximal entanglement is reached for $\theta=\pi/4$ and—as expected—the entanglement vanishes when $\theta=0$ and $\theta=\pi/2$, i.e., when we subtract entirely in either mode $\mathbf{f_1}$ and $\mathbf{f_2}$.

More general settings are shown in Fig. 7, where we show the entanglement creation for unbalanced squeezing, by setting

$$V_{\mathbf{f_1}}=\begin{pmatrix}s_1 & 0\\ 0 & \frac{1}{s_1}\end{pmatrix},\quad\text{and}\quad V_{\mathbf{f_2}}=\begin{pmatrix}s_2 & 0\\ 0 & \frac{1}{s_2}\end{pmatrix}. \tag{283}$$

We compare the case with $s_1=s_2$ to the case with $s_1\neq s_2$ and find that in absence of a mean field one can reach the same maximal amount of entanglement. However, the maximum is attained at a different value of θ when the squeezing is unbalanced. From Eq. (282) we know that in absence of a mean field, the curve for $s_1=s_2$ does not depend on the actual value of squeezing.

Figure 7 also shows the effect of an existing mean field, by probing a mean field in mode $\mathbf{f_1}$ with $\vec{\xi}=(1,0,0,0)^T$ and in mode $\mathbf{f_2}$ with $\vec{\xi}=(0,0,1,0)^T$. Generally speaking, we observe that the mean field reduces the created entanglement. Nevertheless, the unbalance of squeezing ($s_1\neq s_2$) also unbalances the effect of the mean field. The higher squeezing in mode $\mathbf{f_1}$ makes the entanglement creation more resilient to displacements, but a mean field in mode $\mathbf{f_2}$ will reduce the maximal attainable amount of entanglement to the same level as in the balanced case [because in both (c) and (d) mode $\mathbf{f_2}$ is squeezed with $s_2=2$]. In the presence of a mean field, we also find that unbalanced squeezing shifts the value θ for which most entanglement is created. In other words, to achieve maximal entanglement upon photon subtraction in two modes with unequal squeezing, one must subtract in an unbalanced superposition of these modes.

Through this example, we showed that entanglement between previously uncorrelated Gaussian states can be created by a non-Gaussian operation. This entanglement has some additional peculiarities. For example, a quick glance at how this procedure affects Eq. (190) shows that we can split the state in a Gaussian, i.e., $W_{\mathbf{f}}(\vec{x}_{\mathbf{f}})$, and a non-Gaussian part, i.e., $\langle\hat{A}\rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}}/\langle\hat{A}\rangle$. In this case of Eq. (259) the Gaussian part of the state clearly remains fully separable. This means that, in this representation, all entanglement is originating from the non-Gaussian part of the state. Nevertheless, the decomposition, Eq. (190), of the state into a Gaussian and a non-Gaussian part most probably is not unique for mixed states, making it challenging to study such non-Gaussian entanglement in its most general sense.

Yet, common tools that rely on the covariance matrix, such as Refs. [216,217], to characterize entanglement in the photon subtracted states, Eq. (259), are doomed to fail. In Ref. [152] it is explicitly shown that the covariance matrix of a photon-subtracted state is given by the covariance matrix of the initial Gaussian state with a positive matrix added to it. This means that photon subtraction just adds correlated noise to the covariance matrix and if we consider a Gaussian state that has exactly this covariance matrix we can decompose it using Eq. (103). In other

words, when there is no entanglement visible in the covariance matrix of the initial Gaussian state, we do not witness any entanglement based on the covariance matrix of the photon-subtracted state. In this case, the non-Gaussian entanglement is thus genuinely non-Gaussian in the sense that it cannot be detected through Gaussian witnesses. Hence, rather than decomposing the states in a Gaussian and non-Gaussian part, as was done in Eq. (190), it may be more fruitful to define non-Gaussian entanglement as any entanglement that cannot be witnessed based solely on the covariance matrix of the state. This approach also offers a natural connection to the framework of Gaussian passivity on quantum thermodynamics [245].

Another peculiarity that was presented in Refs. [152, 153] is the intrinsic nature of this non-Gaussian entanglement. When we transform the system into a different mode basis, there will still be entanglement in the system. The entanglement is said to be intrinsic because the state is entangled in every possible mode basis. As we saw in Eq. (249) Gaussian entanglement is never intrinsic as there always exists a basis in which a Gaussian state is separable.

Figure 7(b) gives a rather interesting approach to understanding the intrinsic nature of non-Gaussian entanglement. In this sketch, the second beam splitter is intended to undo the superposition θ and return to the initial mode basis with modes $\mathbf{f_1}$ and $\mathbf{f_2}$. Changing this beam splitter thus implies a basis change. If we remove this beam splitter entirely, we find ourselves in the entanglement distillation scenario of Sec. 2. In this case, the photon subtraction is fully local, but it happens on a state with Gaussian entanglement. The photon subtraction can then increase the Rényi entanglement by a maximal amount of $\log 2$. When we change to a mode basis where there is no Gaussian entanglement and the entanglement is created through a nonlocal photon subtraction, we create a maximal amount of Rényi entanglement given by $\log 2$. Changing the mode basis in a different way will combine the physics of these two extreme cases such that there will always be entanglement, regardless of the basis.

Extending these ideas to more general non-Gaussian operations on more general Gaussian mixed states is a hard and currently open problem. This reflects the general status of entanglement theory in CV systems: we lack a structured theoretical understanding of this phenomenon and as a consequence we also lack good tools to detect it.

D. Non-Gaussianity and Bell inequalities

In this final subsection of our study of quantum correlations in non-Gaussian states, we study Bell inequalities. First of all, we argue that it is impossible to violate Bell inequalities when both of the states and all the measurements involved can be described by positive Wigner functions. Then, we show that the Wigner function of the state can itself be used to formulate a Bell inequality when we allow for nonpositive Wigner functions.

The general setup for studying nonlocality in CV revolves around a multimode state with Wigner function $W(\vec{x}_\mathbf{f} \oplus \vec{x}_\mathbf{g})$ defined on a phase space $\mathbb{R}^{2m} = \mathbb{R}^{2l} \oplus \mathbb{R}^{2l'}$. Bell nonlocality entails that some measurements on this state cannot be described by a local hidden variable model of the type, Eq. (238). In a quantum framework, the local measurements with POVM elements $\{\hat{A}_j\}$ (on the modes in $\mathbf{f}$) and $\{\hat{B}_j\}$ (on the modes in $\mathbf{g}$) can also be described by Wigner functions $W_{A_j}(\vec{x}_\mathbf{f})$ and $W_{B_j}(\vec{x}_\mathbf{g})$. Because we are dealing with a POVM, we find that

$$(4\pi)^l \sum_j W_{A_j}(\vec{x}_\mathbf{f}) = (4\pi)^{l'} \sum_j W_{B_j}(\vec{x}_\mathbf{g}) = 1. \tag{284}$$

Note that this equality holds for all possible coordinates $\vec{x}_\mathbf{f}$ and $\vec{x}_\mathbf{g}$. Here we assume that the measurement outcomes A_j and B_j are discrete, but by correctly defining resolutions of the identity we can also deal with more general probability distributions, e.g., homodyne measurements.

The probability to get the joint measurement result (A_j, B_k) is given by

$$P(A_j, B_k) = (4\pi)^m \int_{\mathbb{R}^{2l}} \int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_\mathbf{f} \mathrm{d}\vec{x}_\mathbf{g} W(\vec{x}_\mathbf{f} \oplus \vec{x}_\mathbf{g}) W_{A_j}(\vec{x}_\mathbf{f}) W_{B_k}(\vec{x}_\mathbf{g}). \tag{285}$$

Now let us assume that all these Wigner functions are positive. Because they are normalized, this implies that $W(\vec{x}_\mathbf{f} \oplus \vec{x}_\mathbf{g})$ is a probability distribution on the entire phase space $\mathbb{R}^{2m}$, and $W_{A_j}(\vec{x}_\mathbf{f})$ and $W_{B_j}(\vec{x}_\mathbf{g})$ are probability distributions on the reduced phase spaces $\mathbb{R}^{2l}$ and $\mathbb{R}^{2l'}$, respectively. However, the model, Eq. (238), does not require probability distributions on phase space, but rather on the possible measurement outcomes.

This is where Eq. (284) comes into play. Because $W_{A_j}(\vec{x}_\mathbf{f})$ and $W_{B_j}(\vec{x}_\mathbf{g})$ are positive, more than just treat them as probability distributions in phase space we can also consider $P_{\vec{x}_\mathbf{f}}(A_j) = (4\pi)^l W_{A_j}(\vec{x}_\mathbf{f})$ and $P_{\vec{x}_\mathbf{g}}(B_j) = (4\pi)^{l'} W_{B_j}(\vec{x}_\mathbf{g})$ as the probability of getting the measurement outcomes A_j and B_k, respectively. Because of Eq. (284) we find that these probabilities are correctly normalized

$$\sum_j P_{\vec{x}_\mathbf{f}}(A_j) = \sum_j P_{\vec{x}_\mathbf{g}}(B_j) = 1, \tag{286}$$

and because the Wigner functions are positive, we also find that $P_{\vec{x}_\mathbf{f}}(A_j), P_{\vec{x}_\mathbf{g}}(B_j) \geqslant 0$. Note that the phase-space coordinates $\vec{x}_\mathbf{f}$ and $\vec{x}_\mathbf{g}$ are no longer treated as the variable, but rather as a label. The set $\{P_{\vec{x}_\mathbf{f}}(A_j) \mid \vec{x}_\mathbf{f} \in \mathbb{R}^{2l}\}$ denotes a family of different probability distributions on

the space of measurement outcomes $\{A_1, A_2, \dots\}$. The set $\{P_{\vec{x}_\mathbf{g}}(B_j) \mid \vec{x}_\mathbf{g} \in \mathbb{R}^{2l'}\}$ can be interpreted analogously.

We can thus recast Eq. (285) in the following form:

$$P(A_j, B_k) = \int \mathrm{d}\vec{x}_\mathbf{f}\mathrm{d}\vec{x}_\mathbf{g} W(\vec{x}_\mathbf{f} \oplus \vec{x}_\mathbf{g}) P_{\vec{x}_\mathbf{f}}(A_j) P_{\vec{x}_\mathbf{g}}(B_k). \quad (287)$$

Because $W(\vec{x}_\mathbf{f} \oplus \vec{x}_\mathbf{g})$ is a positive and normalized Wigner function, it is a joint probability distribution on the coordinates $\vec{x}_\mathbf{f}$ and $\vec{x}_\mathbf{g}$. These coordinates label families of probability distributions $\{P_{\vec{x}_\mathbf{f}}(A_j) \mid \vec{x}_\mathbf{f} \in \mathbb{R}^{2l}\}$ and $\{P_{\vec{x}_\mathbf{g}}(B_j) \mid \vec{x}_\mathbf{g} \in \mathbb{R}^{2l'}\}$ for the measurement outcomes. The expression (287) is thus fully consistent with Bell's local hidden variable model, Eq. (238). As a consequence, we cannot violate any Bell inequalities when the system is prepared in a state with a positive Wigner function and when we only have access to POVM that have Wigner representations with positive Wigner functions.

Let us emphasize that there is generally no reason to assume that the probabilities $P_{\vec{x}_\mathbf{f}}(A_j)$ and $P_{\vec{x}_\mathbf{g}}(B_j)$ are also consistent with quantum mechanics. In other words, there is not necessarily any state $\hat{\rho}$ such that $P_{\vec{x}_\mathbf{f}}(A_j) = \mathrm{tr}[\hat{\rho}\hat{A}_j]$. However, because we are dealing with Bell nonlocality, we do not need this to be the case, since Eq. (238) allows for arbitrary local probability distributions.

To make a long story short, we have shown that Wigner negativity is necessary for witnessing Bell nonlocality. The interested reader can consult works such as Ref. [64] that relate Wigner negativity to the more general concept of quantum contextuality. However, the topic of contextuality in CV systems is still a matter of scientific debate [246].

There has been a significant body of work about the violation of Bell inequalities in CV setups [247–249]. It is evident that this is an arduous task once one approaches a realistic experimental setting [250]. Here, we focus on one particular suggestion to test Bell nonlocality based on a state's Wigner function [251,252].

The starting point of this approach is the CHSH inequality

$$\left|\langle \hat{X}\hat{Y}\rangle - \langle \hat{X}\hat{Y}'\rangle + \langle \hat{X}'\hat{Y}\rangle + \langle \hat{X}'\hat{Y}'\rangle\right| \leqslant 2.$$

As we discussed in Sec. 4, this inequality relies on some assumptions for the observables X, X', Y, and Y'. In particular, we must assume that the measurement outcomes are either -1 or 1. In a CV setting, where we generally deal with a continuum of possible measurement outcomes, this seems like a serious constraint. Nevertheless, we have already encountered some natural examples during this Tutorial. For example, photon counters yield a discrete number of possible measurement outcomes. Here, we choose a related observable that takes us all the way back to Sec. B, where we encountered the observable

$$\hat{\Delta}(\vec{x}) = \hat{D}(-\vec{x})(-\mathbb{1})^{\hat{N}}\hat{D}(\vec{x}).$$

This displaced parity operator has a rich structure, but when it comes to actual measurement outcomes is will return either -1 or 1. This means that we can choose X, X', Y, and Y' to be parity operators. First of all, let us note that

$$\hat{\Delta}(\vec{x}_\mathbf{f} \oplus \vec{x}_\mathbf{g}) = \hat{\Delta}(\vec{x}_\mathbf{f}) \otimes \hat{\Delta}(\vec{x}_\mathbf{g}). \quad (288)$$

To see this, one can first show that $(-\mathbb{1})^{\hat{N}_m} = (-\mathbb{1})^{\hat{N}_l + \hat{N}_{l'}} = (-\mathbb{1})^{\hat{N}_l} \otimes (-\mathbb{1})^{\hat{N}_{l'}}$ and subsequently use $\hat{D}(\vec{x}_\mathbf{f} \oplus \vec{x}_\mathbf{g}) = \hat{D}(\vec{x}_\mathbf{f}) \otimes \hat{D}(\vec{x}_\mathbf{g})$ (displacements in different modes are independent from each other).

Now we can identify the observables as follows:

$$\begin{aligned} X &= \hat{\Delta}(\vec{x}_\mathbf{f}), \quad X' = \hat{\Delta}(\vec{x}'_\mathbf{f}), \\ Y &= \hat{\Delta}(\vec{x}_\mathbf{g}), \quad Y' = \hat{\Delta}(\vec{x}'_\mathbf{g}), \end{aligned} \quad (289)$$

and therefore the CHSH inequality is transformed into

$$\left|\langle \hat{\Delta}(\vec{x}_\mathbf{f} \oplus \vec{x}_\mathbf{g})\rangle - \langle \hat{\Delta}(\vec{x}_\mathbf{f} \oplus \vec{x}'_\mathbf{g})\rangle + \langle \hat{\Delta}(\vec{x}'_\mathbf{f} \oplus \vec{x}_\mathbf{g})\rangle + \langle \hat{\Delta}(\vec{x}'_\mathbf{f} \oplus \vec{x}'_\mathbf{g})\rangle\right| \leqslant 2. \quad (290)$$

As a next step, we use Eq. (78) to write

$$\langle \hat{\Delta}(\vec{x}_\mathbf{f} \oplus \vec{x}_\mathbf{g})\rangle = (2\pi)^m W(\vec{x}_\mathbf{f} \oplus \vec{x}_\mathbf{g}), \quad (291)$$

such that the inequality (290) can be recast as

$$\left|W(\vec{x}_\mathbf{f} \oplus \vec{x}_\mathbf{g}) - W(\vec{x}_\mathbf{f} \oplus \vec{x}'_\mathbf{g}) + W(\vec{x}'_\mathbf{f} \oplus \vec{x}_\mathbf{g}) + W(\vec{x}'_\mathbf{f} \oplus \vec{x}'_\mathbf{g})\right| \leqslant \frac{2}{(2\pi)^m}. \quad (292)$$

Any state with a Wigner function that violates this inequality for some choice of coordinates $\vec{x}_\mathbf{f}, \vec{x}'_\mathbf{f}, \vec{x}_\mathbf{g}$, and $\vec{x}'_\mathbf{g}$ possesses some form of Bell nonlocality. In Refs. [251,252] it is argued that the inequality (292) can be violated by sending a single photon through a beam splitter, but also by an EPR state. The fact that a Gaussian Wigner function suffices to violate Eq. (292) sometimes comes as a surprise, because we previously argued that one needs Wigner negativity to violate Bell inequalities. The reason why one can detect Bell nonlocality with this inequality even when the Wigner function is positive stems from our choice of observable $\hat{\Delta}(\vec{x}_\mathbf{f} \oplus \vec{x}_\mathbf{g})$. The POVM elements that correspond to the measurement outcomes 1 and -1 have Wigner functions that are strongly Wigner negative. As a consequence the necessary Wigner negativity is baked into Eq. (292) by construction.

In practice, the inequality (292) is highly sensitive to impurities and can often be hard to violate with experimentally reconstructed Wigner functions. The hunt for good new techniques to show Bell nonlocality in CV systems

is therefore still open. However, this subsection clearly showed us that Wigner negativity is necessary to observe one of the most exotic features in quantum physics. This negativity might be baked into the state, but it could just as well be induced by measurements. The conditional methods of Sec. B also highlight this duality, where Wigner negativity in the measurement is used to induce Wigner negativity in the state. It should come as no surprise that Wigner negativity is also a necessary ingredient for the most exotic quantum protocols. However, it should also be highlighted that even with a little extra trust, it is possible to design protocols that do not require Wigner negativity to witness quantum correlations [201]. In the next section, we discuss its importance for reaching a quantum advantage with CV systems.

VI. NON-GAUSSIAN QUANTUM ADVANTAGES

It has been long known that systems that are entirely built with Gaussian building blocks are easy to simulate [25]. It should perhaps not come as a surprise that efficient numerical tools exist to sample numbers from a multivariate Gaussian distribution. The discrete variable analog of this result comes across as less intuitive and goes by the name "Gottesman-Knill theorem" [253]. Yet, it turns out that something stronger than mere non-Gaussian elements is required to render a system hard to simulate.

In Sec. D, we encountered the power of Wigner negativity by realizing that it is a necessary requirement for Bell nonlocality. This connection between Wigner negativity and the most exotic types of quantum correlations shows us that Wigner negativity is key to giving CV systems their most prominent quantum features. It is then perhaps not a surprise that such Wigner negativity is also a necessary requirement for implementing any type of protocol that cannot be efficiently simulated by a classical computer [33,254,255]. We thus start our discussion of quantum advantages by explaining the result of Ref. [33]. To show the necessity of Wigner negativity, we show an explicit simulation algorithm for general quantum protocols without Wigner negativity.

Any quantum protocol ultimately relies on the measurement of a certain set of measurement operators $\{\hat{E}_j\}$ (typically a POVM) of a system prepared on a state $\hat{\rho}$. In a Wigner function formalism, we then find

$$p_j = (4\pi)^m \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\; W_{E_j}(\vec{x}) W(\vec{x}). \tag{293}$$

Furthermore, the fact that the set $\{\hat{E}_j\}$ forms a POVM implies that

$$(4\pi)^m \sum_j W_{E_j}(\vec{x}) = 1. \tag{294}$$

As we already discussed in Sec. D, surrounding Eq. (284), it is crucial that the normalization condition (294) holds for any phase-space coordinate $\vec{x}$. In the present context, we want to show that there is an efficient method for a classical device to sample values from the probability distribution $\{p_j\}$ when all involved Wigner functions are positive.

Let us start by assuming that the Wigner functions that describe the POVM elements are all positive. When combined with the POVM condition (294), this implies that we can identify a set of probabilities $P_{\vec{x}}(e_j) = (4\pi)^m W_{E_j}(\vec{x})$ as the probability to obtain the measurement outcome e_j, associated with the POVM element $\hat{E}_j$. These $P_{\vec{x}}(e_j)$ depend on a parameter $\vec{x}$, we can thus form a family of probability distributions $\{P_{\vec{x}}(e_j) \mid \vec{x} \in \mathbb{R}^{2m}\}$ that describe the probability of obtaining the different results e_j, depending on a chosen phase-space point. The normalization condition (294) now states that $\sum_j P_{\vec{x}}(e_j) = 1$ for all $\vec{x}$. Let us emphasize that this family of probabilities would not be well defined if $W_{E_j}(\vec{x})$ were not positive Wigner functions, as some of the probabilities would be negative.

Going back to the initial Eq. (293), we now find that

$$p_j = \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\; P_{\vec{x}}(e_j) W(\vec{x}). \tag{295}$$

To find the actual probability of getting the jth outcome is thus given by "averaging" the probabilities $P_{\vec{x}}(e_j)$ over the different phase-space coordinates. Generally speaking, this is not a real average, unless the Wigner function $W(\vec{x})$ of the state is an actual probability distribution on phase space. The latter is exactly the case when $W(\vec{x})$ is positive. Then, we can simply think of the probability p_j for obtaining event e_j as $p_j = \mathbb{E}_W[P_{\vec{x}}(e_j)]$, where $\mathbb{E}_W$ is the expectation value over the probability distribution $W(\vec{x})$.

Hence, when all Wigner functions are positive, the algorithm to simulate our relevant quantum process can simply be expressed by the following steps:

1. Sample a phase-space coordinate $\vec{x}$ from the probability distribution $W(\vec{x})$.
2. Construct the probability distribution $P_{\vec{x}}(e_j)$ for the sampled value $\vec{x}$.
3. Sample an outcome e_j from the probability distribution $P_{\vec{x}}(e_j)$.

Even though this is the general idea behind our sampling protocol, there are some major hidden assumptions. First, we assume here that the Wigner function for the state and the measurement are known. Furthermore, we also assume that we can simply sample points from any distribution on phase space and from any distribution of measurement outcomes $P_{\vec{x}}(e_j)$. In particular, for the sampling aspects it is not at all clear that these are reasonable assumptions to make. Standard sampling protocols for multivariate probability distributions tend to get highly inefficient once the

probability distributions become too exotic such that it is dangerous to assume that we can "just sample."

To address this point Ref. [33] makes more assumptions on the exact setup we are trying to simulate. First, we assume that the detection is done by a series of single-mode detectors, such that our label j now become a tuple $\mathbf{j} = (j_1, j_2, \ldots, j_m)$ where j_k denotes the outcome $e_{j_k}^{(k)}$ for the detector on the kth mode. We can thus write the POVM element as $\hat{E}_{\mathbf{j}} = \hat{E}_{j_1}^{(1)} \otimes \cdots \otimes \hat{E}_{j_m}^{(m)}$, such that

$$W_{E_{\mathbf{j}}}(\vec{x}) = W_{E_{j_1}^{(1)}}(x_1, p_1) W_{E_{j_2}^{(2)}}(x_2, p_2) \ldots W_{E_{j_m}^{(m)}}(x_m, p_m). \tag{296}$$

We assume that each detector has been accurately calibrated, such that all the individual Wigner functions are known. This implies that for a given point in phase space $\vec{x} = (x_1, p_1, \ldots, x_m, p_m)^T$, we can simply evaluate the probabilities for each detector to produce a certain outcome. Thus, we calculate the probability distributions $P_{(x_k,p_k)}(e_j^{(k)}) = 4\pi W_{E_j^{(k)}}(x_k, p_k)$ for all the possible measurement outcomes for that specific mode. We assume that sampling outcomes $e_j^{(k)}$ from these probability distributions $P_{(x_k,p_k)}(e_j^{(k)})$ is a feasible task. For typical detectors in quantum optics experiments this is a very reasonable assumption.

The Wigner function $W(\vec{x})$ that describes the state is more subtle as it also includes all correlations between modes. If the state is Gaussian, a measurement of the covariance matrix would be sufficient to know the full Wigner function. Because of its Gaussian features, there are efficient tools to directly sample phase-space points. Yet, for more general non-Gaussian positive Wigner functions this sampling may be much harder. Therefore, Ref. [33] makes an essential assumption: it assumes that we know a protocol that combines local operations to design the state $W(\vec{x})$ from a known initial state with no correlations between the modes. The notion of "locality" should here be understood in the sense of acting on a small set of modes while leaving the others fully untouched. These local operations are also supposed to be represented by positive Wigner functions, which depend only on the phase-space coordinates of the subset of modes on which they act.

Generally speaking, such Wigner positive operations $\Xi : \mathcal{H}^{\text{in}} \to \mathcal{H}^{\text{out}}$ map a state $\hat{\rho}$ to a new state $\Xi[\hat{\rho}]$. In Ref. [33], the Choi representation [256,257] is used to represent Ξ as a state on a larger Hilbert space $\mathcal{H}^{\text{in}} \otimes \mathcal{H}^{\text{out}}$. This becomes particularly appealing when we go to a Wigner representation, where the Choi representation of Ξ is given by a Wigner function $W_\Xi(\vec{x}^{\,\text{in}} \oplus \vec{x}^{\,\text{out}})$. The action of Ξ on a state with Wigner function $W(\vec{x}^{\,\text{in}})$ is then given by

$$W_{\text{out}}(\vec{x}^{\,\text{out}}) = (4\pi)^m \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}^{\,\text{in}}\, W_\Xi(\vec{x}^{\,\text{in}} \oplus T\vec{x}^{\,\text{out}}) W(\vec{x}^{\,\text{in}}), \tag{297}$$

where m is the number of modes of the input state. For technical reasons, we must include the transposition operator T, Eq. (257), in the action of the channel. Because this operation must be trace preserving, we on top get the property that

$$(4\pi)^m \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}^{\,\text{out}}\, W_\Xi(\vec{x}^{\,\text{in}} \oplus T\vec{x}^{\,\text{out}}) = 1. \tag{298}$$

When we now assume that the operation Ξ has a Wigner-Choi representation $W_\Xi(\vec{x}^{\,\text{in}} \oplus \vec{x}^{\,\text{out}})$, which is a positive function, it immediately follows that the operation Ξ turns a Wigner positive initial state $W(\vec{x}^{\,\text{in}})$ into a Wigner positive output state $W_{\text{out}}(\vec{x}^{\,\text{out}})$.

It is useful to note that such operations, Eq. (297), can be trivially embedded in a larger space. Let us assume that we consider a state $W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$, we can simply let the operation act on the modes $\mathbf{g}$ by taking

$$\begin{aligned} W_{\text{out}}(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}^{\,\text{out}}) = (4\pi)^{l'} \int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}}^{\,\text{in}}\, W_\Xi(\vec{x}_{\mathbf{g}}^{\,\text{in}} \oplus T\vec{x}_{\mathbf{g}}^{\,\text{out}}) \\ \times W(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}}^{\,\text{in}}). \end{aligned} \tag{299}$$

Notationally, this may seem a little complicated, but, in essence, we just carry out the integration over a subset of the full phase space. We call these operations local Wigner positive operations.

In our simulation protocol, we thus assume that $W(\vec{x})$ is created by a series of such local Wigner positive operations of $\Xi_1, \ldots, \Xi_t$ on a noncorrelated input state $W_{\text{in}}(\vec{x}_{\text{in}}) = W_{\text{in}}^{(1)}(x_1, p_1) W_{\text{in}}^{(2)}(x_2, p_2) \ldots W_{\text{in}}^{(m)}(x_m, p_m)$.

$$\begin{aligned} W(\vec{x}) = (4\pi)^{mt} \int_{\mathrm{R}^{2m}} \mathrm{d}\vec{x}_t \ldots \int_{\mathrm{R}^{2l}} \mathrm{d}\vec{x}_t W_{\Xi_t}(\vec{x}_t \oplus T\vec{x}) \ldots \\ \times W_{\Xi_2}(\vec{x}_1 \oplus T\vec{x}_2) W_{\Xi_1}(\vec{x}_{\text{in}} \oplus T\vec{x}_1) \\ \times W_{\text{in}}(\vec{x}_{\text{in}}). \end{aligned} \tag{300}$$

We assume on top that each operation is local over a small number of modes $l \ll m$. To model this with Eq. (299), it suffices to split $\vec{x}_{t_{k-1}} = \vec{x}^{\,l}_{t_{k-1}} \oplus \vec{x}^{\,l'}_{t_{k-1}}$ and $\vec{x}_{t_k} = \vec{x}^{\,l}_{t_k} \oplus \vec{x}^{\,l'}_{t_k}$, such that

$$\begin{aligned} &(4\pi)^m W_{\Xi_{t_k}}(\vec{x}_{t_{k-1}} \oplus T\vec{x}_{t_k}) \\ &\quad = (4\pi)^l W_{\Xi_{t_k}}(\vec{x}^{\,l}_{t_{k-1}} \oplus T\vec{x}^{\,l}_{t_k}) \delta(\vec{x}^{\,l'}_{t_{k-1}} - \vec{x}^{\,l'}_{t_k}). \end{aligned} \tag{301}$$

Even though the notation is complicated, it simply describes that we act on an l-mode subspace with the operation Ξ_{t_k} and leave the other l' modes untouched.

The normalization condition (298) now has an important consequence, since it allows us to identify a probability distribution on phase space $P_{\vec{x}_{t_{k-1}}}(\vec{x}_{t_k}) = (4\pi)^m W_{\Xi_{t_k}}(\vec{x}_{t_{k-1}} \oplus T\vec{x}_{t_k})$. It gives us the probability of choosing a phase space value $\vec{x}_{t_k}$, given that we know $\vec{x}_{t_{k-1}}$. Because the operations are local, Eq. (301) allows us to keep most of the phase-space coordinates constant from step to step. Furthermore, the first step is simple. Every pair (x_k, p_k) of the initial coordinate $\vec{x}_{\text{in}}$ can be sampled independently because $W_{\text{in}}(\vec{x}_{\text{in}})$ factorizes. This now gives us the following new algorithm:

1. Take the initial Wigner function $W_{\text{in}}(\vec{x}^{\text{ in}}) = W_{\text{in}}^{(1)}(x_1,p_1)W_{\text{in}}^{(2)}(x_2,p_2)\ldots W_{\text{in}}^{(m)}(x_m,p_m)$ and sample a pair (x_k,p_k) from every single-mode probability distribution $W_{\text{in}}^{(k)}(x_k,p_k)$. Put all these pairs together to obtain $\vec{x}^{\text{ in}} = (x_1,p_1,\ldots,x_m,p_m)$.
2. Update the coordinate by sampling new coordinates based on $P_{\vec{x}_{t_{k-1}}}(\vec{x}_{t_k}) = (4\pi)^m W_{\Xi_{t_k}}(\vec{x}_{t_{k-1}} \oplus T\vec{x}_{t_k})$. Because Ξ_{t_k} are local operations, it suffices to only locally update coordinates. Let us make this clear through an example. Say we have $\vec{x}_{t_{k-1}} = (x_1^{(k-1)},p_1^{(k-1)},\ldots,x_m^{(k-1)},p_m^{(k-1)})^T$ and operation Ξ_{t_k} acts locally on modes with labels $2,5$, and 7. Take $\vec{x}^{\,l}_{t_{k-1}} = (x_2^{(k-1)},p_2^{(k-1)},x_5^{(k-1)},p_5^{(k-1)},x_7^{(k-1)},p_7^{(k-1)})$ and use it to evaluate $P_{\vec{x}_{t_{k-1}}}(\vec{x}^{\,l}_{t_k}) = (4\pi)^l W_{\Xi_{t_k}}(\vec{x}^{\,l}_{t_{k-1}} \oplus T\vec{x}^{\,l}_{t_k})$. Now sample a new vector $\vec{x}^{\,l}_{t_k} = (x_2^{(k)},p_2^{(k)},x_5^{(k)},p_5^{(k)},x_7^{(k)},p_7^{(k)})$ from this probability distribution. Then construct the new vector $\vec{x}_{t_k}$ by taking $\vec{x}_{t_{k-1}}$ and updating the coordinates associated to modes $2,5$, and 7 to the newly sampled coordinates
3. After the operations $\Xi_1,\ldots,\Xi_t$ have been implemented by updating the phase-space coordinate, take the final phase-space coordinate $\vec{x} = (x_1,p_1,\ldots,x_m,p_m)^T$ and the Wigner function describing the detectors $W_{E_{\mathbf{j}}}(\vec{x}) = W_{E_{j_1}^{(1)}}(x_1,p_1)W_{E_{j_2}^{(2)}}(x_2,p_2)\ldots W_{E_{j_m}^{(m)}}(x_m,p_m)$. For each detector k, use the phase-space coordinate $\vec{x}$ to generate the probability distribution $P_{(x_k,p_k)}(e_j^{(k)}) = W_{E_j^{(k)}}(x_k,p_k)$.
4. Sample an outcome $e_j^{(k)}$ from the distribution $P_{(x_k,p_k)}(e_j^{(k)})$ for every detector.

Sampling the final phase-space coordinate $\vec{x}$ by using a Monte-Carlo-style update rule is time consuming, but if the operations are local it can be done efficiently. This procedure implicitly assumes that we do not just know the state we are sampling from, but that we know the circuit of local operations that is used to create the state from local resources. Ultimately, when one considers the circuit representation of quantum algorithms, this is also how a quantum algorithm works. For example, Sec. A exactly shows that any unitary CV circuit can be built with single- and two-mode gates. The algorithm outlined in this section shows that we can efficiently simulate any protocol where the local input state, the circuit's operations, and the measurements are described by positive Wigner functions.

One may wonder whether any positive Wigner function $W(\vec{x})$ can be constructed through such a circuit and, if so, whether there is an efficient way to design such a circuit when we know the Wigner function. If we assume that not only the state $W(\vec{x})$ but also all its marginals are known, it is possible to construct a stepwise sampling procedure through the chain rule of probability theory:

$$W(\vec{x}) = W(x_m,p_m \mid x_{m-1},p_{m-1},\ldots x_1,p_1)\times\ldots \\ W(x_3,p_3 \mid x_2,p_2,x_1,p_1)W(x_2,p_2 \mid x_1,p_1)W(x_1,p_1). \tag{302}$$

This process effectively executes a type of random walk with memory. In each step of this walk, we then sample the phase-space coordinates for one mode. Nevertheless, this process only works when we have access to all these conditional probabilities, which practically implies having access to all the marginals of the distribution. In practical setups, this will often not be the case. Nevertheless, it is quickly seen that this setup can be efficiently used to sample from Gaussian Wigner functions where these conditional distributions have a particularly simple form.

Thus, we have shown that it is impossible to obtain a quantum computational advantage by using only local states, measurements, and operations with positive Wigner functions. This means that Wigner negativity is necessary to reach a quantum advantage in such setups. However, Wigner negativity is certainly not sufficient since there are many setups of quantum systems that involve negative Wigner functions that can be efficiently simulated [258]. It is thus interesting to take the opposite approach and explore a setup that is known to lead to a quantum advantage. In the spirit of CV setups, the most logical choice for such a discussion is Gaussian boson sampling [259]. In the literature, this setup has been studied mainly from the point of view of complexity theory [260,261], but here we rather focus on its physical building blocks.

Boson sampling [262] is a problem in which one injects a set of N bosons (generally photons) into an m-mode interferometer. On the output ports of this interferometer, photodetectors are mounted to count the particles at the output. Simulating this type of quantum Galton board is a computationally hard task, implying that a quantum advantage could be reached by implementing the setup in a quantum optics experiment. On the other hand, it turns out that the required number of photons to implement such an experiment is also hard to come by. This was the motivation for developing a new approach, where the input photons are replaced by squeezed states that are injected

in each of the interferometer inputs. Because an interferometer, built out of phase shifters and beam splitters, is a Gaussian transformation the output state will remain Gaussian. We can thus effectively say that we are sampling photons from a state with Wigner function $W_G(\vec{x})$. In addition, there is no mean field in the setup such that the entire state is characterized by its covariance matrix V.

When we assume that the detectors resolve photon numbers, the probability to detect a string of counts $\mathbf{n} = (n_1, \ldots, n_m)$ can we written as

$$P(\mathbf{n}) = (4\pi)^m \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\; W_{\mathbf{n}}(\vec{x}) W_G(\vec{x}). \tag{303}$$

We can then use Eq. (118) to write

$$W_{\mathbf{n}}(\vec{x}) = W_{n_1}(x_1, p_1) \ldots W_{n_m}(x_m, p_m). \tag{304}$$

Even though the integral (303) is hard to compute, it is insightful in the light of Eq. (295) and our discussion regarding the necessity of Wigner negativity. Indeed, we see immediately that the detectors form a crucial element in rendering the setup hard to simulate. The same holds when we replace the number-resolving detectors with their on-off counterparts [263] such that $n_k = \{0, 1\}$ and the Wigner functions are given by $W_{n_k}(x_k, p_k) = \{1 - 2\exp[-(x_k^2 + p_k^2)/2]\}/(4\pi)$.

When we stick with number-resolving detectors that project on Fock states, it is practical to reformulate the problem in terms of P functions and Q functions, such that

$$P(\mathbf{n}) = \int_{\mathbb{R}^{2m}} \mathrm{d}\vec{x}\; P_{\mathbf{n}}(\vec{x}) Q_G(\vec{x}). \tag{305}$$

For the detailed calculation, we refer to Ref. [259]. It turns out that the probabilities $P(\mathbf{n})$ can be expressed in terms of the Hafnian of a matrix [264], which establishes a connection to the problem of finding perfect matchings in graph theory. This connection has led to several suggested applications for Gaussian Boson sampling [265–267].

In the light of this Tutorial, the most interesting application of Gaussian Boson sampling is its potential role in quantum state engineering [205]. When only a subset of modes are measured, we can see Gaussian Boson sampling as a generalization of photon subtraction (and even as a generalization of "generalized photon subtraction" [157]). The idea is reasonably simply explained in the light of Sec. B: when we split the system in two parts $\mathbb{R}^{2m} = \mathbb{R}^{2l} \oplus \mathbb{R}^{2l'}$, such that the Gaussian state that comes out of the interferometer now takes the form $W_G(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$, we can postselect on a measurement outcome $\mathbf{n} = (n_1, \ldots n_{l'})$ for the second subsystem. We thus project on a state $W_{\mathbf{n}}(\vec{x}_{\mathbf{g}})$, which is a product of l' Fock states, and from Eq. (190), we obtain that the conditional state on the remaining modes is given by

$$W_{\mathbf{f}|\mathbf{n}}(\vec{x}_{\mathbf{f}}) = \frac{\langle [|n_1\rangle \langle n_1| \otimes \cdots \otimes |n_{l'}\rangle \langle n_{l'}|] \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}}}{\langle [|n_1\rangle \langle n_1| \otimes \cdots \otimes |n_{l'}\rangle \langle n_{l'}|] \rangle} W_{\mathbf{f}}(\vec{x}_{\mathbf{f}}), \tag{306}$$

with $W_{\mathbf{f}}(\vec{x}_{\mathbf{f}})$ defined by Eq. (182). From Eq. (192) we recall the expression

$$\begin{aligned} &\langle [|n_1\rangle \langle n_1| \otimes \cdots \otimes |n_{l'}\rangle \langle n_{l'}|] \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}} \\ &\quad = (4\pi)^{l'} \int_{\mathbb{R}^{2l'}} \mathrm{d}\vec{x}_{\mathbf{g}}\; W_{\mathbf{n}}(\vec{x}_{\mathbf{g}}) W_G(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}}), \end{aligned} \tag{307}$$

and because the initial state $W_G(\vec{x}_{\mathbf{f}} \oplus \vec{x}_{\mathbf{g}})$ is Gaussian, we find that the conditional probability distribution $W_G(\vec{x}_{\mathbf{g}} \mid \vec{x}_{\mathbf{f}})$ is given by Eq. (193). Ironically, to evaluate $\langle [|n_1\rangle \langle n_1| \otimes \cdots \otimes |n_{l'}\rangle \langle n_{l'}|] \rangle_{\mathbf{g}|\vec{x}_{\mathbf{f}}}$ and $\langle [|n_1\rangle \langle n_1| \otimes \cdots \otimes |n_{l'}\rangle \langle n_{l'}|] \rangle$ we must essentially solve the same hard problem as for the implementation of Gaussian Boson sampling itself. Therefore, the exact description of the resulting states is generally complicated.

Nevertheless, in idealized scenarios, even small Gaussian Boson sampling circuits can be used to prepare interesting non-Gaussian states [205]. In particular, the capacity of Gaussian Boson sampling to produce GKP states has taken up a prominent place in a recent blueprint for photonic quantum computation [24]. Furthermore, the results in Sec. C suggest that states created by performing Gaussian Boson sampling on a subset of modes can have additional non-Gaussian entanglement. Yet, to be able to use this procedure to produce highly resourceful Wigner negative states, Sec. 2 highlights that the initial Gaussian state needs to be such that the modes in $\mathbf{f}$ can steer the modes in $\mathbf{g}$. This condition can be seen as a basic quality requirement for the Gaussian Boson samplers that are used in Ref. [24].

Finally, the experimental imperfections are also detrimental for the quantum advantage that is produced in Gaussian Boson sampling. Clearly, when the Gaussian state $W_G(\vec{x})$ can be written as a Gaussian mixture of coherent states (meaning that no mode basis exists in which the quadrature noise is below vacuum noise), the sampling can be simulated efficiently. Because multimode coherent states are always just a tensor product of single-mode coherent states, it suffices to sample a coherent state from the mixture, calculate all the individual probabilities for the output detectors, and sample independent detector outputs according to these probabilities. The presence of entanglement in the Gaussian state from which we sample is thus crucial. In addition, detector efficiencies must be sufficiently high such that their Wigner functions remain nonpositive, otherwise the protocol of Ref. [33] renders the setup easy to simulate (as explained in the first part

of this section). A more thorough analysis of how different experimental imperfections render Gaussian Boson sampling easier to simulate can be found in Ref. [268].

There are clearly still many aspects of the relation between non-Gaussian features of quantum states on the one hand, and the ability to achieve a quantum computational advantage on the other hand, that are not yet fully understood. The Gaussian Boson sampling setup clearly emphasizes the importance of entanglement in combination with Wigner negativity. Furthermore, there is the implicit fact that a simulation scheme such as Ref. [33] requires knowledge of the actual circuit of local operations that was used to create the state. It does make sense to assume that we actually have some ideas of the quantum protocol that we are attempting to simulate, but yet one may wonder whether there could be a reasonable setting (in the sense that we are actually implementing a well-controlled protocol) in which the assumptions of Ref. [33] do not hold. This clearly shows that many fundamental theoretical aspects of CV quantum computation remain to be uncovered.

VII. EXPERIMENTAL REALIZATIONS

Now that we have provided an overview of some theoretical aspects of non-Gaussian quantum states, we interpret the "where to find them" part of the title in a very literal sense. Non-Gaussian states are generally rather fragile, as one should expect from quantum central limit theorem and the fact that thermal states in free bosonic theories are Gaussian. Producing and analyzing non-Gaussian states in a laboratory setting is indeed challenging, but nevertheless it has been done numerous times. Our main focus in Sec. A is quantum optics, which is the historical testbed for CV quantum physics. However, in recent years there has been increased attention for CV approaches in other setting such as optomechanics, superconducting circuits, and trapped ions.

A. Quantum optics experiments

This section provides an overview of some of the most important milestones in the generation of non-Gaussian states in optics. For more details, we refer the reader to a specialized review [81].

Historically, one might argue that the first experimental realizations of non-Gaussian states in optical setups relied on sufficiently sensitive photon detectors. Initial demonstrations primarily used photoemission of atoms [269,270], which are prepared in excited states (e.g., by electron bombardment) or via resonance fluorescence in ions [271]. The development of spontaneous parametric down-conversion (SPDC) made it possible to create a single-photon state using only bulk optical elements [202]. However, all these early non-Gaussian states were characterized through counting statistics, which means that we generally classify them as DV experiments.

It is perhaps intriguing to note that SPDC is also the process that lies at the basis of the creation of squeezed states of light [272], which are Gaussian. These states play a key role in the generation of single-photon states, simply because a weakly squeezed vacuum is mainly a superposition of vacuum and a photon pair. By detecting one photon of the pair, the presence of the second photon is heralded. Hence, the approach of Ref. [202] is a basic implementation of a conditional scheme for the generation of non-Gaussian states as presented in Sec. B.

A genuine CV treatment of such non-Gaussian states would only be achieved much later in a work that presents the first tomographic reconstruction of a state with Wigner negativity in optics [273]. Due to the developments of an easily implementable maximum-likelihood algorithm for state reconstruction, homodyne tomography became one of the main tools to study non-Gaussian states in CV quantum optics [274]. It did not take long before this also led to the reconstruction of a displaced single-photon Fock state [275] and a two-photon Fock state [276]. The combination of increased squeezing with type-II SPDC and an array of photon detectors to increase the number of heralded photons more recently made it possible to resolve the Wigner function of a three-photon Fock state [277]. Similar ideas of multiplexed photon detection have also been used to generate superpositions of Fock states [278].

For non-Gaussian states beyond Fock states, photon subtraction, as described in Sec. 2, is a common experimental tool. Its first experimental implementation successfully showed the capability of generating non-Gaussian statistics in the homodyne measurements, but it failed to demonstrate Wigner negativity [148]. Later experiments improved the quality of the generated states, demonstrating Wigner negativity and creating so-called "Schrödinger kittens" [129,279,280]. The terminology is chosen because these states resemble cat states proportional to $|\alpha\rangle - |-\alpha\rangle$ for small values of the mean field α. Even though such Schrödinger kittens are ultimately not very different from squeezed single-photon states, the nomenclature makes more sense in the context of experiments that "breed" cat states [281]. Here, one mixes two Schrödinger kittens on a beam splitter and performs homodyne detection on one output port. By conditioning on instances where this homodyne detector registers values close to zero, one effectively heralds a larger cat state (the value of α has increased). A variation of photon subtraction has also been used to create a type of CV qubit [282].

As an alternative to photon subtraction, one can also add a photon [159]. Even though this operation theoretically equates to applying a creation operator on the state, it is experimentally much harder to implement than photon subtraction as it requires nonlinear optics. However, photon subtraction can only produce Wigner negativity

when the initial state is squeezed. Photon addition, on the other hand, provides the advantage of always creating a Wigner negative state. A simple way to see this is by applying a creation operator to the state and evaluating the Q function (70). When a photon is added to the mode g, the Q function after photon addition has the property $Q^+(\vec{\alpha}) \sim (\vec{\alpha}^T\vec{g})^2 Q_G(\vec{\alpha})$, where $Q_G(\vec{\alpha})$ is the Q function of the initial Gaussian state. This relation implies automatically that the Q function will be exactly zero for $\vec{\alpha} = \vec{0}$, and a zero of the Q function implies Wigner negativity. This means that one can apply photon addition to highly classical states, such as a coherent state or a thermal state, and still end up creating Wigner negativity. Such photon-added coherent states were also used to experimentally measure [283] non-Gaussianity $\delta(\hat{\rho})$ as defined in Eq. (110). Remarkably, combining photon addition and photon subtraction operations in both possible orders provides a way to experimentally verify the canonical commutation relations $[\hat{a}, \hat{a}^\dagger] = 1$, as was shown in Ref. [160].

The above methods are all based on Gaussian states as initial resources to generate non-Gaussian states. The non-Gaussian states that are created as such can in turn serve as useful resources to create more intricate non-Gaussian states. Fock states are a commonly used type of input state, for example, in the first demonstration of a large Schrödinger's cat state [284]. Intriguingly, by using non-Gaussian initial states, it suffices to use homodyne detection as the conditional operation. This setup can then be extended to a cat breeding scheme [285]. Another method to create large cat states in optics relies on making the light field interact with an atom [286]. The presence of entanglement between the "macroscopic" coherent state and the "microscopic" atomic degrees of freedom make for an experiment that resembles Schrödinger's original though experiment [106]. Once the atom and the coherent light are entangled, a spin rotation of the atom is followed by a measurement to project the state of the light field in either an even or an odd cat state. This reflects the general idea that atoms still induce much larger nonlinearities than nonlinear crystals. These nonlinearities are the direct source of non-Gaussian effect, but they are also much harder to control. At present, experiments that rely on such higher-order nonlinearities to create non-Gaussian states remain rare in the optical regime.

The above methods all focus on the creation of single-mode non-Gaussian states. For multimode systems, much of the experimental progress has concentrated on two-mode systems. As we extensively discussed throughout this tutorial, an important feature in such multimode systems are quantum correlations. Some of the first experimental demonstrations of non-Gaussian quantum correlations were based on the Bell inequality (292). Homodyne tomography and a single photon, delocalized over two modes by a beam splitter, suffices to violate the inequality [287,288]. However, these works also teach us that extreme high purities are required to do so.

Motivated by photon-subtraction experiments and challenged by the no-go theorem of [225–227], entanglement distillation soon became a new focus for non-Gaussian quantum optics experiments. Some of these experiments have focused on adding some form of non-Gaussian noise on the initial state to circumvent the no-go theorem [228, 229]. Entanglement distillation through local photon subtraction from the entangled modes of a Gaussian input state would later be demonstrated in Ref. [289]. Earlier, it had already been shown that Gaussian entanglement can be increased by photon subtraction in a superposition of the entangled mode [290]. Interestingly, in the latter case, the photon is effectively subtracted in a nonentangled mode such that the setup is essentially equivalent to mixing a squeezed vacuum and a photon-subtracted squeezed vacuum on a beam splitter. A similar photon subtraction in a coherent superposition of modes was later carried out to entangle two Schrödinger kittens [291]. This can probably be seen as the first realizations of purely non-Gaussian entanglement in CV.

Photon addition has also been considered as a tool for creating entanglement between pairs of previously uncorrelated modes [292]. The resulting state can be seen as a hybrid entangled state proportional to $|0\rangle|\alpha\rangle + |1\rangle|-\alpha\rangle$, such states have also been produced using techniques similar to photon subtraction [293]. For two modes, photon addition can be implemented in a mode-selective way [294]. This setup is particularly useful to create entanglement between coherent states by adding a photon in a superposition of displaced modes.

Going beyond two modes has always remained a challenging task. For mode-selective photon subtraction from a multimode field, one must abandon the typical implementation based on a beam splitter. For two modes, such an alternative photon subtraction scheme was, for example, realized in the time-frequency domain, by subtracting a photon from a sideband [295]. Yet, going to a genuine multimode scenario required the design of a whole new photon subtractor based on sum-frequency generation [296,297]. This finally permitted the first demonstration of multimode non-Gaussian state in a CV setting, demonstrating non-Gaussian features in up to four entangled modes [298].

Such highly multimode states of more than two modes are confronted with a considerable problem: the exponential scaling of the required number of measurements for a full state tomography. This makes it highly challenging to demonstrate non-Gaussian features such as Wigner negativity in multimode non-Gaussian states. For single-photon-subtracted states, it has been pointed out that good analytical models can be used to train machine-learning

algorithms to recognize Wigner negativity based on single-mode measurements [299]. Furthermore, the techniques of Ref. [128] combined with Ref. [300] should also make it possible to use multiplexed double homodyne detection to witness Wigner negativity in certain classes of multimode states.

In multimode systems, we are confronted with the limitations of homodyne tomography. Recently, it has been shown that machine-learning techniques can be used to implement an improved form of CV tomography based on homodyne measurements [301]. Even though this setup is computationally heavy in single-mode setups, it uses a smaller set of states as a basis for state reconstruction, which might make multimode versions of the protocol more scalable. Alternatively, one can also bypass homodyne measurements all together. Photon-number-resolving detectors such as transition edge sensors [302] make it possible to use the identity, Eq. (78), to directly measure the Wigner function [76]. Intriguingly, this implies that a photon-number-resolving detector and a setup to generate displacements of the state in arbitrary modes makes it possible to directly measure the full multimode Wigner function. Nevertheless, such a multimode protocol has so far not been realized in any experiment.

B. Other experimental setups

Given all the experimental work in CV quantum optics, it is perhaps surprising that the first experimental demonstrations of quantum states with Wigner negativity happened in different fields. The very first realization of such a state was achieved with trapped ions. Even though one often uses the atomic transitions in these systems to isolate qubits for potential quantum computers, trapped ions also have interesting motional degrees of freedom. By exploiting a Jaynes-Cummings type interaction between the atom and the trapping field, it is possible to use the ions' internal atomic degrees of freedom to create well-controlled non-Gaussian states such as a Fock state [78] and a Schrödinger's cat state [303] in the motional degrees of freedom.

Mathematically, this setup is equivalent to cavity QED, where it was shown that photons in a cavity can be manipulated through interactions with atoms [304] and the Rabi oscillations of the injected Rydberg atoms can in turn be used to probe the field within the cavity [305]. These methods would then be combined to experimentally generate a single-photon Fock state of the microwave field in a cavity [306], confirm its Wigner negativity [79], and probe its full Wigner function [80]. A few years later, similar techniques were used to finally generate Schrödinger cat states and higher-order Fock states [307].

A third setup with very similar physics is found in circuit QED. In this field, the macroscopic microwave cavities are replaced by superconducting circuits, and nonlinearities are induced by Josephson junctions rather than atoms [140]. Even though these setups are often used in a DV approach, the microwave fields involved can equally be treated in a CV approach. The large nonlinearities rather naturally create non-Gaussian states, but getting a good sense of control over them can be challenging. Nevertheless, a wide range of non-Gaussian states such as Fock state [308] and large Schrödinger cat states [309] have been experimentally realized. The latter have furthermore been stabiized by engineering the decoherence processes in the system [310]. Very recently these systems have also been used to demonstrate the deterministic generation of photon triplets [142].

In recent years, both, trapped ions [311] and superconducting circuits [312] were used to achieve another important milestone in CV quantum computing: the experimental generation of a GKP state. These highly non-Gaussian states are useful for encoding a fault-tolerant qubit in a CV degree of freedom. By exploiting the redundancy that is offered by the infinite dimension Hilbert space of a CV system, one can create a qubit with a certain degree of robustness. This effectively makes it possible to implement error-correction routines, as shown in Ref. [312]. In other words, these systems have managed to generate CV states that are so non-Gaussian that they can be effectively used as fault-tolerant DV states.

A final field that has shown much potential over the last decades is cavity optomechanics. Here, an optical field is injected into a cavity with one moving mirror (more generally also other types of "dynamic cavities" can be used). The goal is to cool this mirror to its ground state to observe its quantum features. This way, one hopes to create nonclassical states of motion in reasonably large objects. A wide variety of such optomechanical devices exist [313]. Several theoretical schemes have been proposed to generate non-Gaussian states in such an optomechanical setup [314,315]. Even though quantum features such as photon-phonon entanglement have been demonstrated in such systems [316,317], it remains highly challenging to obtain good experimental control over the motional quantum state. Nevertheless, some CV non-Gaussian states states in the form of superpositions between vacuum and a single-phonon Fock state have been experimentally realized [318].

A common problem in these setups is the creation of entanglement between the CV degrees of freedom in different modes. Some degree of such CV entanglement has been experimentally achieved in trapped ion [319] and circuit QED setups [320]. However, the number of entangled modes is much lower than what has been achieved in optics [18–21,220], where even non-Gaussian entangled states of more than two modes have been created [298]. This shows clearly how different experimental setups have different strengths and weaknesses. Optics comes with the advantage of spatial, temporal, and spectral mode

manipulations, which allows the creation of large entangled states. However, the resilience of optical setups to decoherence is due to limited interaction with the environment. The latter implies that it is also difficult to find controlled ways to make these systems strongly non-Gaussian. On the other hand, the other setups, which we discussed, require much more significant shielding from environmental degrees of freedom. When this coupling to other degrees of freedom can be controlled, it provides the means to create non-Gaussian quantum states. In this context, it is appealing to combine the advantages of different regimes. Optomechanics offers a potential pathway to achieve this by converting between microwave and optical degrees of freedom [321,322].

As a last remark, it is interesting to mention that phase-space descriptions and non-Gaussian states also appear in atomic ensembles. This framework relies on the fact that an ensemble of a large number of atoms can be described by collective observables that behave very similar to bosonic systems. The associated phase space behaves differently from the optical phase space, in the sense that it is compact. More specifically, the phase space will cover a sphere and the radius of this sphere will depend on the number of atoms. Effectively, we would recover a bosonic system in the limit of an infinite number of atoms. However, the compactness of phase space for a finite ensemble comes with interesting side effects: a sufficiently high amount of spin squeezing can create non-Gaussian states. We do not go into details for these systems, but it should nevertheless be highlighted that non-Gaussian spin states have received considerable attention in the literature [323] and have been produced in a range of experiments [324–327].

VIII. CONCLUSIONS AND OUTLOOK

In this Tutorial, we have presented a framework based on phase-space representations to study continuous-variable quantum systems. We then focused on the various aspects of non-Gaussian states, where we first represented different ways to structure the space of continuous-variable states in a single mode in Fig. 1. Whenever possible, we generalized results from the literature to a multimode setting. However, for certain properties such as the stellar rank, these generalizations become insufficient to classify all possible quantum states.

We introduced two paradigms to create non-Gaussian states, where one is a deterministic approach based on unitary transformations, reminiscent of the circuit approach for quantum-information processing. The second approach is conditional, in the sense that it relies on conditioning one part of a state on measurement outcomes for another part of a state, which is more narrowly related to a measurement-based approach to quantum protocols. Throughout the remainder of the Tutorial, we have largely focused on conditional operations, since it is the most commonly used approach in experiments. It also provides a natural avenue to start studying the relation between quantum correlations and non-Gaussian features. We show how the conditional approach requires certain correlations in the initial Gaussian state to be able to induce certain type of non-Gaussianity in the conditional state, as summarized in Fig. 5.

On the other hand, non-Gaussian operations can also create a type of non-Gaussian entanglement as introduced in Sec. C. This kind of entanglement is particular as it can not be identified with typical techniques that rely on the state's covariance matrix. Nevertheless, we use Rényi-2 entanglement as a measure to illustrate the existence of such purely non-Gaussian quantum correlations in photon-subtracted states. Even though its existence is known from pure-state examples, it has only received limited attention in both theoretical and experimental work. One possible reason is the difficulty of studying this type of entanglement for mixed states, since convex roof constructions tend to become highly intractable for non-Gaussian states.

As a final theoretical aspect of the Tutorial, we highlight the need of Wigner negativity to achieve some of the most striking features in quantum technologies. On the one hand, we show that Wigner negativity in either the state or the measurement is necessary to violate a Bell inequality. This observation can be understood in the broader context of nonlocality and contextuality: Wigner negativity is often seen as a manifestation of the contextual behavior of quantum systems, and nonlocality can be understood as a type of contextuality of measurements on different subsystems. On the other hand, we also present results that show how Wigner negativity is a requirement to achieve a quantum computational advantage. Intuitively, it is perhaps not surprising that states, operations, and measurements that can all be described by probability distributions on phase space can be efficiently simulated on a classical computer. However, as we showed in Sec. VI, the actual simulations protocol contains many subtle points. Here, too, we conclude that there are still many open questions surrounding the physics of quantum computational advantages in continuous-variable setups.

As a last step of this Tutorial, we provided an overview of the experimental realizations of non-Gaussian states with continuous variables. Quantum states of light are indeed the usual suspects for continuous-variable quantum-information processing, but it turns out to be remarkably challenging to engineer highly non-Gaussian states in such setups. We highlighted how trapped ions, cavity QED, and circuit QED have proven to be better equipped for this task, but in return they are confronted with other problems. Optomechanics presents itself as an ideal translator between these two regimes, which may soon make it possible to combine the scalability of

optical setups with the high nonlinearities of the microwave domain.

In a more general sense, there are definitely many open question to be resolved in the domain of continuous-variable quantum physics. In this Tutorial, we have focused extensively on questions related to non-Gaussian features, notably in multimode systems. In the greater scheme of things, this is only one of the many challenges in the field. The recent demonstration of a quantum computational advantage with Gaussian Boson sampling has set an important milestone for continuous-variable quantum technologies [23], but we are still far away from useful computational protocols as set out in the roadmap of Ref. [24]. Even though the quest for a Gottesman-Kitaev-Preskill state [31] is one of the main experimental priorities, there are still many open challenges in designing the Gaussian operations that form the basis of such a setup [328,329].

Beyond universal fault-tolerant quantum computers, there are many other potential applications for continuous-variable systems. They are widely used in quantum communications for quantum key distribution [330] and and secret sharing [331]. These protocols are largely based on Gaussian states and measurements, such that also the best possible attacks to these systems are Gaussian [332]. Nevertheless, non-Gaussian protocols for quantum key distribution, based on photon subtraction, have been proposed [333]. Such non-Gaussian quantum computation protocols and their security still involve many open questions.

Continuous-variable systems also provide a natural link to other bosonic systems, which is why they have been suggested as a platform to simulate molecular vibronic spectra [334]. The continuous-variable approach also plays an important role in quantum algorithms for other chemistry-related problems such as drug discovery through molecular docking [335] and the simulation of electron transport [336].

Furthermore, the continuous-variable setting is also suitable to implement certain elements for quantum machine learning such as quantum neural networks [337]. Even though this is a promising platform for tackling a wide range of problems, the proposal is highly ambitious on several points. In the context of this Tutorial, we emphasize the need of non-Gaussian unitary transformations. In principle, neural networks require linear couplings between different "neurons," which each implement some form of nonlinear operation. Non-Gaussian operations play the role of this nonlinear element, making them a crucial step in the scheme. To implement such continuous-variable neural networks we thus require either new developments on the implementation of non-Gaussian operations, or theoretical modifications in the protocol to make it fit for implementable conditional non-Gaussian operations. It should be highlighted that other machine-learning approaches exist, such as reservoir computing, which can be entirely based on Gaussian states [338].

A final quantum technology that may benefit from the use of non-Gaussian states is quantum metrology, as was recently demonstrated with motional Fock states of trapped ions [339]. Even though early work has shown that there is no clear benefit in using non-Gaussian operations such as photon subtraction for parameter estimation [154], there may still be other settings where such states are beneficial. Non-Gaussian entanglement could, for example, have a formal metrological advantage that is reflected in the quantum Fisher information [340]. On the other hand, ideas from quantum metrology also provide a possible approach for measuring non-Gaussian quantum steering [341]. The effects of non-Gaussian features on the sensitivity of the state can in principle be captured by higher moments of the quadrature operators [342]. It was recently shown that postselected measurements could, indeed, offer a quantum advantage for metrology [343]. This result is narrowly connected to the field of weak measurements and makes a connection to yet another phase-space representation: the Kirkwood-Dirac distribution [70,344]. Hence, we circle back to the fundamental physics of continuous-variable systems and conclude that there are still many connections to be made.

Beyond the technological applications that continuous-variable systems may have to offer, there is an important down-to-earth perspective that must be emphasized. With the improvement of detectors throughout the years, we have reached a point where theory and experiment can be considered mature to tackle single-mode problems. In multimode systems, the same cannot be said. With the exponential scaling of standard homodyne tomography, experimental tools for studying large multimode states beyond the Gaussian regime are limited. We may have to accept that the full quantum state is out of reach for experimental measurements. Even theoretically, highly multimode Wigner functions quickly become cumbersome to handle. Treating them with numerical integration techniques becomes a near-impossible task, once the number of modes is drastically increased. This makes even numerical simulations challenging. How then can we understand and even detect the non-Gaussian features of these systems?

One clear and important future research goal in this field is to provide an answer to this question. For quantum technologies, this may provide us with new ways to benchmark our systems, but more fundamentally it might teach us something new about the physics of these systems. One place where one might look for inspiration is the field of statistical mechanics, where statistical methods show that even highly complex systems can produce clear emergent signatures. We recently took a first step in exploring such ideas by looking at emergent network structures for continuous-variable non-Gaussian states [345]. The most exciting lesson from such preliminary work is that there

is still much to be learned about non-Gaussian quantum states.

ACKNOWLEDGMENTS

First, I want to express sincere gratitude to the anonymous referees for providing very thorough reports and many useful suggestions that significantly improved the clarity and completeness of this Tutorial. I also thank M. Genoni for several useful suggestions. More generally speaking, the content of this text was influenced by stimulating discussions throughout the years with many colleagues, notably F. Grosshans, R. Filip, Q. He, M. Gessner, and G. Ferrini. Furthermore, I am very grateful to M. Fannes for teaching me the mathematical foundations that lie at the basis bosonic quantum systems. This point of view was complemented by colleagues in the multimode quantum optics group of the Laboratoire Kastler Brossel, V. Parigi, N. Treps, and C. Fabre, who have introduced me to the wonderful world of experimental quantum optics. Still, the main source of inspiration for this Tutorial are the many excellent students and postdocs that I have worked with in the last few years. Their questions and struggles have been essential to highlight the barriers that I try to overcome in this Tutorial. Among these students and postdocs, I want to express explicit gratitude to U. Chabaud, G. Sorelli, and D. Barral for their careful and detailed reading of the Tutorial and for their useful comments. I also acknowledge the many useful discussions with K. Zhang, who notably made me aware of the possibility of reducing entanglement through photon subtraction (here shown in Fig. 7). Last but not least, I want to express special thanks to C. Lopetegui, who started reading this Tutorial as a newcomer to the field of CV quantum optics and thus had the perfect point of view to help fine tune the content.

APPENDIX: MATHEMATICAL REMARKS

Here we present some important well-known mathematical concepts that are regularly used in the Tutorial to make the text more self-contained. The comments and definitions given here are not very rigorous and mainly aim at giving the reader an intuitive understanding, for a more formal introduction one should consult a standard textbook [41,346,347].

1. Topological vector spaces

Throughout this Tutorial, we often deal implicitly with topological vector spaces. Vector spaces are well known from linear algebra and can be thought of as sets of mathematical objects called vectors, which can be added together in a commutative way and multiplied by scalars. When we consider a vector space $\mathcal{V}$ on a field $\mathcal{F}$, this means that for any $\vec{v}_1, \vec{v}_2 \in \mathcal{V}$ and any $\alpha_1, \alpha_2 \in \mathcal{F}$ the object $\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 \in \mathcal{V}$. This means that the vector space is closed under addition and scalar multiplication. In this Tutorial, the field $\mathcal{F}$ is either identified as $\mathbb{R}$ (for phase space) or $\mathbb{C}$ (for Hilbert spaces).

The spaces that are considered in the Tutorial have much more structure than what is given by the vector space. First of all, we generally deal with normed spaces, which means that our vector spaces are topological vector spaces in the sense that there is a notion of distance defined upon them. Generally speaking, topological vector spaces can be equipped with exotic topologies, but here we simply deal with norms. On top, we again add an additional structure when we assume that these norms are generated by inner products (depending on exact properties, these inner products go by different names such as "positive-definite sesquilinear form", which is what we typically consider in quantum mechanics).

As we deal with infinite-dimensional spaces to describe bosonic quantum states and Fock space, it is important to set some terminology straight. When we talk about a Hilbert space, there is the assumption that the space is complete. In an infinite-dimensional inner-product space, we can define sequences of elements in $\mathcal{V}$. If we consider a sequence $(\vec{v}_j)_{j\in\mathbb{N}}$ such that for any ϵ we can find a value $N > 0$ such that $\|\vec{v}_j - \vec{v}_k\| < \epsilon$ for all $j, k > N$, we call the sequence a Cauchy sequence. In other words, the distance between elements in the Cauchy sequence shrinks as we proceed further into the sequence. The fact that a Hilbert space is closed means that all Cauchy sequences converge in the sense that $\lim_{j\to\infty} \vec{v}_j = \vec{v} \in \mathcal{V}$. Finite-dimensional inner-product spaces automatically have this property, but for infinite-dimensional spaces it must be imposed explicitly.

Another class of structured vector space, that is often encountered in the Tutorial, is a real symplectic space. These spaces appear when we consider phase space, and they are given by a real vector space with an additional symplectic form σ instead of the usual inner product. In the mathematical literature, one often encounters the notation $(\mathcal{V}, \sigma)$ for a symplectic space, where the symplectic form has the following properties: we consider $\vec{v}_1.\vec{v}_2 \in \mathcal{V}$ and find that $\sigma(\vec{v}_1, \vec{v}_2) \in \mathbb{R}$, σ is bilinear, and $\sigma(\vec{v}_1, \vec{v}_2) = -\sigma(\vec{v}_2, \vec{v}_1)$. In all cases in this Tutorial, we also consider that σ is nondegenerate, which means that $\sigma(\vec{v}_1, \vec{v}_2) = 0$ for all $\vec{v}_1 \in \mathcal{V}$ if and only if $\vec{v}_2 = \vec{0}$. When the symplectic space is finite dimensional, it is often practical to represent the symplectic form in terms of a matrix. In the Tutorial this is done by associating $\sigma(\vec{v}_1, \vec{v}_2) = \vec{v}_1^T \Omega \vec{v}_2$.

In principle, a real symplectic space is all that is needed to develop the mathematical framework of the CCR algebra. However, it is often natural when dealing with bosonic systems to include an additional structure in the form of an inner product. In the Tutorial, this is done implicitly by also using the standard inner product $\vec{v}_1^T\vec{v}_2$ on phase space. This allows us to ultimately get the isomorphism (44). In the quantum statistical mechanics literature, it is common to

see references to a "pre-Hilbert space", rather than a phase space or a symplectic space. When we refer to a pre-Hilbert space, we consider an inner-product space, which is not necessarily complete and one must consider the closure to be guaranteed to obtain a full Hilbert space. The reason is that phase space, as a real vector space $\mathcal{V}$ with an inner product, given by a bilinear form $s(.,.)$, and a symplectic form $\sigma(.,.)$ is equivalent to a complex pre-Hilbert space $\mathcal{H}$. For finite-dimensional spaces, the equivalence between the vector spaces is obtained via isomorphism (44):

$$\vec{f} \in \mathcal{V} \mapsto |\psi_f\rangle = \sum_j (f_{2j-1} + if_{2j}) |\varphi_j\rangle \in \mathcal{H}, \tag{A1}$$

where $|\varphi_j\rangle$ for a basis of $\mathcal{H}$. As we are talking about an isomorphism between structured vectors spaces, we also need an identity between additional structures, which is given by

$$\langle \psi_{f_1} \mid \psi_{f_2} \rangle = s(\vec{f}_1, \vec{f}_2) - i\sigma(\vec{f}_1, \vec{f}_2). \tag{A2}$$

This isomorphism holds very generally and can be extended to infinite-dimensional spaces. It provides a very formal connection between the single-particle Hilbert space for a many-boson system and its phase space associated with the modes of the bosonic field. Technically, we note that the phase space is equivalent to a pre-Hilbert space, and the closure of this space is the single-particle Hilbert space. Whenever the phase space (and thus the single-particle Hilbert space) is finite dimensional, the pre-Hilbert space is closed such that phase space and single-particle Hilbert space really are equivalent. For a very rigorous treatment on all these points, we refer to Ref. [42].

2. Span

Throughout the Tutorial, we often refer to the "span" of a certain set of vectors. These vectors can be members of a vector space, symplectic space, topological vectors space, pre-Hilbert space, or Hilbert space, the definition of the span is always the same. Let us here assume that $\mathcal{V}$ denotes any type of vector space over a field $\mathcal{F}$ and consider a set $\vec{v}_1, \ldots \vec{v}_n \in \mathcal{V}$. We can now define the span of this set of vectors as the set of all linear combinations that can be made with these vectors

$$\text{span}\{\vec{v}_1, \ldots \vec{v}_n\} := \{\alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n \mid \alpha_1, \ldots, \alpha_n \in \mathcal{F}\}. \tag{A3}$$

We emphasize that there is no need for the set $\vec{v}_1, \ldots \vec{v}_n$ to form a basis, nor for the vectors to be linearly independent, nor for the vectors to be normalized, nor for the vectors to be orthogonal to one another.

Throughout the Tutorial, the vector spaces we encounter are either real (in the case of phase space) such that $\mathcal{F} = \mathbb{R}$ or complex (in the case of Hilbert spaces for quantum systems) such that $\mathcal{F} = \mathbb{C}$. In the case where the vector spaces have some topological structure (which we can colloquially understand as a mathematical sense of distance that allows limits to be defined), it can make sense to consider the closure of a span, denoted by

$$\overline{\text{span}\{\vec{v}_1, \ldots \vec{v}_n\}}, \tag{A4}$$

such that any convergent sequence built with elements of the span has its limit also included in the closure.

[1] E. Schrödinger, Der stetige übergang von der mikro- zur makromechanik, Naturwissenschaften **14**, 664 (1926).

[2] H. Araki and E. J. Woods, Representations of the canonical commutation relations describing a nonrelativistic infinite free bose gas, J. Math. Phys. **4**, 637 (1963).

[3] D. W. Robinson, The ground state of the Bose gas, Commun.Math. Phys. **1**, 159 (1965).

[4] A. Verbeure, *Many-Body Boson Systems: Half a Century Later*, Theoretical and mathematical physics (Springer, London ; New York, 2011).

[5] R. J. Glauber, Coherent and incoherent states of the radiation field, Phys. Rev. **131**, 2766 (1963).

[6] E. C. G. Sudarshan, Equivalence of Semiclassical and Quantum Mechanical Descriptions of Statistical Light Beams, Phys. Rev. Lett. **10**, 277 (1963).

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

[8] 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).

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

[10] J. Williamson, On the algebraic problem concerning the normal forms of linear dynamical systems, American J. Math. **58**, 141 (1936).

[11] S. L. Braunstein, Squeezing as an irreducible resource, Phys. Rev. A **71**, 055801 (2005).

[12] M. M. Wolf, G. Giedke, and J. I. Cirac, Extremality of Gaussian Quantum States, Phys. Rev. Lett. **96**, 080502 (2006).

[13] A. S. Holevo, M. Sohma, and O. Hirota, Capacity of quantum gaussian channels, Phys. Rev. A **59**, 1820 (1999).

[14] M. Yukawa, R. Ukai, P. van Loock, and A. Furusawa, Experimental generation of four-mode continuous-variable cluster states, Phys. Rev. A **78**, 012301 (2008).

[15] X. Su, Y. Zhao, S. Hao, X. Jia, C. Xie, and K. Peng, Experimental preparation of eight-partite cluster state for photonic qumodes, Opt. Lett. **37**, 5178 (2012).

[16] D. Barral, M. Walschaers, K. Bencheikh, V. Parigi, J. A. Levenson, N. Treps, and N. Belabas, Versatile Photonic Entanglement Synthesizer in the Spatial Domain, Phys. Rev. Appl. **14**, 044025 (2020).

[17] J. Roslund, R. M. de Araújo, S. Jiang, C. Fabre, and N. Treps, Wavelength-multiplexed quantum networks with ultrafast frequency combs, Nat Photon **8**, 109 (2014).

[18] M. Chen, N. C. Menicucci, and O. Pfister, Experimental Realization of Multipartite Entanglement of 60 Modes of a Quantum Optical Frequency Comb, Phys. Rev. Lett. **112**, 120505 (2014).

[19] Y. Cai, J. Roslund, G. Ferrini, F. Arzani, X. Xu, C. Fabre, and N. Treps, Multimode entanglement in reconfigurable graph states using optical frequency combs, Nat. Commun. **8**, 15645 (2017).

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

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

[22] Z. Yang, M. Jahanbozorgi, D. Jeong, S. Sun, O. Pfister, H. Lee, and X. Yi, A squeezed quantum microcomb on a chip, arXiv:2103.03380 [physics.optics] (2021).

[23] H.-S. Zhong *et al.*, Quantum computational advantage using photons, Science 370, 1460 (2020), https://science.sciencemag.org/content/370/6523/1460.full.pdf.

[24] J. E. Bourassa, R. N. Alexander, M. Vasmer, A. Patil, I. Tzitrin, T. Matsuura, D. Su, B. Q. Baragiola, S. Guha, G. Dauphinais, K. K. Sabapathy, N. C. Menicucci, and I. Dhand, Blueprint for a scalable photonic fault-tolerant quantum computer, Quantum **5**, 392 (2021).

[25] S. D. Bartlett, B. C. Sanders, S. L. Braunstein, and K. Nemoto, Efficient Classical Simulation of Continuous Variable Quantum Information Processes, Phys. Rev. Lett. **88**, 097904 (2002).

[26] S. Lloyd and S. L. Braunstein, Quantum Computation Over Continuous Variables, Phys. Rev. Lett. **82**, 1784 (1999).

[27] N. C. Menicucci, P. van Loock, M. Gu, C. Weedbrook, T. C. Ralph, and M. A. Nielsen, Universal Quantum Computation with Continuous-Variable Cluster States, Phys. Rev. Lett. **97**, 110501 (2006).

[28] M. Gu, C. Weedbrook, N. C. Menicucci, T. C. Ralph, and P. van Loock, Quantum computing with continuous-variable clusters, Phys. Rev. A **79**, 062318 (2009).

[29] N. C. Menicucci, Fault-Tolerant Measurement-Based Quantum Computing with Continuous-Variable Cluster States, Phys. Rev. Lett. **112**, 120504 (2014).

[30] F. Arzani, N. Treps, and G. Ferrini, Polynomial approximation of non-gaussian unitaries by counting one photon at a time, Phys. Rev. A **95**, 052352 (2017).

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

[32] B. Q. Baragiola, G. Pantaleoni, R. N. Alexander, A. Karanjai, and N. C. Menicucci, All-Gaussian Universality and Fault Tolerance with the Gottesman-Kitaev-Preskill Code, Phys. Rev. Lett. **123**, 200502 (2019).

[33] A. Mari and J. Eisert, Positive Wigner Functions Render Classical Simulation of Quantum Computation Efficient, Phys. Rev. Lett. **109**, 230503 (2012).

[34] R. Hudson, When is the Wigner quasi-probability density non-negative?, Rep. Math. Phys. **6**, 249 (1974).

[35] F. Soto and P. Claverie, When is the Wigner function of multidimensional systems nonnegative?, J. Math. Phys. **24**, 97 (1983).

[36] A. Mandilara, E. Karpov, and N. J. Cerf, Extending Hudson's theorem to mixed quantum states, Phys. Rev. A **79**, 062302 (2009).

[37] L. c. v. Lachman, I. Straka, J. Hloušek, M. Ježek, and R. Filip, Faithful Hierarchy of Genuine *n*-Photon Quantum Non-Gaussian Light, Phys. Rev. Lett. **123**, 043601 (2019).

[38] U. Chabaud, D. Markham, and F. Grosshans, Stellar Representation of Non-Gaussian Quantum States, Phys. Rev. Lett. **124**, 063605 (2020).

[39] M. Walschaers, V. Parigi, and N. Treps, Practical framework for conditional non-gaussian quantum state preparation, PRX Quantum **1**, 020305 (2020).

[40] D. Petz, *An Invitation to the Algebra of Canonical Comutation Relations*, Leuven notes in mathematical and theoretical physics Series A No. 2 (Leuven Univ. Press, Leuven, 1990).

[41] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics 1* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1987).

[42] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics Equilibrium States. Models in Quantum Statistical Mechanics* (Springer, Berlin, 1997).

[43] M. H. Stone, Linear transformations in Hilbert space: III. Operational methods and group theory, PNAS **16**, 172 (1930).

[44] J. von Neumann, Die eindeutigkeit der Schrödingerschen operatoren, Math. Ann. **104**, 570 (1931).

[45] M. H. Stone, On one-parameter unitary groups in Hilbert space, Ann. Math. **33**, 643 (1932).

[46] J. von Neumann, Über einen satz von herrn M. H. Stone, Ann. Math. **33**, 567 (1932).

[47] M. C. Tichy, F. Mintert, and A. Buchleitner, Essential entanglement for atomic and molecular physics, J. Phys. B: At. Mol. Opt. Phys. **44**, 192001 (2011).

[48] F. Benatti, R. Floreanini, F. Franchini, and U. Marzolino, Entanglement in indistinguishable particle systems, Phys. Rep. **878**, 1 (2020), entanglement in indistinguishable particle systems.

[49] N. Killoran, M. Cramer, and M. B. Plenio, Extracting Entanglement from Identical Particles, Phys. Rev. Lett. **112**, 150501 (2014).

[50] N. Killoran, F. E. S. Steinhoff, and M. B. Plenio, Converting Nonclassicality Into Entanglement, Phys. Rev. Lett. **116**, 080402 (2016).

[51] R. Lo Franco and G. Compagno, Indistinguishability of Elementary Systems as a Resource for Quantum Information Processing, Phys. Rev. Lett. **120**, 240403 (2018).

[52] B. Morris, B. Yadin, M. Fadel, T. Zibold, P. Treutlein, and G. Adesso, Entanglement between Identical Particles is a Useful and Consistent Resource, Phys. Rev. X **10**, 041012 (2020).

[53] M. C. Tichy, Interference of identical particles from entanglement to boson-sampling, J. Phys. B: At. Mol. Opt. Phys. **47**, 103001 (2014).

[54] Note that the notation is not a coincidence, we can indeed construct a formal linear map from the single-particle Hilbert space into the operator algebra of linear operators on the Fock space. This mapping takes single-particle state vectors $\psi \in \mathcal{H}$ and maps them to a creation operators. The mapping is linear in the sense that $\hat{a}^\dagger(x\psi + y\varphi) = x\hat{a}^\dagger(\psi) + y\hat{a}^\dagger(\varphi)$, for all $x, y \in \mathbb{C}$.

[55] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995).

[56] M. Walschaers, Signatures of many-particle interference, J. Phys. B: At. Mol. Opt. Phys. **53**, 043001 (2020).

[57] W. P. Schleich, *Quantum Optics in Phase Space* (Wiley VCH, Berlin, 2001).

[58] C. Fabre and N. Treps, Modes and states in quantum optics, Rev. Mod. Phys. **92**, 035005 (2020).

[59] A. M. O. de Almeida, The weyl representation in classical and quantum mechanics, Phys. Rep. **295**, 265 (1998).

[60] In the literature, one will encounter various different choices for the symplectic structure, which correspond to different forms of ordering the amplitude and phase quadratures. In general, any matrix J that satisfies the conditions $J^T = -J$ and $J^2 = -\mathbb{1}$ defines a symplectic structure. A popular alternative to the ordering $(E_1^{(x)}, E_1^{(p)}, \ldots, E_m^{(x)}, E_m^{(p)})$ that we follow in this Tutorial is the ordering $(E_1^{(x)}, \ldots, E_m^{(x)}, E_1^{(p)}, \ldots, E_m^{(p)})$. The choice in this Tutorial is motivated by the study of entanglement, where it is more convenient to group the quadratures that correspond to the same modes together.

[61] E. Wigner, On the quantum correction for thermodynamic equilibrium, Phys. Rev. **40**, 749 (1932).

[62] M. Hillery, R. O'Connell, M. Scully, and E. Wigner, Distribution functions in physics: Fundamentals, Phys. Rep. **106**, 121 (1984).

[63] W. P. Schleich, in *Quantum Optics in Phase Space* (John Wiley and Sons, Ltd, 2001), Chap. 3, p. 67.

[64] R. W. Spekkens, Negativity and Contextuality are Equivalent Notions of Nonclassicality, Phys. Rev. Lett. **101**, 020401 (2008).

[65] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Berlin, Heidelberg, 1955).

[66] H. Weyl, Quantenmechanik und gruppentheorie, Zeitschrift für Physik **46**, 1 (1927).

[67] J. B. Conway, in *A Course in Functional Analysis* (Springer New York, New York, NY, 1985), p. 310.

[68] M. Plancherel, Contribution à l'étude de la représentation d'une fonction arbitraire par des intégrales définies, Rendiconti del Circolo Matematico di Palermo (1884-1940) **30**, 289 (1910).

[69] K. E. Cahill and R. J. Glauber, Density operators and quasiprobability distributions, Phys. Rev. **177**, 1882 (1969).

[70] P. A. M. Dirac, On the analogy between classical and quantum mechanics, Rev. Mod. Phys. **17**, 195 (1945).

[71] J. Sperling and W. Vogel, Quasiprobability distributions for quantum-optical coherence and beyond, Physica Scripta **95**, 034007 (2020).

[72] A. Royer, Wigner function as the expectation value of a parity operator, Phys. Rev. A **15**, 449 (1977).

[73] K. Banaszek and K. Wódkiewicz, Direct Probing of Quantum Phase Space by Photon Counting, Phys. Rev. Lett. **76**, 4344 (1996).

[74] K. Banaszek, C. Radzewicz, K. Wódkiewicz, and J. S. Krasiński, Direct measurement of the wigner function by photon counting, Phys. Rev. A **60**, 674 (1999).

[75] K. Laiho, K. N. Cassemiro, D. Gross, and C. Silberhorn, Probing the Negative Wigner Function of a Pulsed Single Photon Point by Point, Phys. Rev. Lett. **105**, 253603 (2010).

[76] R. Nehra, A. Win, M. Eaton, R. Shahrokhshahi, N. Sridhar, T. Gerrits, A. Lita, S. W. Nam, and O. Pfister, State-independent quantum state tomography by photon-number-resolving measurements, Optica **6**, 1356 (2019).

[77] L. G. Lutterbach and L. Davidovich, Method for Direct Measurement of the Wigner Function in Cavity qed and ion Traps, Phys. Rev. Lett. **78**, 2547 (1997).

[78] D. Leibfried, D. M. Meekhof, B. E. King, C. Monroe, W. M. Itano, and D. J. Wineland, Experimental Determination of the Motional Quantum State of a Trapped Atom, Phys. Rev. Lett. **77**, 4281 (1996).

[79] G. Nogues, A. Rauschenbeutel, S. Osnaghi, P. Bertet, M. Brune, J. M. Raimond, S. Haroche, L. G. Lutterbach, and L. Davidovich, Measurement of a negative value for the wigner function of radiation, Phys. Rev. A **62**, 054101 (2000).

[80] P. Bertet, A. Auffeves, P. Maioli, S. Osnaghi, T. Meunier, M. Brune, J. M. Raimond, and S. Haroche, Direct Measurement of the Wigner Function of a One-Photon Fock State in a Cavity, Phys. Rev. Lett. **89**, 200402 (2002).

[81] A. I. Lvovsky, P. Grangier, A. Ourjoumtsev, V. Parigi, M. Sasaki, and R. Tualle-Brouri, Production and applications of non-gaussian quantum states of light, arXiv:2006.16985 [quant-ph] (2020).

[82] F. Flamini, N. Spagnolo, and F. Sciarrino, Photonic quantum information processing: A review, Rep. Prog. Phys. **82**, 016001 (2018).

[83] R. Simon, N. Mukunda, and B. Dutta, Quantum-noise matrix for multimode systems: U(n) invariance, squeezing, and normal forms, Phys. Rev. A **49**, 1567 (1994).

[84] H. P. Robertson, The uncertainty principle, Phys. Rev. **34**, 163 (1929).

[85] Arvind, B. Dutta, N. Mukunda, and R. Simon, The real symplectic groups in quantum mechanics and optics, Pramana **45**, 471 (1995).

[86] J. Eisert and M. M. Wolf, in *Quantum Information with Continuous Variables of Atoms and Light* (World Scientific, 2007), p. 23.

[87] B. Demoen, P. Vanheuverzwijn, and A. Verbeure, Completely positive maps on the CCR-algebra, Lett Math Phys **2**, 161 (1977).

[88] B. Demoen, P. Vanheuverzwijn, and A. Verbeure, Completely positive quasi-free maps of the CCR-algebra, Rep. Math. Phys. **15**, 27 (1979).

[89] R. Filip and L. Mišta, Detecting Quantum States with a Positive Wigner Function beyond Mixtures of Gaussian States, Phys. Rev. Lett. **106**, 200401 (2011).

[90] A. Kenfack and K. Życzkowski, Negativity of the wigner function as an indicator of non-classicality, J. Opt. B: Quantum Semiclassical Opt. **6**, 396 (2004).

[91] J. L. W. V. Jensen, Sur les fonctions convexes et les inégalités entre les valeurs moyennes, Acta Math. **30**, 175 (1906).

[92] G. Vidal and R. F. Werner, Computable measure of entanglement, Phys. Rev. A **65**, 032314 (2002).

[93] M. B. Hastings, Superadditivity of communication capacity using entangled inputs, Nat. Phys. **5**, 255 (2009).

[94] C. D. Cushen and R. L. Hudson, A quantum-mechanical central limit theorem, J. Appl. Probab. **8**, 454 (1971).

[95] J. Quaegebeur, A noncommutative central limit theorem for ccr-algebras, J. Funct. Anal. **57**, 1 (1984).

[96] D. Goderis, A. Verbeure, and P. Vets, Non-commutative central limits, Probab. Theory Relat. Fields **82**, 527 (1989).

[97] M. G. Genoni, M. G. A. Paris, and K. Banaszek, Measure of the non-gaussian character of a quantum state, Phys. Rev. A **76**, 042327 (2007).

[98] M. G. Genoni, M. G. A. Paris, and K. Banaszek, Quantifying the non-gaussian character of a quantum state by quantum relative entropy, Phys. Rev. A **78**, 060303 (2008).

[99] M. G. Genoni and M. G. A. Paris, Quantifying non-gaussianity for quantum information, Phys. Rev. A **82**, 052341 (2010).

[100] F. Albarelli, M. G. Genoni, M. G. A. Paris, and A. Ferraro, Resource theory of quantum non-gaussianity and wigner negativity, Phys. Rev. A **98**, 052350 (2018).

[101] P. Marian and T. A. Marian, Relative entropy is an exact measure of non-gaussianity, Phys. Rev. A **88**, 012322 (2013).

[102] J. S. Ivan, M. S. Kumar, and R. Simon, A measure of non-gaussianity for quantum states, Quantum Inf. Process. **11**, 853 (2012).

[103] N. Lütkenhaus and S. M. Barnett, Nonclassical effects in phase space, Phys. Rev. A **51**, 3340 (1995).

[104] A function f is said to have an order of growth r when there are constant $a, b \in \mathbb{C}$ such that $|f(z)| \leqslant a\,\mathrm{e}\,xp(b|z|^r)$ for all $z \in \mathbb{C}$.

[105] B. Yurke and D. Stoler, Generating Quantum Mechanical Superpositions of Macroscopically Distinguishable States via Amplitude Dispersion, Phys. Rev. Lett. **57**, 13 (1986).

[106] E. Schrödinger, Die gegenwärtige situation in der quantenmechanik, Naturwissenschaften **23**, 807 (1935).

[107] P. T. Cochrane, G. J. Milburn, and W. J. Munro, Macroscopically distinct quantum-superposition states as a bosonic code for amplitude damping, Phys. Rev. A **59**, 2631 (1999).

[108] Z. Leghtas, G. Kirchmair, B. Vlastakis, R. J. Schoelkopf, M. H. Devoret, and M. Mirrahimi, Hardware-Efficient Autonomous Quantum Memory Protection, Phys. Rev. Lett. **111**, 120501 (2013).

[109] B. C. Sanders, Quantum dynamics of the nonlinear rotator and the effects of continual spin measurement, Phys. Rev. A **40**, 2417 (1989).

[110] A. N. Boto, P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, and J. P. Dowling, Quantum Interferometric Optical Lithography: Exploiting Entanglement to Beat the Diffraction Limit, Phys. Rev. Lett. **85**, 2733 (2000).

[111] H. Lee, P. Kok, and J. P. Dowling, A quantum rosetta stone for interferometry, J. Mod. Opt. **49**, 2325 (2002).

[112] B. Yurke, Input States for Enhancement of Fermion Interferometer Sensitivity, Phys. Rev. Lett. **56**, 1515 (1986).

[113] Y. Israel, I. Afek, S. Rosen, O. Ambar, and Y. Silberberg, Experimental tomography of noon states with large photon numbers, Phys. Rev. A **85**, 022115 (2012).

[114] M. G. Genoni, M. L. Palma, T. Tufarelli, S. Olivares, M. S. Kim, and M. G. A. Paris, Detecting quantum non-Gaussianity via the Wigner function, Phys. Rev. A **87**, 062104 (2013).

[115] C. Hughes, M. G. Genoni, T. Tufarelli, M. G. A. Paris, and M. S. Kim, Quantum non-Gaussianity witnesses in phase space, Phys. Rev. A **90**, 013810 (2014).

[116] J. Park, J. Zhang, J. Lee, S.-W. Ji, M. Um, D. Lv, K. Kim, and H. Nha, Testing Nonclassicality and Non-Gaussianity in Phase Space, Phys. Rev. Lett. **114**, 190402 (2015).

[117] L. Happ, M. A. Efremov, H. Nha, and W. P. Schleich, Sufficient condition for a quantum state to be genuinely quantum non-Gaussian, New J. Phys. **20**, 023046 (2018).

[118] B. Kühn and W. Vogel, Quantum non-Gaussianity and quantification of nonclassicality, Phys. Rev. A **97**, 053823 (2018).

[119] M. Bohmann, E. Agudelo, and J. Sperling, Probing nonclassicality with matrices of phase-space distributions, Quantum **4**, 343 (2020).

[120] I. Straka, A. Predojević, T. Huber, L. c. v. Lachman, L. Butschek, M. Miková, M. Mičuda, G. S. Solomon, G. Weihs, M. Ježek, and R. Filip, Quantum Non-Gaussian Depth of Single-Photon States, Phys. Rev. Lett. **113**, 223603 (2014).

[121] I. Straka, L. Lachman, J. Hloušek, M. Miková, M. Mičuda, M. Ježek, and R. Filip, Quantum non-Gaussian multiphoton light, Npj Quantum Inf. **4**, 4 (2018).

[122] C. N. Gagatsos and S. Guha, Efficient representation of Gaussian states for multimode non-Gaussian quantum state engineering via subtraction of arbitrary number of photons, Phys. Rev. A **99**, 053816 (2019).

[123] U. Chabaud, G. Roeland, M. Walschaers, F. Grosshans, V. Parigi, D. Markham, and N. Treps, Certification of non-Gaussian states with operational measurements, PRX Quantum **2**, 020333 (2021).

[124] P. Lelong and L. Gruman, *Entire Functions of Several Complex Variables* (Springer-Verlag Berlin Heidelberg, 1986).

[125] U. Chabaud, G. Ferrini, F. Grosshans, and D. Markham, Classical simulation of Gaussian quantum circuits with non-Gaussian input states, Phys. Rev. Res. **3**, 033018 (2021).

[126] R. Takagi and Q. Zhuang, Convex resource theory of non-Gaussianity, Phys. Rev. A **97**, 062337 (2018).

[127] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, The resource theory of stabilizer quantum computation, New J. Phys. **16**, 013009 (2014).

[128] U. Chabaud, P.-E. Emeriau, and F. Grosshans, Witnessing Wigner negativity, Quantum **5**, 471 (2021).

[129] A. Ourjoumtsev, R. Tualle-Brouri, J. Laurat, and P. Grangier, Generating optical Schrödinger kittens for quantum information processing, Science **312**, 83 (2006).

[130] A. Mari, K. Kieling, B. M. Nielsen, E. S. Polzik, and J. Eisert, Directly Estimating Nonclassicality, Phys. Rev. Lett. **106**, 010403 (2011).
[131] R. Loudon and P. Knight, Squeezed light, J. Mod. Opt. **34**, 709 (1987).
[132] L.-A. Wu, H. J. Kimble, J. L. Hall, and H. Wu, Generation of Squeezed States by Parametric Down Conversion, Phys. Rev. Lett. **57**, 2520 (1986).
[133] D. P. DiVincenzo, Quantum computation, Science **270**, 255 (1995).
[134] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Experimental Realization of any Discrete Unitary Operator, Phys. Rev. Lett. **73**, 58 (1994).
[135] P. Marek, R. Filip, and A. Furusawa, Deterministic implementation of weak quantum cubic nonlinearity, Phys. Rev. A **84**, 053802 (2011).
[136] K. Marshall, R. Pooser, G. Siopsis, and C. Weedbrook, Repeat-until-success cubic phase gate for universal continuous-variable quantum computation, Phys. Rev. A **91**, 032321 (2015).
[137] K. Miyata, H. Ogawa, P. Marek, R. Filip, H. Yonezawa, J.-i. Yoshikawa, and A. Furusawa, Implementation of a quantum cubic gate by an adaptive non-Gaussian measurement, Phys. Rev. A **93**, 022301 (2016).
[138] R. Yanagimoto, T. Onodera, E. Ng, L. G. Wright, P. L. McMahon, and H. Mabuchi, Engineering a Kerr-Based Deterministic Cubic Phase Gate via Gaussian Operations, Phys. Rev. Lett. **124**, 240503 (2020).
[139] Y. Zheng, O. Hahn, P. Stadler, P. Holmvall, F. Quijandría, A. Ferraro, and G. Ferrini, Gaussian conversion protocols for cubic phase state generation, PRX Quantum **2**, 010327 (2021).
[140] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, and M. H. Devoret, Manipulating the quantum state of an electrical circuit, Science **296**, 886 (2002).
[141] M. H. Devoret and R. J. Schoelkopf, Superconducting circuits for quantum information: An outlook, Science **339**, 1169 (2013).
[142] C. W. S. Chang, C. Sabín, P. Forn-Díaz, F. Quijandría, A. M. Vadiraj, I. Nsanzineza, G. Johansson, and C. M. Wilson, Observation of Three-Photon Spontaneous Parametric Down-Conversion in a Superconducting Parametric Cavity, Phys. Rev. X **10**, 011011 (2020).
[143] K. Kraus, General state changes in quantum theory, Ann. Phys. (N. Y) **64**, 311 (1971).
[144] Statistics of quantum measurements, in *Statistical Structure of Quantum Theory* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2001) p. 39.
[145] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, New York, NY, 2017), 2nd ed. corrected reprint ed.
[146] L. Lami, A. Serafini, and G. Adesso, Gaussian entanglement revisited, New J. Phys. **20**, 023030 (2018).
[147] D. E. Browne, J. Eisert, S. Scheel, and M. B. Plenio, Driving non-Gaussian to Gaussian states with linear optics, Phys. Rev. A **67**, 062320 (2003).
[148] J. Wenger, R. Tualle-Brouri, and P. Grangier, Non-Gaussian Statistics from Individual Pulses of Squeezed Light, Phys. Rev. Lett. **92**, 153601 (2004).
[149] 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 J. Phys. **10**, 123006 (2008).
[150] V. Averchenko, C. Jacquard, V. Thiel, C. Fabre, and N. Treps, Multimode theory of single-photon subtraction, New J. Phys. **18**, 083042 (2016).
[151] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2009).
[152] M. Walschaers, C. Fabre, V. Parigi, and N. Treps, Entanglement and Wigner Function Negativity of Multimode Non-Gaussian States, Phys. Rev. Lett. **119**, 183601 (2017).
[153] M. Walschaers, C. Fabre, V. Parigi, and N. Treps, Statistical signatures of multimode single-photon-added and -subtracted states of light, Phys. Rev. A **96**, 053835 (2017).
[154] D. Braun, P. Jian, O. Pinel, and N. Treps, Precision measurements with photon-subtracted or photon-added gaussian states, Phys. Rev. A **90**, 013821 (2014).
[155] K. Mølmer, Non-Gaussian states from continuous-wave Gaussian light sources, Phys. Rev. A **73**, 063804 (2006).
[156] M. Dakna, T. Anhut, T. Opatrný, L. Knöll, and D.-G. Welsch, Generating Schrödinger-cat-like states by means of conditional measurements on a beam splitter, Phys. Rev. A **55**, 3184 (1997).
[157] K. Takase, J.-i. Yoshikawa, W. Asavanant, M. Endo, and A. Furusawa, Generation of optical Schrödinger cat states by generalized photon subtraction, Phys. Rev. A **103**, 013710 (2021).
[158] A. O. C. Davis, M. Walschaers, V. Parigi, and N. Treps, Conditional preparation of non-Gaussian quantum optical states by mesoscopic measurement, New J. Phys. **23**, 063039 (2021).
[159] A. Zavatta, S. Viciani, and M. Bellini, Quantum-to-classical transition with single-photon-added coherent states of light, Science **306**, 660 (2004).
[160] V. Parigi, A. Zavatta, M. Kim, and M. Bellini, Probing quantum commutation rules by addition and subtraction of single photons to/from a light field, Science **317**, 1890 (2007).
[161] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).
[162] R. Uola, A. C. S. Costa, H. C. Nguyen, and O. Gühne, Quantum steering, Rev. Mod. Phys. **92**, 015001 (2020).
[163] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).
[164] One could introduce more general and rigorous notation based on measure theory, but here we restrict to a simpler, though less general, formulation for pedagogical purposes.
[165] R. Jozsa, Quantum effects in algorithms, Chaos, Solitons Fractals **10**, 1657 (1999).
[166] This is true for general "states" in classical probability theory, but in the CV context of this Tutorial it is appealing to

phrase the argument in terms of probability distributions of phase space.

[167] R. F. Werner, Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).

[168] J. B. Conway, in *A Course in Functional Analysis* (Springer New York, New York, NY, 2007), p. 63.

[169] Sohail and U. Sen, Witnessing nonseparability of bipartite quantum operations, Phys. Lett. A **404**, 127411 (2021).

[170] H. F. Hofmann and S. Takeuchi, Violation of local uncertainty relations as a signature of entanglement, Phys. Rev. A **68**, 032103 (2003).

[171] O. Gittsovich, O. Gühne, P. Hyllus, and J. Eisert, Unifying several separability conditions using the covariance matrix criterion, Phys. Rev. A **78**, 052319 (2008).

[172] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev. **47**, 777 (1935).

[173] E. Schrödinger, Discussion of probability relations between separated systems, Math. Proc. Cambridge Philosophical Soc. **31**, 555 (1935).

[174] E. Schrödinger, Probability relations between separated systems, Math. Proc. Cambridge Philosophical Soc. **32**, 446 (1936).

[175] M. D. Reid, Demonstration of the einstein-podolsky-rosen paradox using nondegenerate parametric amplification, Phys. Rev. A **40**, 913 (1989).

[176] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Steering, Entanglement, Nonlocality, and the Einstein-Podolsky-Rosen Paradox, Phys. Rev. Lett. **98**, 140402 (2007).

[177] D. Cavalcanti and P. Skrzypczyk, Quantum steering: A review with focus on semidefinite programming, Rep. Prog. Phys. **80**, 024001 (2017).

[178] E. G. Cavalcanti, S. J. Jones, H. M. Wiseman, and M. D. Reid, Experimental criteria for steering and the einstein-podolsky-rosen paradox, Phys. Rev. A **80**, 032112 (2009).

[179] I. Kogias, A. R. Lee, S. Ragy, and G. Adesso, Quantification of Gaussian Quantum Steering, Phys. Rev. Lett. **114**, 060403 (2015).

[180] Q. Y. He, Q. H. Gong, and M. D. Reid, Classifying Directional Gaussian Entanglement, Einstein-Podolsky-Rosen Steering, and Discord, Phys. Rev. Lett. **114**, 060402 (2015).

[181] X. Deng, Y. Xiang, C. Tian, G. Adesso, Q. He, Q. Gong, X. Su, C. Xie, and K. Peng, Demonstration of Monogamy Relations for Einstein-Podolsky-Rosen Steering in Gaussian Cluster States, Phys. Rev. Lett. **118**, 230501 (2017).

[182] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering, Phys. Rev. A **85**, 010301 (2012).

[183] J. S. Bell, On the einstein podolsky rosen paradox, Physics Physique Fizika **1**, 195 (1964).

[184] J. S. BELL, On the problem of hidden variables in quantum mechanics, Rev. Mod. Phys. **38**, 447 (1966).

[185] N. D. Mermin, Hidden variables and the two theorems of john bell, Rev. Mod. Phys. **65**, 803 (1993).

[186] A. Peres, All the bell inequalities, Found. Phys. **29**, 589 (1999).

[187] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[188] J. F. Clauser and M. A. Horne, Experimental consequences of objective local theories, Phys. Rev. D **10**, 526 (1974).

[189] S. J. Freedman and J. F. Clauser, Experimental Test of Local Hidden-Variable Theories, Phys. Rev. Lett. **28**, 938 (1972).

[190] A. Aspect, P. Grangier, and G. Roger, Experimental Tests of Realistic Local Theories via Bell's Theorem, Phys. Rev. Lett. **47**, 460 (1981).

[191] A. Aspect, J. Dalibard, and G. Roger, Experimental Test of Bell's Inequalities Using Time-Varying Analyzers, Phys. Rev. Lett. **49**, 1804 (1982).

[192] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, Violation of Bell's Inequality under Strict Einstein Locality Conditions, Phys. Rev. Lett. **81**, 5039 (1998).

[193] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres, Nature **526**, 682 (2015).

[194] M. Giustina *et al.*, Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons, Phys. Rev. Lett. **115**, 250401 (2015).

[195] L. K. Shalm *et al.*, Strong Loophole-Free Test of Local Realism, Phys. Rev. Lett. **115**, 250402 (2015).

[196] C. Abellán *et al.*, T. B. B. T. Collaboration, Challenging local realism with human choices, Nature **557**, 212 (2018).

[197] D. Rauch, J. Handsteiner, A. Hochrainer, J. Gallicchio, A. S. Friedman, C. Leung, B. Liu, L. Bulla, S. Ecker, F. Steinlechner, R. Ursin, B. Hu, D. Leon, C. Benn, A. Ghedina, M. Cecconi, A. H. Guth, D. I. Kaiser, T. Scheidl, and A. Zeilinger, Cosmic Bell Test Using Random Measurement Settings from High-Redshift Quasars, Phys. Rev. Lett. **121**, 080403 (2018).

[198] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-Independent Security of Quantum Cryptography against Collective Attacks, Phys. Rev. Lett. **98**, 230501 (2007).

[199] F. Buscemi, All Entangled Quantum States are Nonlocal, Phys. Rev. Lett. **108**, 200401 (2012).

[200] C. Branciard, D. Rosset, Y.-C. Liang, and N. Gisin, Measurement-Device-Independent Entanglement Witnesses for all Entangled Quantum States, Phys. Rev. Lett. **110**, 060405 (2013).

[201] P. Abiuso, S. Bäuml, D. Cavalcanti, and A. Acín, Measurement-Device-Independent Entanglement Detection for Continuous-Variable Systems, Phys. Rev. Lett. **126**, 190502 (2021).

[202] C. K. Hong and L. Mandel, Experimental Realization of a Localized One-Photon State, Phys. Rev. Lett. **56**, 58 (1986).

[203] G. S. Thekkadath, B. A. Bell, I. A. Walmsley, and A. I., Engineering Schrödinger cat states with a photonic even-parity detector, Quantum **4**, 239 (2020).

[204] M. Eaton, R. Nehra, and O. Pfister, Non-gaussian and gottesman–kitaev–preskill state preparation by photon catalysis, New J. Phys. **21**, 113034 (2019).
[205] 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).
[206] S. Yu, Q. Chen, C. Zhang, C. H. Lai, and C. H. Oh, All Entangled Pure States Violate a Single Bell's Inequality, Phys. Rev. Lett. **109**, 120402 (2012).
[207] R. F. Werner and M. M. Wolf, Bound Entangled Gaussian States, Phys. Rev. Lett. **86**, 3658 (2001).
[208] Y. Xiang, S. Liu, J. Guo, Q. Gong, N. Treps, Q. He, and M. Walschaers, Quantification of wigner negativity remotely generated via einstein-podolsky-rosen steering, arXiv:2104.00451 [quant-ph] (2021).
[209] M. D. Reid, Monogamy inequalities for the einstein-podolsky-rosen paradox and quantum steering, Phys. Rev. A **88**, 062108 (2013).
[210] S.-W. Ji, M. S. Kim, and H. Nha, Quantum steering of multimode gaussian states by gaussian measurements: monogamy relations and the peres conjecture, J. Phys. A: Math. Theor. **48**, 135301 (2015).
[211] X. Deng, Y. Xiang, C. Tian, G. Adesso, Q. He, Q. Gong, X. Su, C. Xie, and K. Peng, Demonstration of Monogamy Relations for Einstein-Podolsky-Rosen Steering in Gaussian Cluster States, Phys. Rev. Lett. **118**, 230501 (2017).
[212] E. Shchukin and W. Vogel, Inseparability Criteria for Continuous Bipartite Quantum States, Phys. Rev. Lett. **95**, 230502 (2005).
[213] J. Sperling and W. Vogel, Verifying continuous-variable entanglement in finite spaces, Phys. Rev. A **79**, 052313 (2009).
[214] J. S. Kim and B. C. Sanders, Monogamy of multi-qubit entanglement using rényi entropy, J. Phys. A: Math. Theor. **43**, 445305 (2010).
[215] G. Adesso, D. Girolami, and A. Serafini, Measuring Gaussian Quantum Information and Correlations Using the rényi Entropy of Order 2, Phys. Rev. Lett. **109**, 190502 (2012).
[216] R. Simon, Peres-Horodecki Separability Criterion for Continuous Variable Systems, Phys. Rev. Lett. **84**, 2726 (2000).
[217] L.-M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, Inseparability Criterion for Continuous Variable Systems, Phys. Rev. Lett. **84**, 2722 (2000).
[218] P. van Loock and A. Furusawa, Detecting genuine multipartite continuous-variable entanglement, Phys. Rev. A **67**, 052315 (2003).
[219] P. Hyllus and J. Eisert, Optimal entanglement witnesses for continuous-variable systems, New J. Phys. **8**, 51 (2006).
[220] S. Gerke, J. Sperling, W. Vogel, Y. Cai, J. Roslund, N. Treps, and C. Fabre, Full Multipartite Entanglement of Frequency-Comb Gaussian States, Phys. Rev. Lett. **114**, 050501 (2015).
[221] A. A. Valido, F. Levi, and F. Mintert, Hierarchies of multipartite entanglement for continuous-variable states, Phys. Rev. A **90**, 052321 (2014).
[222] Note that there is also a more profound connection between the transpose and time reversal through Wigner's theorem. The latter is here effectively implemented by changing the sign of the momentum variables. However, this connection is beyond the scope of this Tutorial and interested readers are invited to indulge in the literature on quantum chaos instead.
[223] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels, Phys. Rev. Lett. **76**, 722 (1996).
[224] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Concentrating partial entanglement by local operations, Phys. Rev. A **53**, 2046 (1996).
[225] J. Eisert, S. Scheel, and M. B. Plenio, Distilling Gaussian States with Gaussian Operations is Impossible, Phys. Rev. Lett. **89**, 137903 (2002).
[226] J. Fiurášek, Gaussian Transformations and Distillation of Entangled Gaussian States, Phys. Rev. Lett. **89**, 137904 (2002).
[227] G. Giedke and J. Ignacio Cirac, Characterization of gaussian operations and distillation of gaussian states, Phys. Rev. A **66**, 032316 (2002).
[228] R. Dong, M. Lassen, J. Heersink, C. Marquardt, R. Filip, G. Leuchs, and U. L. Andersen, Experimental entanglement distillation of mesoscopic quantum states, Nat. Phys. **4**, 919 (2008).
[229] B. Hage, A. Samblowski, J. DiGuglielmo, A. Franzen, J. Fiurášek, and R. Schnabel, Preparation of distilled and purified continuous-variable entangled states, Nat. Phys. **4**, 915 (2008).
[230] L.-M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, Entanglement Purification of Gaussian Continuous Variable Quantum States, Phys. Rev. Lett. **84**, 4002 (2000).
[231] L.-M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, Physical implementation for entanglement purification of gaussian continuous-variable quantum states, Phys. Rev. A **62**, 032304 (2000).
[232] J. Fiurášek, Distillation and purification of symmetric entangled gaussian states, Phys. Rev. A **82**, 042331 (2010).
[233] E. T. Campbell, M. G. Genoni, and J. Eisert, Continuous-variable entanglement distillation and noncommutative central limit theorems, Phys. Rev. A **87**, 042330 (2013).
[234] J. Dias and T. C. Ralph, Quantum repeaters using continuous-variable teleportation, Phys. Rev. A **95**, 022312 (2017).
[235] G. Y. Xiang, T. C. Ralph, A. P. Lund, N. Walk, and G. J. Pryde, Heralded noiseless linear amplification and distillation of entanglement, Nat. Photonics **4**, 316 (2010).
[236] T. Opatrný, G. Kurizki, and D.-G. Welsch, Improvement on teleportation of continuous variables by photon subtraction via conditional measurement, Phys. Rev. A **61**, 032302 (2000).
[237] P. T. Cochrane, T. C. Ralph, and G. J. Milburn, Teleportation improvement by conditional measurements on the two-mode squeezed vacuum, Phys. Rev. A **65**, 062306 (2002).

[238] S. Olivares, M. G. A. Paris, and R. Bonifacio, Teleportation improvement by inconclusive photon subtraction, Phys. Rev. A **67**, 032314 (2003).
[239] A. Kitagawa, M. Takeoka, M. Sasaki, and A. Chefles, Entanglement evaluation of non-gaussian states generated by photon subtraction from squeezed states, Phys. Rev. A **73**, 042310 (2006).
[240] Y. Yang and F.-L. Li, Entanglement properties of non-gaussian resources generated via photon subtraction and addition and continuous-variable quantum-teleportation improvement, Phys. Rev. A **80**, 022315 (2009).
[241] C. Navarrete-Benlloch, R. García-Patrón, J. H. Shapiro, and N. J. Cerf, Enhancing quantum entanglement by photon addition and subtraction, Phys. Rev. A **86**, 012328 (2012).
[242] T. Das, R. Prabhu, A. Sen(De), and U. Sen, Superiority of photon subtraction to addition for entanglement in a multimode squeezed vacuum, Phys. Rev. A **93**, 052313 (2016).
[243] M. Walschaers, S. Sarkar, V. Parigi, and N. Treps, Tailoring Non-Gaussian Continuous-Variable Graph States, Phys. Rev. Lett. **121**, 220501 (2018).
[244] K. Zhang, J. Jing, N. Treps, and M. Walschaers, Maximal entanglement distillation with single-photon subtraction, arXiv:2103.09197 [quant-ph] (2021).
[245] E. G. Brown, N. Friis, and M. Huber, Passivity and practical work extraction using gaussian operations, New J. Phys. **18**, 113028 (2016).
[246] R. S. Barbosa, T. Douce, P.-E. Emeriau, E. Kashefi, and S. Mansfield, Continuous-variable nonlocality and contextuality, arXiv:1905.08267 [quant-ph] (2019).
[247] A. Acín, N. J. Cerf, A. Ferraro, and J. Niset, Tests of multimode quantum nonlocality with homodyne measurements, Phys. Rev. A **79**, 012112 (2009).
[248] R. García-Patrón, J. Fiurášek, N. J. Cerf, J. Wenger, R. Tualle-Brouri, and P. Grangier, Proposal for a Loophole-Free Bell Test Using Homodyne Detection, Phys. Rev. Lett. **93**, 130409 (2004).
[249] D. Klyshko, The bell and ghz theorems: A possible three-photon interference experiment and the question of nonlocality, Phys. Lett. A **172**, 399 (1993).
[250] W. N. Plick, F. Arzani, N. Treps, E. Diamanti, and D. Markham, Violating bell inequalities with entangled optical frequency combs and multipixel homodyne detection, Phys. Rev. A **98**, 062101 (2018).
[251] K. Banaszek and K. Wódkiewicz, Nonlocality of the einstein-podolsky-rosen state in the wigner representation, Phys. Rev. A **58**, 4345 (1998).
[252] K. Banaszek and K. Wódkiewicz, Testing Quantum Nonlocality in Phase Space, Phys. Rev. Lett. **82**, 2009 (1999).
[253] S. Aaronson and D. Gottesman, Improved simulation of stabilizer circuits, Phys. Rev. A **70**, 052328 (2004).
[254] V. Veitch, N. Wiebe, C. Ferrie, and J. Emerson, Efficient simulation scheme for a class of quantum optics experiments with non-negative wigner representation, New J. Phys. **15**, 013037 (2013).
[255] S. Rahimi-Keshari, T. C. Ralph, and C. M. Caves, Sufficient Conditions for Efficient Classical Simulation of Quantum Optics, Phys. Rev. X **6**, 021039 (2016).
[256] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Appl. **10**, 285 (1975).
[257] A. S. Holevo, The choi–jamiolkowski forms of quantum gaussian channels, J. Math. Phys. **52**, 042202 (2011).
[258] L. García-Álvarez, C. Calcluth, A. Ferraro, and G. Ferrini, Efficient simulatability of continuous-variable circuits with large wigner negativity, Phys. Rev. Res. **2**, 043322 (2020).
[259] C. S. Hamilton, R. Kruse, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex, Gaussian Boson Sampling, Phys. Rev. Lett. **119**, 170501 (2017).
[260] 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).
[261] A. Deshpande, A. Mehta, T. Vincent, N. Quesada, M. Hinsche, M. Ioannou, L. Madsen, J. Lavoie, H. Qi, J. Eisert, D. Hangleiter, B. Fefferman, and I. Dhand, Quantum computational supremacy via high-dimensional gaussian boson sampling, arXiv:2102.12474 [quant-ph] (2021).
[262] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, Theory of Computing **9**, 143 (2013).
[263] N. Quesada, J. M. Arrazola, and N. Killoran, Gaussian boson sampling using threshold detectors, Phys. Rev. A **98**, 062322 (2018).
[264] E. R. Caianiello, On quantum field theory —i: Explicit solution of dyson's equation in electrodynamics without use of feynman graphs, Il Nuovo Cimento (1943-1954) **10**, 1634 (1953).
[265] K. Brádler, P.-L. Dallaire-Demers, P. Rebentrost, D. Su, and C. Weedbrook, Gaussian boson sampling for perfect matchings of arbitrary graphs, Phys. Rev. A **98**, 032310 (2018).
[266] M. Schuld, K. Brádler, R. Israel, D. Su, and B. Gupt, Measuring the similarity of graphs with a gaussian boson sampler, Phys. Rev. A **101**, 032314 (2020).
[267] T. R. Bromley, J. M. Arrazola, S. Jahangiri, J. Izaac, N. Quesada, A. D. Gran, M. Schuld, J. Swinarton, Z. Zabaneh, and N. Killoran, Applications of near-term photonic quantum computers: Software and algorithms, Quantum Sci. Technol. **5**, 034010 (2020).
[268] H. Qi, D. J. Brod, N. Quesada, and R. García-Patrón, Regimes of Classical Simulability for Noisy Gaussian Boson Sampling, Phys. Rev. Lett. **124**, 100502 (2020).
[269] J. F. Clauser, Experimental distinction between the quantum and classical field-theoretic predictions for the photoelectric effect, Phys. Rev. D **9**, 853 (1974).
[270] H. J. Kimble, M. Dagenais, and L. Mandel, Photon Antibunching in Resonance Fluorescence, Phys. Rev. Lett. **39**, 691 (1977).
[271] F. Diedrich and H. Walther, Nonclassical Radiation of a Single Stored ion, Phys. Rev. Lett. **58**, 203 (1987).
[272] R. Schnabel, Squeezed states of light and their applications in laser interferometers, Phys. Rep. **684**, 1 (2017).
[273] A. I. Lvovsky, H. Hansen, T. Aichele, O. Benson, J. Mlynek, and S. Schiller, Quantum State Reconstruction of the Single-Photon Fock State, Phys. Rev. Lett. **87**, 050402 (2001).

[274] A. I. Lvovsky and M. G. Raymer, Continuous-variable optical quantum-state tomography, Rev. Mod. Phys. **81**, 299 (2009).
[275] A. I. Lvovsky and S. A. Babichev, Synthesis and tomographic characterization of the displaced fock state of light, Phys. Rev. A **66**, 011801 (2002).
[276] A. Ourjoumtsev, R. Tualle-Brouri, and P. Grangier, Quantum Homodyne Tomography of a Two-Photon Fock State, Phys. Rev. Lett. **96**, 213601 (2006).
[277] M. Cooper, L. J. Wright, C. Söller, and B. J. Smith, Experimental generation of multi-photon fock states, Opt. Express **21**, 5309 (2013).
[278] M. Yukawa, K. Miyata, T. Mizuta, H. Yonezawa, P. Marek, R. Filip, and A. Furusawa, Generating superposition of up-to three photons for continuous variable quantum information processing, Opt. Express **21**, 5529 (2013).
[279] J. S. Neergaard-Nielsen, B. M. Nielsen, C. Hettich, K. Mølmer, and E. S. Polzik, Generation of a Superposition of odd Photon Number States for Quantum Information Networks, Phys. Rev. Lett. **97**, 083604 (2006).
[280] K. Wakui, H. Takahashi, A. Furusawa, and M. Sasaki, Photon subtracted squeezed states generated with periodically poled ktiopo4, Opt. Express **15**, 3568 (2007).
[281] D. V. Sychev, A. E. Ulanov, A. A. Pushkina, M. W. Richards, I. A. Fedorov, and A. I. Lvovsky, Enlargement of optical schrödinger's cat states, Nat. Photonics **11**, 379 (2017).
[282] J. S. Neergaard-Nielsen, M. Takeuchi, K. Wakui, H. Takahashi, K. Hayasaka, M. Takeoka, and M. Sasaki, Optical Continuous-Variable Qubit, Phys. Rev. Lett. **105**, 053602 (2010).
[283] M. Barbieri, N. Spagnolo, M. G. Genoni, F. Ferreyrol, R. Blandino, M. G. A. Paris, P. Grangier, and R. Tualle-Brouri, Non-gaussianity of quantum states: An experimental test on single-photon-added coherent states, Phys. Rev. A **82**, 063833 (2010).
[284] A. Ourjoumtsev, H. Jeong, R. Tualle-Brouri, and P. Grangier, Generation of optical 'schrödinger cats'from photon number states, Nature **448**, 784 (2007).
[285] J. Etesse, M. Bouillard, B. Kanseri, and R. Tualle-Brouri, Experimental Generation of Squeezed cat States with an Operation Allowing Iterative Growth, Phys. Rev. Lett. **114**, 193602 (2015).
[286] B. Hacker, S. Welte, S. Daiss, A. Shaukat, S. Ritter, L. Li, and G. Rempe, Deterministic creation of entangled atom–light schrödinger-cat states, Nat. Photonics **13**, 110 (2019).
[287] S. A. Babichev, J. Appel, and A. I. Lvovsky, Homodyne Tomography Characterization and Nonlocality of a Dual-Mode Optical Qubit, Phys. Rev. Lett. **92**, 193601 (2004).
[288] M. D'Angelo, A. Zavatta, V. Parigi, and M. Bellini, Tomographic test of bell's inequality for a time-delocalized single photon, Phys. Rev. A **74**, 052114 (2006).
[289] H. Takahashi, J. S. Neergaard-Nielsen, M. Takeuchi, M. Takeoka, K. Hayasaka, A. Furusawa, and M. Sasaki, Entanglement distillation from Gaussian input states, Nat. Photon **4**, 178 (2010).
[290] A. Ourjoumtsev, A. Dantan, R. Tualle-Brouri, and P. Grangier, Increasing Entanglement between Gaussian States by Coherent Photon Subtraction, Phys. Rev. Lett. **98**, 030502 (2007).
[291] A. Ourjoumtsev, F. Ferreyrol, R. Tualle-Brouri, and P. Grangier, Preparation of non-local superpositions of quasi-classical light states, Nat. Phys. **5**, 189 (2009).
[292] H. Jeong, A. Zavatta, M. Kang, S.-W. Lee, L. S. Costanzo, S. Grandi, T. C. Ralph, and M. Bellini, Generation of hybrid entanglement of light, Nat. Photonics **8**, 564 (2014).
[293] O. Morin, K. Huang, J. Liu, H. Le Jeannic, C. Fabre, and J. Laurat, Remote creation of hybrid entanglement between particle-like and wave-like optical qubits, Nat. Photonics **8**, 570 (2014).
[294] N. Biagi, L. S. Costanzo, M. Bellini, and A. Zavatta, Entangling Macroscopic Light States by Delocalized Photon Addition, Phys. Rev. Lett. **124**, 033604 (2020).
[295] T. Serikawa, J.-i. Yoshikawa, S. Takeda, H. Yonezawa, T. C. Ralph, E. H. Huntington, and A. Furusawa, Generation of a cat State in an Optical Sideband, Phys. Rev. Lett. **121**, 143602 (2018).
[296] V. A. Averchenko, V. Thiel, and N. Treps, Nonlinear photon subtraction from a multimode quantum field, Phys. Rev. A **89**, 063808 (2014).
[297] Y.-S. Ra, C. Jacquard, A. Dufour, C. Fabre, and N. Treps, Tomography of a Mode-Tunable Coherent Single-Photon Subtractor, Phys. Rev. X **7**, 031012 (2017).
[298] Y.-S. Ra, A. Dufour, M. Walschaers, C. Jacquard, T. Michel, C. Fabre, and N. Treps, Non-gaussian quantum states of a multimode light field, Nat. Phys. **16**, 144 (2020).
[299] V. Cimini, M. Barbieri, N. Treps, M. Walschaers, and V. Parigi, Neural Networks for Detecting Multimode Wigner Negativity, Phys. Rev. Lett. **125**, 160504 (2020).
[300] U. Chabaud, F. Grosshans, E. Kashefi, and D. Markham, Efficient verification of boson sampling, arXiv:2006.03520 [quant-ph] (2021).
[301] E. S. Tiunov, V. V. T. (Vyborova), A. E. Ulanov, A. I. Lvovsky, and A. K. Fedorov, Experimental quantum homodyne tomography via machine learning, Optica **7**, 448 (2020).
[302] A. E. Lita, A. J. Miller, and S. W. Nam, Counting near-infrared single-photons with 95% efficiency, Opt. Express **16**, 3032 (2008).
[303] C. Monroe, D. M. Meekhof, B. E. King, and D. J. Wineland, A "schrödinger cat" superposition state of an atom, Science **272**, 1131 (1996).
[304] M. Brune, S. Haroche, J. M. Raimond, L. Davidovich, and N. Zagury, Manipulation of photons in a cavity by dispersive atom-field coupling: Quantum-nondemolition measurements and generation of "Schrödinger cat" states, Phys. Rev. A **45**, 5193 (1992).
[305] M. Brune, F. Schmidt-Kaler, A. Maali, J. Dreyer, E. Hagley, J. M. Raimond, and S. Haroche, Quantum Rabi Oscillation: A Direct Test of Field Quantization in a Cavity, Phys. Rev. Lett. **76**, 1800 (1996).
[306] G. Nogues, A. Rauschenbeutel, S. Osnaghi, M. Brune, J. M. Raimond, and S. Haroche, Seeing a single photon without destroying it, Nature **400**, 239 (1999).

[307] S. Deléglise, I. Dotsenko, C. Sayrin, J. Bernu, M. Brune, J.-M. Raimond, and S. Haroche, Reconstruction of non-classical cavity field states with snapshots of their decoherence, Nature **455**, 510 (2008).

[308] M. Hofheinz, E. M. Weig, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O'Connell, H. Wang, J. M. Martinis, and A. N. Cleland, Generation of Fock states in a superconducting quantum circuit, Nature **454**, 310 (2008).

[309] B. Vlastakis, G. Kirchmair, Z. Leghtas, S. E. Nigg, L. Frunzio, S. M. Girvin, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, Deterministically encoding quantum information using 100-photon Schrödinger cat states, Science **342**, 607 (2013).

[310] Z. Leghtas, S. Touzard, I. M. Pop, A. Kou, B. Vlastakis, A. Petrenko, K. M. Sliwa, A. Narla, S. Shankar, M. J. Hatridge, M. Reagor, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret, Confining the state of light to a quantum manifold by engineered two-photon loss, Science **347**, 853 (2015).

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

[312] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, L. Frunzio, M. Mirrahimi, and M. H. Devoret, Quantum error correction of a qubit encoded in grid states of an oscillator, Nature **584**, 368 (2020).

[313] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, Cavity optomechanics, Rev. Mod. Phys. **86**, 1391 (2014).

[314] U. B. Hoff, J. Kollath-Bönig, J. S. Neergaard-Nielsen, and U. L. Andersen, Measurement-Induced Macroscopic Superposition States in Cavity Optomechanics, Phys. Rev. Lett. **117**, 143601 (2016).

[315] M. Brunelli, O. Houhou, D. W. Moore, A. Nunnenkamp, M. Paternostro, and A. Ferraro, Unconditional preparation of nonclassical states via linear-and-quadratic optomechanics, Phys. Rev. A **98**, 063801 (2018).

[316] T. A. Palomaki, J. D. Teufel, R. W. Simmonds, and K. W. Lehnert, Entangling mechanical motion with microwave fields, Science **342**, 710 (2013).

[317] R. Riedinger, S. Hong, R. A. Norte, J. A. Slater, J. Shang, A. G. Krause, V. Anant, M. Aspelmeyer, and S. Gröblacher, Non-classical correlations between single photons and phonons from a mechanical oscillator, Nature **530**, 313 (2016).

[318] A. P. Reed, K. H. Mayer, J. D. Teufel, L. D. Burkhart, W. Pfaff, M. Reagor, L. Sletten, X. Ma, R. J. Schoelkopf, E. Knill, and K. W. Lehnert, Faithful conversion of propagating quantum information to mechanical motion, Nat. Phys. **13**, 1163 (2017).

[319] J. D. Jost, J. P. Home, J. M. Amini, D. Hanneke, R. Ozeri, C. Langer, J. J. Bollinger, D. Leibfried, and D. J. Wineland, Entangled mechanical oscillators, Nature **459**, 683 (2009).

[320] E. Flurin, N. Roch, F. Mallet, M. H. Devoret, and B. Huard, Generating Entangled Microwave Radiation Over two Transmission Lines, Phys. Rev. Lett. **109**, 183901 (2012).

[321] M. Forsch, R. Stockill, A. Wallucks, I. Marinković, C. Gärtner, R. A. Norte, F. van Otten, A. Fiore, K. Srinivasan, and S. Gröblacher, Microwave-to-optics conversion using a mechanical oscillator in its quantum ground state, Nat. Phys. **16**, 69 (2020).

[322] R. Lescanne, S. Deléglise, E. Albertinale, U. Réglade, T. Capelle, E. Ivanov, T. Jacqmin, Z. Leghtas, and E. Flurin, Irreversible Qubit-Photon Coupling for the Detection of Itinerant Microwave Photons, Phys. Rev. X **10**, 021038 (2020).

[323] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Quantum metrology with nonclassical states of atomic ensembles, Rev. Mod. Phys. **90**, 035005 (2018).

[324] B. Dubost, M. Koschorreck, M. Napolitano, N. Behbood, R. J. Sewell, and M. W. Mitchell, Efficient Quantification of Non-Gaussian Spin Distributions, Phys. Rev. Lett. **108**, 183602 (2012).

[325] F. Haas, J. Volz, R. Gehr, J. Reichel, and J. Estève, Entangled states of more than 40 atoms in an optical fiber cavity, Science **344**, 180 (2014).

[326] H. Strobel, W. Muessel, D. Linnemann, T. Zibold, D. B. Hume, L. Pezzè, A. Smerzi, and M. K. Oberthaler, Fisher information and entanglement of non-Gaussian spin states, Science **345**, 424 (2014).

[327] R. McConnell, H. Zhang, J. Hu, S. Ćuk, and V. Vuletić, Entanglement with negative Wigner function of almost 3,000 atoms heralded by one photon, Nature **519**, 439 (2015).

[328] M. V. Larsen, X. Guo, C. R. Breum, J. S. Neergaard-Nielsen, and U. L. Andersen, Deterministic multi-mode gates on a scalable photonic quantum computing platform, arXiv:2010.14422 [quant-ph] (2020).

[329] J. M. Arrazola *et al.*, Quantum circuits with many photons on a programmable nanophotonic chip, Nature **591**, 54 (2021).

[330] P. Jouguet, S. Kunz-Jacques, A. Leverrier, P. Grangier, and E. Diamanti, Experimental demonstration of long-distance continuous-variable quantum key distribution, Nat. Photonics **7**, 378 (2013).

[331] B. A. Bell, D. Markham, D. A. Herrera-Martí, A. Marin, W. J. Wadsworth, J. G. Rarity, and M. S. Tame, Experimental demonstration of graph-state quantum secret sharing, Nat. Commun. **5**, 5480 (2014).

[332] M. Navascués, F. Grosshans, and A. Acín, Optimality of Gaussian Attacks in Continuous-Variable Quantum Cryptography, Phys. Rev. Lett. **97**, 190502 (2006).

[333] Y. Guo, Q. Liao, Y. Wang, D. Huang, P. Huang, and G. Zeng, Performance improvement of continuous-variable quantum key distribution with an entangled source in the middle via photon subtraction, Phys. Rev. A **95**, 032304 (2017).

[334] J. Huh, G. G. Guerreschi, B. Peropadre, J. R. McClean, and A. Aspuru-Guzik, Boson sampling for molecular vibronic spectra, Nat. Photonics **9**, 615 (2015).

[335] L. Banchi, M. Fingerhuth, T. Babej, C. Ing, and J. M. Arrazola, Molecular docking with Gaussian Boson sampling, Sci. Adv. **6** (2020).

[336] S. Jahangiri, J. M. Arrazola, and A. Delgado, Quantum algorithm for simulating single-molecule electron transport, *The Journal of Physical Chemistry Letters*, J. Phys. Chem. Lett. **12**, 1256 (2021).

[337] N. Killoran, T. R. Bromley, J. M. Arrazola, M. Schuld, N. Quesada, and S. Lloyd, Continuous-variable

quantum neural networks, Phys. Rev. Res. **1**, 033063 (2019).

[338] J. Nokkala, R. Martínez-Peña, G. L. Giorgi, V. Parigi, M. C. Soriano, and R. Zambrini, Gaussian states of continuous-variable quantum systems provide universal and versatile reservoir computing, Commun. Physics **4**, 53 (2021).

[339] F. Wolf, C. Shi, J. C. Heip, M. Gessner, L. Pezzè, A. Smerzi, M. Schulte, K. Hammerer, and P. O. Schmidt, Motional Fock states for quantum-enhanced amplitude and phase measurements with trapped ions, Nat. Commun. **10**, 2929 (2019).

[340] M. Gessner, L. Pezzè, and A. Smerzi, Entanglement and squeezing in continuous-variable systems, Quantum **1**, 17 (2017).

[341] B. Yadin, M. Fadel, and M. Gessner, Metrological complementarity reveals the Einstein-Podolsky-Rosen paradox, Nat. Commun. **12**, 2410 (2021).

[342] M. Gessner, A. Smerzi, and L. Pezzè, Metrological Nonlinear Squeezing Parameter, Phys. Rev. Lett. **122**, 090503 (2019).

[343] D. R. M. Arvidsson-Shukur, N. Yunger Halpern, H. V. Lepage, A. A. Lasek, C. H. W. Barnes, and S. Lloyd, Quantum advantage in postselected metrology, Nat. Commun. **11**, 3775 (2020).

[344] J. G. Kirkwood, Quantum statistics of almost classical assemblies, Phys. Rev. **44**, 31 (1933).

[345] M. Walschaers, N. Treps, B. Sundar, L. D. Carr, and V. Parigi, Emergent complex quantum networks in continuous-variables non-gaussian states, arXiv:2012.15608 [quant-ph] (2021).

[346] H. H. Schaefer, *Topological Vector Spaces* (Springer New York, New York, NY, 1971).

[347] J. B. Conway, *A Course in Functional Analysis* (Springer New York, New York, NY, 2007).