Detecting continuous-variable entanglement in phase space with the Q distribution

Martin Gärttner [ORCID],[1,2,3,4,*] Tobias Haas [ORCID],[5,†] and Johannes Noll [ORCID][3,‡]

[1]*Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany*
[2]*Physikalisches Institut, Universität Heidelberg, Im Neuenheimer Feld 226, 69120 Heidelberg, Germany*
[3]*Kirchhoff-Institut für Physik, Universität Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany*
[4]*Institute of Condensed Matter Theory and Optics, Friedrich-Schiller-University Jena, Max-Wien-Platz 1, D-07743 Jena, Germany*
[5]*Centre for Quantum Information and Communication, École polytechnique de Bruxelles,
CP 165, Université libre de Bruxelles, 1050 Brussels, Belgium*

We prove a general class of continuous variable entanglement criteria based on the Husimi Q distribution, which represents a quantum state in canonical phase space, by employing a theorem by Lieb and Solovej. We discuss their generality, which roots in the possibility to optimize over the set of concave functions, from the perspective of continuous majorization theory and show that with this approach families of entropic as well as second moment criteria follow as special cases. All derived criteria are compared with corresponding marginal based criteria, and the strength of the phase-space approach is demonstrated for a family of prototypical example states where only our criteria flag entanglement. Furthermore, we explore their optimization prospects in two experimentally relevant scenarios characterized by sparse data: Finite detector resolution and finite statistics. In both scenarios optimization leads to clear improvements enlarging the class of detected states and the signal-to-noise ratio of the detection, respectively.

I. INTRODUCTION

The quest for efficient methods for the detection of entanglement in continuous variable systems dates back as far as the Einstein-Podolsky-Rosen (EPR) paradox [1] and has received renewed interest through the rise of optical quantum technologies for which entanglement is a crucial ingredient [2–5]. Over the past two decades, a plethora of entanglement criteria[1] emerged. Many of them root in demonstrating the negativity of the partially transposed state [6,7] by means of violating uncertainty relations of suitable chosen observables [3,4,8].[2]

Widely used are nonlocal EPR-type operators $X_1 + X_2$ and $P_1 - P_2$, which capture correlations between two systems labeled 1 and 2 [1]. For these variables, criteria have been formulated in terms of second moments in seminal works by Duan, Giedke, Cirac, and Zoller (DGCZ) [10] and Mancini, Giovannetti, Vitali, and Tombesi (MGVT) [11,12] (see also Ref. [13]). They have been shown to be necessary and sufficient for separability in case of Gaussian states [2,5,14,15]

but are known to be rather weak beyond because second moments only reveal partial information about a measured distribution. Here, a significant upgrade have been entropic criteria, derived by Walborn, Taketani, Salles, Toscano, and de Matos Filho (WTSTD) in Ref. [16] and their generalization to Rényi entropies by Saboia, Toscano and Walborn (STW) [17]. Strengthened versions of all these criteria to certify steering have been discussed, too [18–22].

Besides other notable approaches utilizing modular variables [23,24], spin observables with discrete spectra [25], moments of the bipartite density operator [26,27], as well as the quantum Fisher matrix [28,29], all mentioned methods rely on the detection of marginal distributions of the Wigner W distribution. However, the Husimi Q distribution also constitutes an adequate phase-space representation of the quantum state [30–32] and can be accessed in a single experimental setting [33,34]. Until recently [35,36], the possibility of formulating entanglement criteria based on the Husimi Q distribution has been overlooked. Here, we derive an extremely general class of entanglement criteria from the Husimi Q distribution based on concave functions, generalizing the work in Ref. [36].

In the spirit of employing uncertainty relations for entanglement detection, we make use of the most general form of the uncertainty principle in phase space: The Lieb-Solovej theorem [37–39]. It arose from generalizing a lower bound on the differential entropy—the so-called Wehrl entropy [40,41]—of the Husimi Q distribution [42–45] and has been formulated for various algebras [46–49]. Ultimately, the Lieb-Solovej theorem is a majorization relation stating that the vacuum (or any coherent) distribution majorizes all other

---

*martin.gaerttner@uni-jena.de

†tobias.haas@ulb.be

‡johannes.noll@stud.uni-heidelberg.de

[1]We use the terms entanglement criteria and separability criteria interchangeably. Such criteria are usually stated as inequalities that are fulfilled for all separable states, and thus their violation shows that a state is entangled.

[2]This ansatz generically excludes the possibility of detecting bound entanglement [9].

distributions (see also [50] for the relation between majorization and entanglement for finite dimensional Hilbert spaces). A similar statement has been conjectured for the Wigner W distribution when restricting to Wigner-positive states [51–54], but is lacking for marginal distributions. In this sense, the Husimi Q distribution offers the unique opportunity to formulate such general relations for uncertainty and entanglement.

However, in experiments, it is of course not possible to measure a distribution over a continuous space to arbitrary precision. Measuring the Husimi Q distribution means to approximate it based on a finite experimental data set [2,5,14,55]. Thus, a relevant question is not only whether our criterion can flag entanglement in the limit of precise knowledge of the distribution, but also how one can, based on a fixed measurement budget and resolution, maximize the statistical significance of the detection. We show in this work that the generality of our witness leads to a significant advantage with regard to this task by optimizing over different choices of the concave function.

A common scheme for measuring the Husimi Q distribution is the application of a coherent displacement followed by the detection of the vacuum projection [56–58]. The experimental capabilities for realizing this scheme have been demonstrated for microwave photons in cavity QED systems [59], including the bipartite setup [60], atomic gases in optical cavities [61,62], and trapped ions [63,64]. In this scheme the value of the Husimi Q distribution is obtained on a grid of points in the four-dimensional phase space, where each grid point corresponds to a separate experimental measurement in which the corresponding displacement operation is applied. Obtaining high-resolution data is thus challenging in terms of the required experimental resources [65], which motivates us to study the prospects of our entanglement criteria in the case where the Husimi Q distribution is only known on a discrete grid with finite resolution.

Another established way of accessing the Husimi Q distribution is heterodyne detection [33,66], where the system modes are split by sending them on a lossless 50 : 50 beam splitter and subsequently each output mode is interfered with a so-called local oscillator field in a highly excited coherent state. The phase of the local oscillator controls which field quadrature is measured for each output mode and allows one to simultaneously detect the local quadratures with minimal uncertainty. This corresponds to direct sampling from the Husimi Q distribution [67]. This scheme has been realized with optical photons [68–71] and more recently with ultracold atomic gases [72], including the bipartite setting [73]. In particular for cold atom experiments the experimental repetition rate is rather low, making it challenging to obtain sample data with high statistics. This motivates us to study the potential of our entanglement criteria for optimizing the signal-to-noise ratio of the entanglement detection for a given budget of experimental samples.

Let us remark that in this work we discuss our entanglement criteria in detail from both a theoretical and a practical perspective. A more concise description of the criteria and their application is provided in Ref. [74].

The remainder of this work is structured as follows: In Sec. II we introduce the necessary background on canonical phase space, including the formulation of the uncertainty principle in terms of the Husimi Q distribution, and on continuous majorization theory. After proving the general entanglement criteria in Sec. III we discuss the effects of the various free parameters entering them. In particular, the meaning of the choice of the concave function f occurring in the criteria is made accessible by means of continuous majorization theory. Subsequently, we examine specific criteria, which includes entropic and second moment criteria. We show that the latter are the strongest possible state-independent criteria that can be derived from our general criteria and that they are necessary and sufficient for Gaussian states after optimizing over scaling parameters. In Sec. IV we provide a comparison of our criteria to well-known criteria based on marginals of the Wigner W distribution. We show that our second moment criteria imply the DGCZ criteria and are neither stronger nor weaker than the MGVT criteria. We relate our criteria to marginal based entropic criteria by WTSTD and STW, establishing a detailed understanding of their strengths and limitations and show an example state for which our criteria detect a class of states that is not detectable by any other of the aforementioned entanglement criteria. The various criteria are generalized to arbitrary coarse-grained measurements in Sec. V, which is accompanied with a comparison for the finite-resolution Husimi Q distribution of the two-mode squeezed vacuum state. Thereupon, we exemplify how the signal-to-noise ratio can be improved by optimizing over the concave function in the case of finite statistics in Sec. VI. A discussion of our results and of possible future directions is given in Sec. VII.

Notation. We employ natural units $\hbar = 1$, write quantum operators with bold letters, e.g., $\boldsymbol{\rho}$, and use a bar, e.g., $\bar{Q}$, to mark vacuum quantities. If not specified differently, integrals run over the Euclidean plane $\mathbb{R}^2$.

II. PRELIMINARIES

We start with introducing the reader to the Husimi Q distribution and discuss in which way it is constrained by the uncertainty principle from the perspective of continuous majorization theory. Let us stress that *all* presented concepts can be generalized to arbitrary simple Lie groups. Readers familiar with these topics may want to omit this section.

A. Conjugate variables and coherent states

Throughout this work, we are concerned with continuous variable quantum systems. In the monopartite setup these are characterized by an infinite-dimensional Hilbert space, $\dim \mathcal{H} = \infty$, and canonical commutation relations

$$[\boldsymbol{X}, \boldsymbol{P}] = i\mathbb{1} \tag{1}$$

for two Hermitian operators $\boldsymbol{X}$ and $\boldsymbol{P}$ with continuous and unbounded spectra. A prime example for this description is the harmonic oscillator, in which case $\boldsymbol{X}$ represents the position while $\boldsymbol{P}$ describes the momentum, but it also applies to the quadratures of an electromagnetic field or suitably chosen spin components in a spinor Bose-Einstein condensate [72,73,75–77]. We also introduce creation and annihilation operators

$$\boldsymbol{a}^{\dagger} = \frac{1}{\sqrt{2}}(\boldsymbol{X} - i\boldsymbol{P}), \quad \boldsymbol{a} = \frac{1}{\sqrt{2}}(\boldsymbol{X} + i\boldsymbol{P}), \tag{2}$$

respectively, which fulfill bosonic commutation relations

$$[\boldsymbol{a}, \boldsymbol{a}^\dagger] = \mathbb{1}, \tag{3}$$

and single out a unique vacuum state $|0\rangle$ via $\boldsymbol{a}\,|0\rangle = 0$.

We construct the set of coherent states axiomatically following the group-theoretic approach, which allows us to associate a set of coherent states to any simple Lie group, for example SU(2) describing quantum spins [78–80]. We start from the Heisenberg-Weyl algebra $\mathbb{H}_4$, which consists out of the four operators $\{\boldsymbol{a}, \boldsymbol{a}^\dagger, \boldsymbol{a}^\dagger\boldsymbol{a}, \mathbb{1}\}$ and is defined via the commutation relation (3). The subgroup leaving the vacuum state $|0\rangle$ invariant up to a phase is $U(1) \otimes U(1)$ as one might apply rotations in the complex plane generated by $\boldsymbol{a}^\dagger\boldsymbol{a}$ or $\mathbb{1}$ without changing the vacuum. Then, the displacement operator

$$\boldsymbol{D}(\alpha) = e^{\alpha \boldsymbol{a}^\dagger - \alpha^* \boldsymbol{a}}, \tag{4}$$

with a complex phase conveniently parametrized as

$$\alpha = \frac{1}{\sqrt{2}}(x + ip), \tag{5}$$

is nothing but a unitary representation of the coset space $\mathbb{H}_4 / U(1) \otimes U(1)$. Applying this operator to the vacuum generates the set of canonical coherent states

$$|\alpha\rangle = \boldsymbol{D}(\alpha)\,|0\rangle. \tag{6}$$

Note that for conjugate operators defined via Eq. (1) one may equally introduce coherent states as eigenvalues of the annihilation operator

$$\boldsymbol{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle. \tag{7}$$

However, the latter definition does not generalize to systems with degrees of freedom described by other algebras.

Coherent states have three interesting properties which are of importance for our later considerations. First, they are *not* orthogonal:

$$|\langle \alpha | \alpha' \rangle|^2 = e^{-|\alpha - \alpha'|^2}, \tag{8}$$

but orthogonality is approximately restored for sufficiently distinct α and α'. Second, they span an overcomplete basis

$$\mathbb{1} = \int \frac{dx\,dp}{2\pi} |\alpha\rangle\langle\alpha|, \tag{9}$$

and third, they minimize *all* uncertainty relations, for example the Heisenberg uncertainty relation formulated in terms of variances,

$$\sigma_x \sigma_p = \tfrac{1}{2}. \tag{10}$$

B. Husimi Q distribution

Equation (9) shows that pure coherent-state projectors $\boldsymbol{E}_\alpha = |\alpha\rangle\langle\alpha|$ constitute a positive operator-valued measure (POVM), which defines the Husimi Q distribution [30,39,66,81]:

$$Q(x, p) = \mathrm{Tr}\{\boldsymbol{\rho}\boldsymbol{E}_\alpha\} = \langle\alpha|\boldsymbol{\rho}|\alpha\rangle, \tag{11}$$

with the parametrization (5) understood. It covers the quantum-mechanical phase space and corresponds to a joint measurement of position x and momentum p with minimum (but still nonzero) uncertainty.

The Husimi Q distribution comes with several desired properties. Most importantly, it is non-negative and bounded from above by unity,

$$0 \leqslant Q(x, p) \leqslant 1, \tag{12}$$

for all $(x, p) \in \mathbb{R}^2$. This is a profound advantage over the Wigner W distribution, which can become negative and hence measures of localization such as entropies are not defined for all states. The two distributions are related via a Weierstrass transform with respect to the vacuum,

$$Q(x, p) = (W * \bar{W})(x, p)$$
$$= 2\pi \int dx'\,dp'\, W(x', p')\bar{W}(x - x', p - p'), \tag{13}$$

with the vacuum distribution

$$\bar{W}(x, p) = \frac{1}{\pi} e^{-(x^2 + p^2)}, \tag{14}$$

which removes the negativities of the Wigner W distribution.

Furthermore, the Husimi Q distribution is normalized to unity in the sense of

$$1 = \int \frac{dx\,dp}{2\pi} Q(x, p), \tag{15}$$

following from $\mathrm{Tr}\{\boldsymbol{\rho}\} = 1$. However, it *cannot* be considered a probability density function in a strict sense as the underlying random variables do not constitute mutually exclusive events by non-orthogonality (8). It is therefore convenient to refer to it as a quasiprobability distribution.

As a final remark, the Husimi Q distribution serves as the quantum-mechanical extension of the Boltzmann distribution, to which it converges in the classical limit $\hbar \to 0$ [40,41].

C. Uncertainty principle

In simple words, the uncertainty principle expresses the fact that noncompatible observables cannot be measured simultaneously with arbitrary precision [82]. Hence, it sets bounds on measures of localization of the distributions obtained when measuring the observables of interest. The simplest example for a measure of localization is the variance [83–86], but over the last decades especially classical entropies became reasonable alternatives [35,51,87–91].

The differential entropy associated with the Husimi Q distribution is the Wehrl entropy [40,41],

$$S(Q) = -\int \frac{dx\,dp}{2\pi} Q(x, p) \ln Q(x, p), \tag{16}$$

which is bounded from below by the Wehrl-Lieb inequality

$$S(Q) \geqslant 1, \tag{17}$$

with equality if and only if the state under consideration is a pure coherent state [37,39,42,44,45].

The latter statement has been generalized to arbitrary concave averages by allowing for a concave function $f : [0, 1] \to \mathbb{R}$ with $f(0) = 0$ in the integrand, which is the Lieb-Solovej

theorem[3] [37],

$$\int \frac{dxdp}{2\pi} f(Q) \geqslant \int \frac{dxdp}{2\pi} f(\bar{Q}), \tag{18}$$

again with equality only for coherent states, containing (17) as a special case for $f(t) = -t \ln t$. Therein, the vacuum Husimi Q distribution is given by

$$\bar{Q}(x, p) = e^{-\frac{1}{2}(x^2+p^2)}, \tag{19}$$

which may be replaced on the right-hand side of (18) by any coherent state. We remark that the condition $f(0) = 0$ ensures the finiteness of both sides (for $f(0) \neq 0$ both sides diverge to $+\infty$ or $-\infty$), while the importance of f being concave becomes apparent in the context of continuous majorization theory, see Sec. II D.

We note that (18) is *not* tight for squeezed coherent states as the unitary squeezing operator Ξ does only reduce to a linear symplectic map $\Xi = \text{diag}(\xi, 1/\xi)$ with $\xi > 0$ in phase space such that $(x, p) \rightarrow (\xi x, p/\xi)$ when applied to the Wigner W distribution. More precisely, the correspondence between Gaussian unitaries on states and affine symplectic maps on the Husimi Q distribution is broken by the Weierstrass transformation (13) in case of squeezing. From a quantum optics perspective, this is due to the vacuum signal in the heterodyne measurement being not equally squeezed with the input signal. Therefore, one has to introduce a squeezing transformation in (18) and minimize with respect to ξ in order to end up with a relation which is tight for *all* pure Gaussian states.

D. Continuous majorization theory

Most generally, measures of localization are described within the framework of majorization theory. For continuous distributions, *localization* replaces the notion of *ordering* in case of discrete probability distributions. Some probability density function $\varphi(x, p)$ is said to be majorized by another distribution $\varphi'(x, p)$, written as $\varphi \prec \varphi'$, if and only if

$$\int \frac{dxdp}{2\pi} f(\varphi) \geqslant \int \frac{dxdp}{2\pi} f(\varphi') \tag{20}$$

for all concave functions $f : \mathcal{I} \rightarrow \mathbb{R}$ with $f(0) = 0$ and $\mathcal{I} = [0, \max\{\varphi, \varphi'\}]$ [52–54,92]. Hence, $\varphi \prec \varphi'$ expresses the intuition that φ' is more localized than φ in most general terms.

Interestingly, the Lieb-Solovej theorem (18) is fundamentally a statement about localization of distributions in phase space and states that no Husimi Q distribution Q is more localized than the vacuum Husimi Q distribution $\bar{Q}$ (or equivalently all coherent ones), i.e., $Q \prec \bar{Q}$. Note that it is sufficient for $Q(x, p)$ to be continuous, non-negative and normalized in order to apply (20) and hence the violation of Kolmogorov's third axiom can be safely disregarded.

To make these insights more explicit, we introduce some techniques from continuous majorization theory. We begin with the level function $m_Q(t)$, which is associated with a given Husimi Q distribution $Q(x, p)$ and is defined as the

phase-space measure of the domain $\mathcal{A}$ for which the value of the distribution exceeds a given threshold t, i.e.,

$$m_Q(t) = \int_{\mathcal{A}} \frac{dxdp}{2\pi} \text{ with } \mathcal{A} = \{(x, p) : Q(x, p) \geqslant t\}, \tag{21}$$

where $t \in \mathcal{I} = [0, 1]$ as a result of the boundedness of the Husimi Q distribution (12). The definition (21) is conveniently rewritten as

$$m_Q(t) = \int \frac{dxdp}{2\pi} \Theta[Q(x, p) - t], \tag{22}$$

with Θ denoting the Heaviside Θ function. The negative derivative of the latter quantity defines the density-level function

$$\mu_Q(t) = -\frac{d}{dt} m_Q(t) = \int \frac{dxdp}{2\pi} \delta[Q(x, p) - t], \tag{23}$$

where δ is the Dirac δ distribution.

With (23) at hand we can simplify the integrals appearing in the Lieb-Solovej theorem (18) as

$$\int \frac{dxdp}{2\pi} f(Q) = \int_{\mathcal{I}} dt f(t) \mu_Q(t). \tag{24}$$

If f is differentiable almost everywhere on $\mathcal{I}$ we can equally write the following with (22):

$$\int \frac{dxdp}{2\pi} f(Q) = \int_{\mathcal{I}} dt \frac{df(t)}{dt} m_Q(t), \tag{25}$$

where we used $f(0) = 0$ and $m_Q(1) = 0$. The latter two equations show that continuous majorization relations ultimately boil down to a comparison of density-level or level functions. Note that these functions remain invariant for all transformations with unit determinant, i.e., those which preserve phase-space area elements, including symplectic transformations such as rotations and squeezings of the coordinate axes as well as displacements [see (74) for the case of Gaussian distributions].

We illustrate all introduced concepts by comparing the Husimi Q distributions and their localization of the vacuum state (19) with the first excited Fock state $|1\rangle$, for which we obtain

$$Q(x, p) = \frac{x^2 + p^2}{2} e^{-\frac{1}{2}(x^2+p^2)}. \tag{26}$$

The two distributions are shown in Figs. 1(a) and 1(b), respectively. The corresponding level functions are obtained by integrating over the area enclosed by the Husimi Q distribution (and dividing by 2π) at a given height, which we sketch for a few heights with black and dark orange circles, respectively. The resulting level and density-level functions are plotted in Fig. 1(c), where the aforementioned values are represented by plot markers. While the vacuum curves take positive values up to $t = 1$, the Fock curves evaluate to zero for $t > 1/e$. If we multiply the density-level functions with a concave function which overweighs small t, for example $f(t) = t^{1/5}$, the integrated area of the Fock curve enclosed with the t axis will be larger than the area below the vacuum curve. Similarly, if we choose a function pronouncing large t, for example $f(t) = -t^5$, the area will again be larger for the Fock curve due to the minus sign. This exemplifies the majorization relation $Q \prec \bar{Q}$ in form of the Lieb-Solovej theorem

---

[3]For generalizations see Refs. [37,46] for SU(2), [47] for symmetric SU(N) and [48,49] for SU(1, 1).
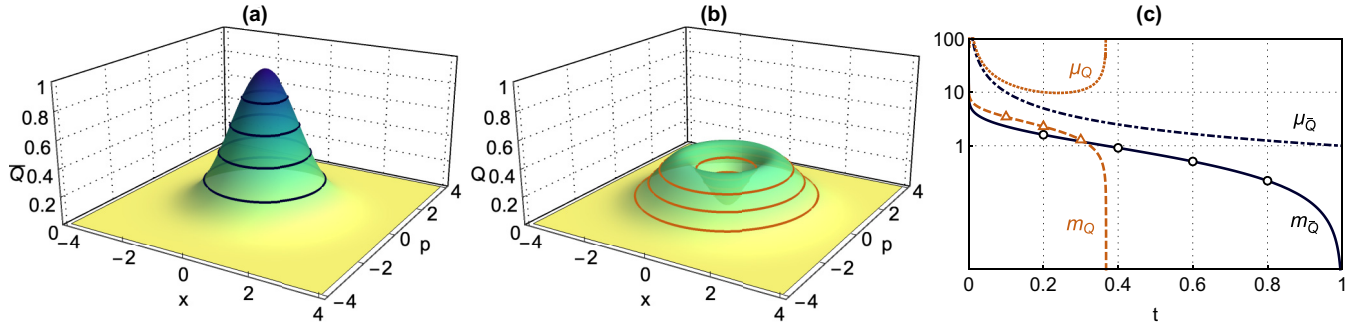
FIG. 1. Panels (a) and (b) show the Husimi Q distributions of the vacuum state (19) and the first Fock state (26), respectively. The areas enclosed by the black and dark orange circles (divided by 2π) give the values for the corresponding level functions at given height, which are depicted as plot markers in panel (c), together with the level and density-level functions in a logarithmic plot. For $t \leqslant 1/e$ we have $\mu_Q(t) \geqslant \mu_{\bar{Q}}(t)$, while for $1/e < t \leqslant 1$ we obtain $\mu_{\bar{Q}}(t) \geqslant \mu_Q(t) = 0$. Irrespective of the concave function $f(t)$ multiplied to the density-level functions, the area below this product for the vacuum state is always smaller than for the Fock state, which illustrates the majorization relation $Q \prec \bar{Q}$.

(18) with (24) understood for two choices for f. The intuition provided here will be useful later when choosing suitable f for certifying entanglement.

III. ENTANGLEMENT CRITERIA

We extend the latter considerations to the bipartite case and derive various classes of entanglement criteria using the Lieb-Solovej theorem (18) for general nonlocal observables.

A. Bipartite setup and nonlocal operators

We now investigate two continuous-variable quantum systems forming a bipartition, which is described by a bipartite quantum state $\boldsymbol{\rho}_{12}$. The degrees of freedom of the two subsystems are encoded in the local operators $\boldsymbol{X}_j, \boldsymbol{P}_j$, where $j \in \{1, 2\}$ labels the subsystem, which fulfill

$$[\boldsymbol{X}_j, \boldsymbol{P}_k] = i\delta_{jk}\mathbb{1}. \quad (27)$$

For the sake of generality we incorporate the effect of rotations in the local subsystems, which will play an important role in the comparison of our criteria to criteria based on marginal distributions. We describe a rotation around an angle $\vartheta \in [0, 2\pi)$ by a rotation matrix

$$\mathcal{T}(\vartheta) = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{pmatrix}, \quad (28)$$

leading to rotated operators via the transformation

$$(\boldsymbol{X}_j, \boldsymbol{P}_j) \rightarrow (\boldsymbol{R}_j, \boldsymbol{S}_j) = \mathcal{T}(\vartheta_j)(\boldsymbol{X}_j, \boldsymbol{P}_j). \quad (29)$$

As the canonical commutation relations (27) are conserved by local rotations, we can associate Husimi Q distributions to the full system as well as to both subsystems along the lines of Secs. II A and II B. More precisely, the global Husimi Q distribution is obtained by applying a POVM formed by local coherent states $\boldsymbol{E}_j = |\alpha_j\rangle \langle \alpha_j|$ to both subsystems

$$Q(r_1, s_1, r_2, s_2) = \mathrm{Tr}\{\boldsymbol{\rho}_{12}(\boldsymbol{E}_1 \otimes \boldsymbol{E}_2)\}$$
$$= (\langle\alpha_1| \otimes \langle\alpha_2|)\boldsymbol{\rho}_{12}(|\alpha_1\rangle \otimes |\alpha_2\rangle)). \quad (30)$$

The resulting distribution is still bounded in the sense of (12), but now normalized to unity with respect to the

four-dimensional phase-space measure $dr_1 ds_1 dr_2 ds_2/4\pi^2$. Local Husimi Q distributions result from integrating out the complementary phase-space variables, for example,

$$Q(r_1, s_1) = \mathrm{Tr}\{\boldsymbol{\rho}_1\boldsymbol{E}_1\}$$
$$= \int \frac{dr_2 ds_2}{2\pi} Q(r_1, s_1, r_2, s_2) \quad (31)$$

for subsystem 1 and analogously for subsystem 2.

Following Refs. [1,10,12], we adopt nonlocal operators for the study of entanglement by adding and subtracting local operators. Additionally, we allow for relative scalings between the four local operators, such that we end up with

$$\boldsymbol{R}_\pm = a_1\boldsymbol{R}_1 \pm a_2\boldsymbol{R}_2, \quad \boldsymbol{S}_\pm = b_1\boldsymbol{S}_1 \pm b_2\boldsymbol{S}_2, \quad (32)$$

with $a_1, b_1, a_2, b_2 \geqslant 0$ and $a_1 b_1 = a_2 b_2$. These operators fulfill the commutation relations

$$[\boldsymbol{R}_\pm, \boldsymbol{S}_\pm] = i(a_1 b_1 + a_2 b_2)\mathbb{1}, \quad [\boldsymbol{R}_\pm, \boldsymbol{S}_\mp] = 0, \quad (33)$$

showing that pairs of operators with equal indices represent independent oscillator modes with canonical commutation relations.

To express the global Husimi Q distribution in terms of the nonlocal variables (32), we employ a variable transformation

$$Q(r_+, s_+, r_-, s_-)$$
$$= \frac{1}{4a_1 b_1 a_2 b_2} Q\left(\frac{r_+ + r_-}{2a_1}, \frac{s_+ + s_-}{2b_1}, \frac{r_+ - r_-}{2a_2}, \frac{s_+ - s_-}{2b_2}\right), \quad (34)$$

where the determinant of the Jacobian matrix evaluates to $1/(4a_1 b_1 a_2 b_2)$. For entanglement criteria, we are particularly interested in the marginals over the mixed variables pairs $(r_\pm, s_\mp)$ of the latter distribution

$$Q_\pm \equiv Q_\pm(r_\pm, s_\mp) = \int \frac{dr_\mp ds_\pm}{2\pi} Q(r_+, s_+, r_-, s_-), \quad (35)$$

which are not constrained by the uncertainty principle as the underlying phase-space operators commute (33).

In contrast, the distributions corresponding to equal signs,

$$Q(r_\pm, s_\pm) = \int \frac{dr_\mp ds_\mp}{2\pi} Q(r_+, s_+, r_-, s_-), \qquad (36)$$

constitute true Husimi Q distributions and may equally be defined via coherent-state projectors with respect to $(\boldsymbol{R}_\pm, \boldsymbol{S}_\pm)$ after a partial trace over the complementary degrees of freedom. Therefore, they are restrained by the uncertainty principle in the form of the Lieb-Solovej theorem (18). However, the different normalization affects the size of the codomain of the Husimi Q distribution (and hence the domain of the concave function f) as well as the form of the vacuum expression (19), which now reads

$$\bar{Q}(r_\pm, s_\pm) = \frac{1}{a_1 b_1 + a_2 b_2} \exp\left(-\frac{1}{2}\frac{r_\pm^2 + s_\pm^2}{a_1 b_1 + a_2 b_2}\right). \quad (37)$$

Adapted to this setup, the Lieb-Solovej theorem states

$$\int \frac{dr_\pm ds_\pm}{2\pi} f(Q) \geqslant \int \frac{dr_\pm ds_\pm}{2\pi} f(\bar{Q}), \qquad (38)$$

for any concave $f : [0, (a_1 b_1 + a_2 b_2)^{-1}] \to \mathbb{R}$ with $f(0) = 0$. As discussed at the end of Sec. II C, we additionally have to allow for an optimization over a squeezing transformation Ξ in the nonlocal Wigner W distribution $W_\pm(r_\pm, s_\mp)$ in order to render the latter inequality tight for all pure Gaussian states. As the tightness of an uncertainty relation is closely related to the detection capabilities of entanglement criteria derived from such a relation, we explicitly allow for this possibility in the following.

B. General criteria

We now show that the distribution $Q_\pm$ is nontrivially constrained for all separable states. A bipartite quantum state $\boldsymbol{\rho}_{12}$ is called separable if it can be written as a convex combination of product states, i.e.,

$$\boldsymbol{\rho}_{12} = \sum_j p_j (\boldsymbol{\rho}_1^j \otimes \boldsymbol{\rho}_2^j), \qquad (39)$$

where $p_j \geqslant 0$ denotes a discrete probability distribution with $\sum_j p_j = 1$. A widespread method to investigate the separability of a given quantum state is to apply a positive but not completely positive trace-preserving map and check whether the resulting operator is still non-negative, i.e., constitutes a valid density operator [3,4]. A prime example for this method is the Peres-Horodecki (PPT) criterion [6,7] which utilizes the partial transpose $\boldsymbol{T}_2$ (conveniently applied to subsystem two) leading to a necessary, but in general not sufficient, condition for separability: For every separable state $\boldsymbol{\rho}_{12}$ the operator $\boldsymbol{\rho}_{12} \to \boldsymbol{\rho}'_{12} = (\mathbb{1}_1 \otimes \boldsymbol{T}_2)(\boldsymbol{\rho}_{12})$ is non-negative $\boldsymbol{\rho}'_{12} \geqslant 0$. In particular, $\boldsymbol{\rho}'_{12}$ is *physical* for separable $\boldsymbol{\rho}_{12}$, and hence the Lieb-Solovej theorem (38) applies also to Husimi Q distributions of $\boldsymbol{\rho}'_{12}$.

To relate these distributions to observable distributions we have to translate the action of the partial transpose $\boldsymbol{T}_2$ on the state $\boldsymbol{\rho}_{12}$ into an action onto the variable pairs $(r_\pm, s_\pm)$ spanning phase space. It is straightforward to show that the partial transpose $\boldsymbol{T}_2$ flips the sign of the local variable $s_2 \to -s_2$, which holds for all quasiprobability distributions covering phase space [13]. On the level of the nonlocal variables

(32) this is equivalent to $s_\pm \to s_\mp$ implying that the partial transpose $\boldsymbol{T}_2$ corresponds to the transformation $Q(r_\pm, s_\pm) \to Q'(r_\pm, s_\pm) = Q_\pm(r_\pm, s_\mp)$.

By the PPT criterion the distribution $Q_\pm(r_\pm, s_\mp)$ is thus constrained by the Lieb-Solovej theorem (38), where the vacuum expression on the right-hand side has to be expressed in the mixed variable pairs,

$$\bar{Q}'_\pm(r_\pm, s_\mp) = \bar{Q}(r_\pm, s_\mp)$$

$$= \frac{1}{a_1 b_1 + a_2 b_2} \exp\left(-\frac{1}{2}\frac{r_\pm^2 + s_\mp^2}{a_1 b_1 + a_2 b_2}\right). \quad (40)$$

To make this statement more explicit we introduce a witness functional $\mathcal{W}_f$,

$$\mathcal{W}_f = \int \frac{dr_\pm ds_\mp}{2\pi} [f(Q_\pm) - f(\bar{Q}'_\pm)], \qquad (41)$$

with concave $f : \mathcal{J} \to \mathbb{R}$ fulfilling $f(0) = 0$ defined over the interval $\mathcal{J} = [0, \max\{\max Q_\pm, (a_1 b_1 + a_2 b_2)^{-1}\}] \subseteq \mathbb{R}^+$. Then, our main result is that $\mathcal{W}_f$ is non-negative for all separable states, i.e.,

$$\boldsymbol{\rho}_{12} \text{ separable} \Rightarrow \mathcal{W}_f \geqslant 0. \qquad (42)$$

If the latter inequality is violated for a given state $\boldsymbol{\rho}_{12}$, entanglement is demonstrated.

Let us stress the generality of our criteria (42), as one may optimize over the two local rotation angles ϑ_1, ϑ_2, the four scaling parameters a_1, b_1, a_2, b_2 under the constraint $a_1 b_1 = a_2 b_2$, the squeezing transformation Ξ and the class of concave functions f with $f(0) = 0$ to witness entanglement. In the following, we discuss the influences of all quantities in detail to develop a systematic understanding of their effect.

C. Rotation angles and scaling parameters

We begin our study with the rotation angles ϑ_1, ϑ_2 and the scaling parameters a_1, b_1, a_2, b_2. While we already used a matrix representation for the former in (28), we write

$$\mathcal{U}(a_1, a_2, b_1, b_2) = \begin{pmatrix} a_1 & 0 & a_2 & 0 \\ 0 & b_1 & 0 & -b_2 \\ a_1 & 0 & -a_2 & 0 \\ 0 & b_1 & 0 & b_2 \end{pmatrix} \qquad (43)$$

for the matrix mixing the local operators in (32), i.e., $(\boldsymbol{R}_+, \boldsymbol{S}_-, \boldsymbol{R}_-, \boldsymbol{S}_+) = \mathcal{U}(a_1, a_2, b_1, b_2)(\boldsymbol{R}_1, \boldsymbol{S}_1, \boldsymbol{R}_2, \boldsymbol{S}_2)$. Then, the four initial local operators $(\boldsymbol{X}_1, \boldsymbol{P}_1, \boldsymbol{X}_2, \boldsymbol{P}_2)$ transform as

$$(\boldsymbol{X}_1, \boldsymbol{P}_1, \boldsymbol{X}_2, \boldsymbol{P}_2) \to (\boldsymbol{R}_+, \boldsymbol{S}_-, \boldsymbol{R}_-, \boldsymbol{S}_+)$$
$$= \mathcal{G}(\vartheta_1, \vartheta_2, a_1, a_2, b_1, b_2)(\boldsymbol{X}_1, \boldsymbol{P}_1, \boldsymbol{X}_2, \boldsymbol{P}_2),$$
$$(44)$$

with the full transformation matrix

$$\mathcal{G}(\vartheta_1, \vartheta_2, a_1, a_2, b_1, b_2)$$
$$= \mathcal{U}(a_1, a_2, b_1, b_2)[\mathcal{T}(\vartheta_1) \oplus \mathcal{T}(\vartheta_2)] \qquad (45)$$

describing the effects of our six parameters of interest.

For equal scaling in the local quadratures, i.e., $a_1 = b_1$ and $a_2 = b_2$ (which includes the most relevant case $a_1 = b_1 = a_2 = b_2 = 1$), the matrix $\mathcal{G}$ can be rewritten as a rotation
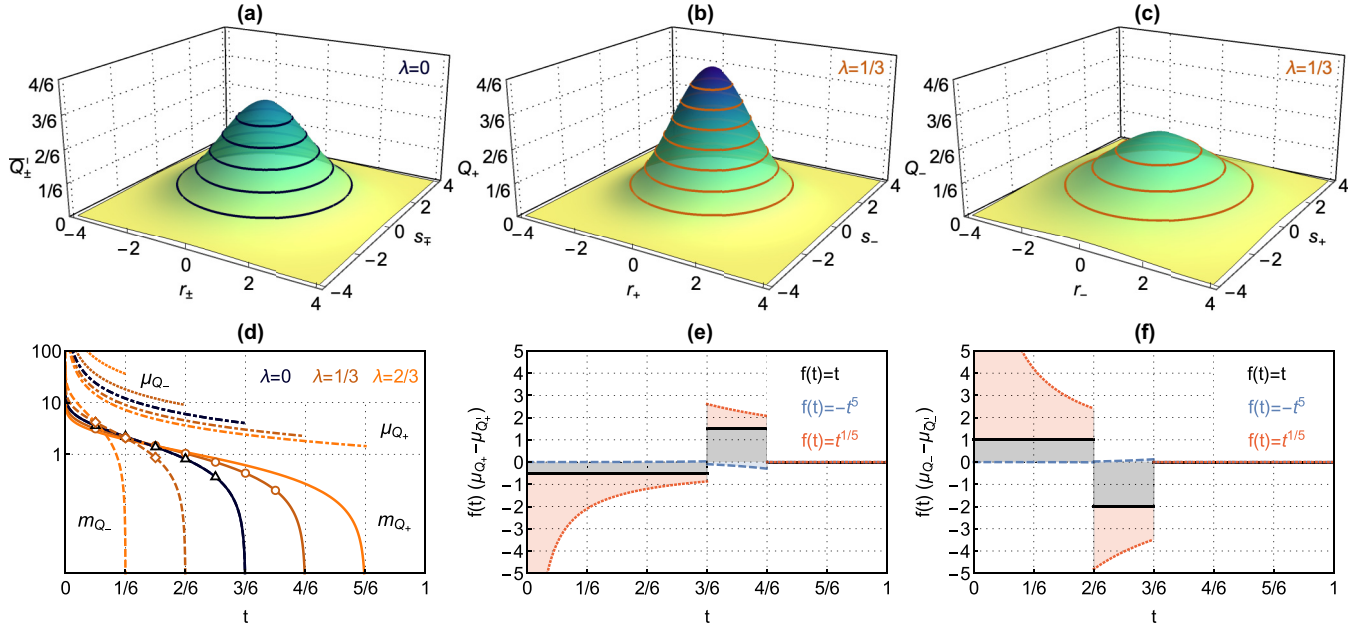
FIG. 2. Distributions of the two-mode squeezed vacuum state (48) with $a_1 = b_1 = a_2 = b_2 = 1$, $\vartheta_1 = \vartheta_2 = 0$, and $\xi = 1$ for (a) $\lambda = 0$ and $\lambda = 1/3$ in (b) (r_+, s_-) and (c) (r_-, s_+) variables, respectively. The areas enclosed by the circles give the values level functions at given height, which are depicted as plot markers in panel (d), together with level and density-level functions from (49) of both variable pairs for selected values of λ. For $\lambda = 1/3$, the integrand of the witness functional (47) is shown for (e) (r_+, s_-) and (f) (r_-, s_+) variables.

in subsystem 1 around an angle ϑ (subsystem 2 is left unchanged) and a transformation to nonlocal variables $(r_\pm, s_\mp)$ followed by a rotation in these variables around $\phi = -\vartheta_2$, i.e.,

$$\mathcal{G}(\vartheta_1, \vartheta_2, a_1, a_2, a_1, a_2)$$
$$= [\mathcal{T}(\phi) \oplus \mathcal{T}(\phi)]\mathcal{G}(\vartheta, 0, a_1, a_2, a_1, a_2), \qquad (46)$$

when choosing $\vartheta = \arctan(\cos\vartheta_1 \cos\vartheta_2 - \sin\vartheta_1 \sin\vartheta_2, \cos\vartheta_1 \sin\vartheta_2 + \sin\vartheta_1 \cos\vartheta_2)$, where $\arctan(x, y)$ denotes the arcus tangens in the Euclidean plane (x, y). As the final rotations $\mathcal{T}(\phi) \oplus \mathcal{T}(\phi)$ leave our criteria (42) invariant, we only need to optimize over *one* angle ϑ. In contrast, criteria based on marginal distributions require angle tomography over *both* angles ϑ_1, ϑ_2, which is substantially more costly in terms of experimental runs.

D. Concave function f

Following the analysis in Sec. II D, our entanglement criteria (42) state that all distributions $Q_\pm$ corresponding to separable states are less localized than the vacuum distribution $\bar{Q}'_\pm$. In contrast, some entangled states have sufficiently localized distributions $Q_\pm$, such that they are detected by the criteria (42) for some f. The corresponding distribution $Q_\pm$ majorizes the vacuum distribution $\bar{Q}'_\pm$ if and only if entanglement can be certified for all f.

Using the definition of the density-level function (24), we can rewrite the criteria as

$$\mathcal{W}_f = \int_{\mathcal{J}} dt\, f(t)[\mu_{Q_\pm}(t) - \mu_{\bar{Q}'_\pm}(t)] \geqslant 0, \qquad (47)$$

showing that also our criteria (42) reduce to a comparison of the density-level functions $\mu_{Q_\pm}(t)$ and $\mu_{\bar{Q}'_\pm}(t)$ and an appropriate choice for $f(t)$. Note that the vacuum expression (40) is bounded from above by $(a_1b_1 + a_2b_2)^{-1}$, such that

the second term in (47) is effectively integrated over the interval $[0, (a_1b_1 + a_2b_2)^{-1}]$. However, some entangled states have sufficiently localized $Q_\pm$ such that their density-level functions take positive values for $t > (a_1b_1 + a_2b_2)^{-1}$.

As a simple example, we consider the two-mode squeezed vacuum (TMSV) state

$$|\psi\rangle = \sqrt{1 - \lambda^2} \sum_{n=0}^{\infty} (-\lambda)^n |n\rangle \otimes |n\rangle, \qquad (48)$$

with $\lambda \in [0, 1]$ being the squeezing parameter, which is entangled for all $\lambda > 0$. For $a_1 = b_1 = a_2 = b_2 = 1$, $\vartheta_1 = \vartheta_2 = 0$, and $\xi = 1$ we obtain a Gaussian distribution with covariance matrix $V_\pm = \frac{2}{1 \pm \lambda}\mathbb{1}$. In Fig. 2(a) we show the vacuum case $\lambda = 0$ (for which $Q_+ = Q_-$), while in Figs. 2(b) and 2(c) we illustrate $\lambda = 1/3$ for Q_+ and Q_-, respectively. Entanglement is only detected in the (r_+, s_-) variables as the corresponding distribution is more localized than the vacuum. The rotational symmetry of these distributions permits analytic calculations of the level and density-level functions when working with polar coordinates, which leads to

$$m_{Q_\pm}(t) = -\det{}^{1/2}V_\pm \ln(t\det{}^{1/2}V_\pm),$$
$$\mu_{Q_\pm}(t) = \frac{\det{}^{1/2}V_\pm}{t}, \qquad (49)$$

for $t \in [0, \det{}^{-1/2}V_\pm]$, respectively. Both are shown for both variable pairs in Fig. 2(d) for the vacuum (black curves), $\lambda = 1/3$ (dark orange curves) and $\lambda = 2/3$ (light orange curves) with solid (dashed) and dot-dashed (dotted) curves for the $+(-)$ level and density-level functions, respectively. The values at the plot markers in t steps of $1/12$ correspond to the phase-space measures, i.e., areas divided by 2π, of the circles in Figs. 2(a)–2(c). Furthermore, we show the integrand of our

witness functional (47) for the distributions of Figs. 2(b) and 2(c) in Figs. 2(e) and 2(f), respectively, for various choices for the concave function $f(t)$, such that the sum of the shaded areas corresponds to the value of the witness functional (47). The black solid curves express normalization, while the blue dashed and red dotted curves illustrate that the absolute value of the witness (47) increases for smaller monomial exponents for the state (48). However, entanglement is detected in Q_+ for all choices of f as the state is Gaussian, which we discuss further in Sec. III F 2.

E. Entropic criteria

Of particular interest are entropic functionals as there exist a variety of criteria formulated in terms of entropies of measurement distributions [16,17,20,22,36,93–97]. In the following, we discuss entropic criteria as special choices for the concave function f. To arrive at entropic criteria, it is additionally necessary to apply a monotonic function to both terms in the witness functional (41). More precisely, we consider a monotonically increasing function $g : \mathbb{R} \to \mathbb{R}$ such that the witness functional becomes

$$\mathcal{W}_{f,g} = g\left[\int \frac{dr_\pm ds_\mp}{2\pi} f(Q_\pm) \right] - g\left[\int \frac{dr_\pm ds_\mp}{2\pi} f(\bar{Q}'_\pm) \right],$$
(50)

which still fulfills the entanglement criteria (42). As the converse statement $\mathcal{W}_f \leqslant 0$ holds for $-f$ (or equivalently for convex f), applying a monotonically *decreasing* function g still results in $\mathcal{W}_{f,g} \geqslant 0$. We note that this procedure does neither strengthen nor weaken the entanglement criteria.

1. Rényi-Wehrl entropies

We start with the family of Rényi-Wehrl entropies

$$S_\beta(Q_\pm) = \frac{1}{1-\beta} \ln \left[\int \frac{dr_\pm ds_\mp}{2\pi} Q_\pm^\beta(r_\pm, s_\mp) \right],$$
(51)

with entropic orders $\beta \in (0, 1) \cup (1, \infty)$. Hence, we choose $f(t) = t^\beta$, which is concave (convex) and a monotonically increasing (decreasing) function $g(t) = \frac{1}{1-\beta} \ln t$ for $\beta < 1$ ($\beta > 1$), in which case (50) implies

$$\mathcal{W}_\beta = S_\beta(Q_\pm) - \frac{\ln \beta}{\beta - 1} - \frac{\ln \det \bar{V}'_\pm}{2} \geqslant 0,$$
(52)

where $\bar{V}'_\pm = (a_1 b_1 + a_2 b_2)\mathbb{1}$ denotes the covariance matrix of the vacuum $\bar{Q}'_\pm$, for all separable states ρ_{12}.

2. Tsallis-Wehrl entropies

For Tsallis-Wehrl entropies

$$S_\gamma(Q_\pm) = \frac{1}{\gamma - 1}\left[1 - \int \frac{dr_\pm ds_\mp}{2\pi} Q_\pm^\gamma(r_\pm, s_\mp) \right],$$
(53)

with $\gamma \in (0, 1) \cup (1, \infty)$[4] we choose again monomials $f(t) = t^\gamma$, but now we take $g(t) = \frac{1-t}{\gamma - 1}$, which is also monotonically

———————

[4]For $\gamma < 0$ we have $f(0) \to \infty$ and hence we have to restrict to Tsallis-Wehrl entropies with $\gamma > 0$.

increasing (decreasing) for $\gamma < 1$ ($\gamma > 1$), leading to the criteria

$$\mathcal{W}_\gamma = S_\gamma(Q_\pm) - \frac{1}{\gamma - 1}\left[1 - \frac{(\det \bar{V}'_\pm)^{\frac{1-\gamma}{2}}}{\gamma} \right] \geqslant 0.$$
(54)

As the function f agrees for both entropic families, the criteria (52) and (54) are equally strong in the sense that they detect the same set of entangled states. Therefore, an appropriate choice regarding which entropic family is being considered can be based on other decisive factors depending on the application: While Tsallis' entropy was designed to generalize statistical mechanics to situations where an entropy is nonextensive, Rényi's aim was to formulate a family of entropies which includes not only Shannon's but also Hartley's and other entropies, which are of particular interest in the context of communication protocols.

3. Wehrl entropy

In the limits $\beta, \gamma \to 1$ we obtain from (52) and (54) the entropic witness

$$\mathcal{W}_1 = S_1(Q_\pm) - 1 - \frac{\ln \det \bar{V}'_\pm}{2} \geqslant 0$$
(55)

in terms of the Wehrl entropy

$$S_1(Q_\pm) = -\int \frac{dr_\pm ds_\mp}{2\pi} Q_\pm(r_\pm, s_\mp) \ln Q_\pm(r_\pm, s_\mp).$$
(56)

Note that this result also follows directly from (41) for the choice $f(t) = -t \ln t$. Note also that for $a_1 = b_1 = a_2 = b_2 = 1$ the corresponding criteria reduce to the criteria reported in Ref. [36].

F. Second moment criteria

Another interesting class of witnesses comprises second moments, which are the simplest quantities revealing information about the localization of a distribution. We derive second moment criteria from our general criteria and discuss them in the context of Gaussian states.

1. Determinant of the covariance matrix

We start from the nonlocal covariance matrix $\gamma_\pm$ of an arbitrary Wigner W distribution

$$\gamma_\pm = \begin{pmatrix} \sigma_{r_\pm}^2 & \sigma_{r_\pm s_\mp} \\ \sigma_{r_\pm s_\mp} & \sigma_{s_\mp}^2 \end{pmatrix},$$
(57)

which contains all three second moments, i.e., the two variances

$$\sigma_{r_\pm}^2 = \int \frac{dr_\pm ds_\mp}{2\pi} r_\pm^2 W_\pm(r_\pm, s_\mp),$$

$$\sigma_{s_\mp}^2 = \int \frac{dr_\pm ds_\mp}{2\pi} s_\mp^2 W_\pm(r_\pm, s_\mp),$$
(58)

as well as the covariance

$$\sigma_{r_\pm s_\mp} = \int \frac{dr_\pm ds_\mp}{2\pi} r_\pm s_\mp W_\pm(r_\pm, s_\mp),$$
(59)

which characterizes the correlations between $r_\pm$ and $s_\mp$. Note here that we have assumed zero expectation values without

loss of generality as the entanglement criteria (42) and all marginal criteria we consider in Sec. IV are invariant under displacements.

Moreover, we write for the covariance matrix $V_\pm$ of the Husimi Q distribution

$$V_\pm = \begin{pmatrix} \Sigma_{r_\pm}^2 & \Sigma_{r_\pm s_\mp} \\ \Sigma_{r_\pm s_\mp} & \Sigma_{s_\mp}^2 \end{pmatrix}, \tag{60}$$

with second moments

$$\Sigma_{r_\pm}^2 = \int \frac{dr_\pm ds_\mp}{2\pi} r_\pm^2 Q_\pm(r_\pm, s_\mp),$$

$$\Sigma_{s_\mp}^2 = \int \frac{dr_\pm ds_\mp}{2\pi} s_\mp^2 Q_\pm(r_\pm, s_\mp),$$

$$\Sigma_{r_\pm s_\mp} = \int \frac{dr_\pm ds_\mp}{2\pi} r_\pm s_\mp Q_\pm(r_\pm, s_\mp). \tag{61}$$

Using (13) adapted to the nonlocal variables, the latter covariance matrix is related to the former via

$$V_\pm = \gamma_\pm + \bar{\gamma}_\pm = \begin{pmatrix} \sigma_{r_\pm}^2 + \frac{a_1^2+a_2^2}{2} & \sigma_{r_\pm s_\mp} \\ \sigma_{r_\pm s_\mp} & \sigma_{s_\mp}^2 + \frac{b_1^2+b_2^2}{2} \end{pmatrix}, \tag{62}$$

which shows that the variances of the Husimi Q distribution are strictly larger than those of the Wigner W distribution, while the covariances agree. Note here that $\bar{\gamma}_\pm$ is obtained when marginalizing the Wigner W distribution of the global vacuum after transforming to nonlocal variables.

To derive second moment criteria, we use the fact that the Wehrl entropy $S_1(Q_\pm)$ is maximized by a Gaussian Husimi Q distribution $Q_\pm$ for a fixed covariance matrix $V_\pm$, i.e.,

$$S_1(Q_\pm) \leqslant 1 + \frac{1}{2} \ln \det V_\pm. \tag{63}$$

With this upper bound the Wehrl entropic witness functional (55) reduces to the second moment witness

$$\mathcal{W}_{\det V_\pm} = \det V_\pm - \det \bar{V}'_\pm \geqslant 0, \tag{64}$$

which is based on the determinant of the covariance matrix $V_\pm$ and thus contains information about all three second moments, in particular about the correlations between $r_\pm$ and $s_\mp$.

2. Gaussian states

Gaussian states are characterized by Gaussian Husimi Q distributions, which are of the form

$$Q_\pm(r_\pm, s_\mp) = \frac{1}{Z} e^{-\frac{1}{2}(r_\pm, s_\mp)^T V_\pm^{-1}(r_\pm, s_\mp)}, \tag{65}$$

with $Z = \det^{1/2} V_\pm$ being a normalization constant. We have again set the expectation values to zero without loss of generality.

Interestingly, for the class of Gaussian distributions (65), the general entanglement criteria (42) are *equivalent* to the second moment criteria $\mathcal{W}_{\det V_\pm} \geqslant 0$ for *all* concave f with $f(0) = 0$. To prove this equivalence, we start from Simon's normal form of the bipartite Wigner covariance matrix [13] (standard form II in Ref. [10]):

$$\gamma_{12} = \begin{pmatrix} m_1 & 0 & m_+ & 0 \\ 0 & m_1 & 0 & m_- \\ m_+ & 0 & m_2 & 0 \\ 0 & m_- & 0 & m_2 \end{pmatrix}, \tag{66}$$

with $m_1, m_2, m_+, m_- \in \mathbb{R}$ (note that these parameters are also constrained by the uncertainty principle formulated for γ_{12}). Every bipartite covariance matrix γ_{12} can be brought into this form by local, single-mode symplectic transformations $\mathcal{S}_1 \otimes \mathcal{S}_2$ with $\mathcal{S}_1, \mathcal{S}_2 \in Sp(2, \mathbb{R})$, which does not alter the separability of the state as such transformations correspond to the class of local operations and classical communications (LOCCs). After adding the global vacuum $\bar{\gamma}_{12}$ to obtain V_{12}, transforming to the nonlocal variables (32) and integrating out the mixed variable pairs according to (35), we find the diagonal matrix

$$V_\pm = \frac{1}{2} \begin{pmatrix} a_1^2(1+2m_1) + a_2^2(1+2m_2) \pm 4a_1 a_2 m_+ & 0 \\ 0 & b_1^2(1+2m_1) + b_2^2(1+2m_2) \mp 4b_1 b_2 m_- \end{pmatrix}. \tag{67}$$

Starting from the vacuum $\bar{Q}'_\pm(r_\pm, s_\mp)$, a general Gaussian distribution $Q_\pm(\tilde{r}_\pm, \tilde{s}_\mp)$ with a covariance matrix of the form (67) can be obtained by applying an affine linear coordinate transformation in phase space

$$(r_\pm, s_\mp) \to (\tilde{r}_\pm, \tilde{s}_\mp) = M_\pm(r_\pm, s_\mp). \tag{68}$$

Comparing (67) with (40) shows that the transformation matrix $M_\pm$ is diagonal and reads

$$M_\pm = \sqrt{V_\pm \bar{V}'^{-1}_\pm}, \tag{69}$$

which is always well defined because all involved components are positive and real by definition. Indeed, with the transfor-

mation (68) we obtain for the quadratic form in the exponent

$$(r_\pm, s_\mp)^T \bar{V}_\pm^{-1}(r_\pm, s_\mp) \to (\tilde{r}_\pm, \tilde{s}_\mp)^T V_\pm^{-1}(\tilde{r}_\pm, \tilde{s}_\mp)$$
$$= (r_\pm, s_\mp)^T M_\pm^T V_\pm^{-1} M_\pm (r_\pm, s_\mp)$$
$$= (r_\pm, s_\mp)^T \bar{V}_\pm^{-1}(r_\pm, s_\mp). \tag{70}$$

Furthermore, (69) implies

$$\det M_\pm^2 = \frac{\det V_\pm}{\det \bar{V}'_\pm}, \tag{71}$$

such that we may obtain a relation between $\det V_\pm$ and $\det \bar{V}'_\pm$ from $\det M_\pm^2$.

Under the coordinate transformation (68), the vacuum distribution $\bar{Q}'_\pm(r_\pm, s_\mp)$ transforms as

$$\bar{Q}'_\pm(r_\pm, s_\mp) \rightarrow Q_\pm(\tilde{r}_\pm, \tilde{s}_\mp) = \frac{\bar{Q}'_\pm(r_\pm, s_\mp)}{\det M_\pm}. \qquad (72)$$

Similarly, the level functions transform as

$$m_{\bar{Q}'_\pm}(t) \rightarrow m_{Q_\pm}(t) = \det M_\pm m_{\bar{Q}'_\pm}(t \det M_\pm), \qquad (73)$$

while for the density-level functions we find

$$\mu_{\bar{Q}'_\pm}(t) \rightarrow \mu_{Q_\pm}(t) = \det M_\pm^2 \mu_{\bar{Q}'_\pm}(t \det M_\pm). \qquad (74)$$

This shows that the level as well as the density-level functions of arbitrary Gaussian distributions are scaled versions of each other, which is also evident in (49) and Fig. 2(d).

Plugging the relation (72) into our witness functional (41) for the first or second term yields

$$\mathcal{W}_f = \int \frac{dr_\pm ds_\mp}{2\pi} \left[\det M_\pm f\left(\frac{\bar{Q}'_\pm}{\det M_\pm} \right) - f(\bar{Q}'_\pm) \right] \qquad (75)$$

or

$$\mathcal{W}_f = \int \frac{d\tilde{r}_\pm d\tilde{s}_\mp}{2\pi} \left[f(Q_\pm) - \frac{f(Q_\pm \det M_\pm)}{\det M_\pm} \right], \qquad (76)$$

respectively. As f is concave with $f(0) = 0$, it is a subadditive function and hence fulfills $f(\kappa t) \geqslant \kappa f(t)$ for all real $\kappa \in [0, 1]$ and all $t \in \mathcal{J}$. Thus, $\det M_\pm \geqslant 1$ implies $\mathcal{W}_f \geqslant 0$ from (75), while $\det M_\pm \leqslant 1$ implies $\mathcal{W}_f \leqslant 0$ from (76). As the latter is equivalent to $\mathcal{W}_f \geqslant 0$ implying $\det M_\pm \geqslant 1$ by contraposition, we can conclude with (71) that the second moment criteria $\mathcal{W}_{\det V_\pm} = \det V_\pm - \det \bar{V}'_\pm \geqslant 0$ and the general criteria $\mathcal{W}_f \geqslant 0$ are equivalent for all Gaussian states and for *all f* under our standard assumptions.

3. Optimality

The considerations of the preceding section have strong implications regarding the optimality of the second moment witness (64): It is optimal in the sense that no stronger state-independent bound on $\det V_\pm$ can be implied from the general criteria (42) as this would be in contradiction with the latter equivalence in case of Gaussian distributions. More precisely, if there existed such a stronger bound, it would have to hold for all states, in particular for Gaussian states. But since the general criteria (42) are equivalent to the second moment criteria from (64) for Gaussian states, there cannot be any such bound. Therefore, taking $f(t) \neq -t \ln t$ and maximizing the witness functional $\mathcal{W}_f$ over $Q_\pm$ for fixed $V_\pm$ cannot lead to stronger second moment criteria.

We illustrate the aforementioned optimality by showing the inferiority of second moment criteria stemming from maximizing Rényi-Wehrl entropies $S_\beta(Q_\pm)$ for fixed $V_\pm$. Depending on the entropic order $\beta > \frac{1}{2}$, Rényi-Wehrl entropies are maximized by the distributions [98,99]

$$Q_\pm(r_\pm, s_\mp)$$

$$= Z(\beta) \left(1 - \frac{1}{2} \frac{\beta - 1}{2\beta - 1} (r_\pm, s_\mp)^T V_\pm^{-1} (r_\pm, s_\mp) \right)_+^{\frac{1}{\beta - 1}}, \quad (77)$$

with $x_+ = \max(x, 0)$ as the support for $\beta > 1$ is defined as the compact domain for which $(r_\pm, s_\mp)^T V_\pm^{-1}(r_\pm, s_\mp) \leqslant \nu$ [for
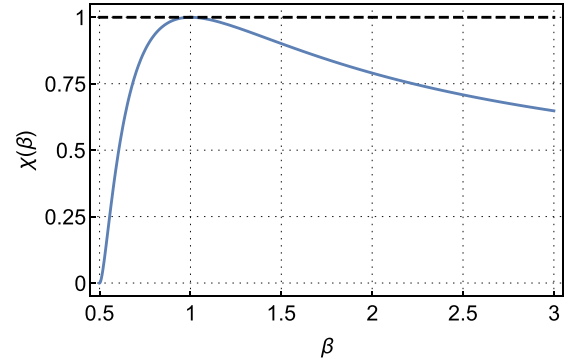

FIG. 3. Additional factor $\chi(\beta)$ in the second moment criteria (79) obtained from maximizing the left-hand side of the Rényi-Wehrl criteria (52) for fixed covariance matrix $V_\pm$ (blue solid curve). We observe $\chi(\beta) \leqslant 1$ with equality for $\beta \rightarrow 1$ (see black dashed line), illustrating the optimality of the main second moment criteria (64).

larger values of $(r_\pm, s_\mp)$ the distribution $Q_\pm$ would become negative], a normalization factor

$$Z(\beta) = \begin{cases} \frac{\Gamma\left(\frac{1}{1-\beta}\right)}{\Gamma\left(\frac{\beta}{1-\beta}\right)} \frac{1-\beta}{2\beta-1} \det^{-1/2} V_\pm & \frac{1}{2} < \beta < 1 \\ \frac{\Gamma\left(\frac{\beta}{\beta-1}+1\right)}{\Gamma\left(\frac{\beta}{\beta-1}\right)} \frac{\beta-1}{2\beta-1} \det^{-1/2} V_\pm & \beta > 1, \end{cases} \qquad (78)$$

and $\Gamma(x)$ denoting the Γ function. These correspond to two-dimensional Student-t and Student-r distributions, for $\frac{1}{2} < \beta < 1$ with scale matrix $\Sigma_\pm = (\nu - 2)V_\pm$ and $\nu = \frac{2\beta}{1-\beta}$ degrees of freedom and for $\beta > 1$ with scale matrix $\Sigma_\pm = \nu V_\pm$ and $\nu = 2\frac{2\beta-1}{\beta-1}$ degrees of freedom, respectively.

Bounding the left-hand side of the Rényi-Wehrl witness (52) from above by the Rényi-Wehrl entropy of the distributions (77) and simplifying the result leads to the second moment criteria

$$\tilde{\mathcal{W}}_{\det V_\pm} = \det V_\pm - \chi(\beta) \det \bar{V}'_\pm \geqslant 0, \qquad (79)$$

with the non-negative function

$$\chi(\beta) = \left(2 - \frac{1}{\beta} \right)^{-\frac{2\beta}{\beta-1}} \beta^{\frac{2}{\beta-1}}, \qquad (80)$$

which is shown in Fig. 3. Given that $\chi(\beta) \leqslant 1$ for all $\beta \in (\frac{1}{2}, 1) \cup (1, \infty)$ with equality in the limit $\beta \rightarrow 1$ [in which case the distributions (77) converge to Gaussian distributions (65)], the criteria (79) are indeed never stronger than (64).

IV. COMPARISON WITH MARGINAL APPROACH

Having discussed various families of entanglement criteria following from our general criteria (42), we now compare the performance of these criteria to their counterparts based on marginal distributions.

A. Second moment criteria

There exists a simple relation between our second moment witness (64) and the MGVT criteria [11,12]. Namely, using that $\det \bar{V}'_\pm = 4 \det \bar{\gamma}'_\pm$ and multiplying out the determinant
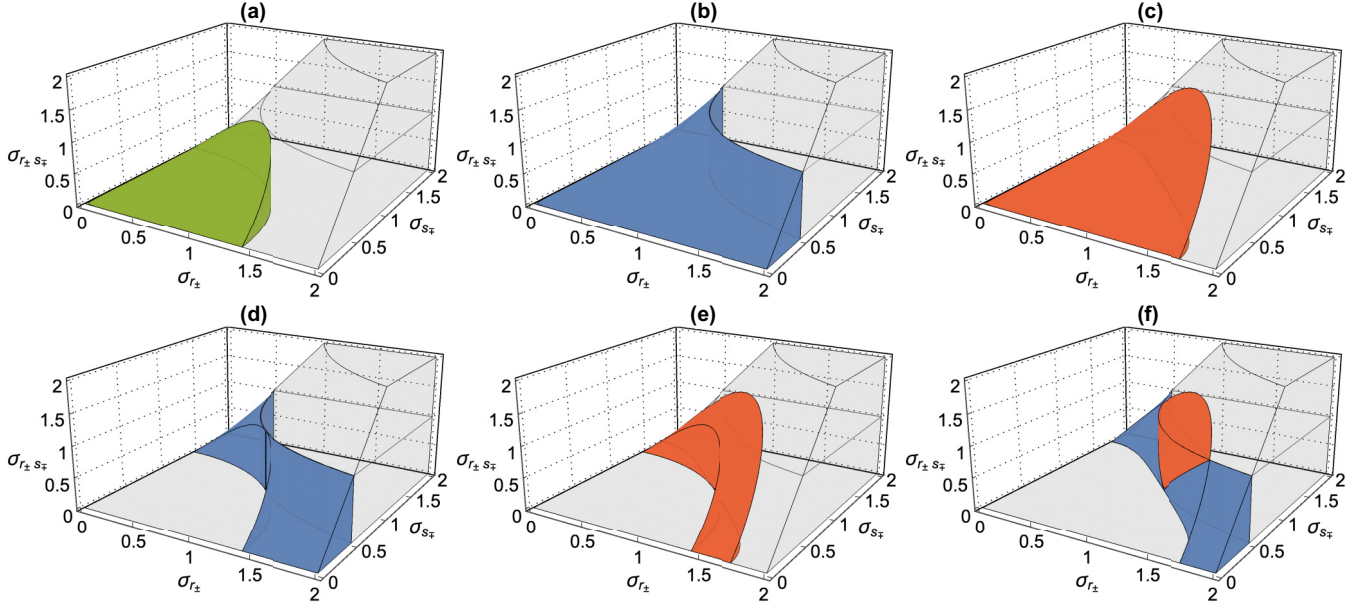
FIG. 4. (upper row) Witnessed regions by the DGCZ, MGVT, and our second moment criteria are shown in panels (a)–(c), respectively, with the light gray regions indicating the allowed regions from non-negativity of the covariance matrix $\gamma_\pm$, i.e., $\sigma_{r_\pm}\sigma_{s_\mp} \geqslant \sigma_{r_\pm s_\mp}$ (overall, we restrict to $\sigma_{r_\pm s_\mp} \geqslant 0$ for simplicity). (lower row) Comparisons of witnessed regions between two criteria. Panels (d) and (e) show the outperformance regions of the MGVT and our criteria, respectively, over the DGCZ criteria, exemplifying that the former two are strictly stronger conditions than the latter. Our criteria are compared with the MGVT criteria in panel (f), highlighting their abilities to account for improper angles and squeezings, respectively. A combined figure is provided in Ref. [74].

with (62) we find

$$\mathcal{W}_{\det V_\pm} = \mathcal{W}_{\text{MGVT}} + \frac{b_1^2 + b_2^2}{2}\sigma_{r_\pm}^2 + \frac{a_1^2 + a_2^2}{2}\sigma_{s_\mp}^2$$
$$+ \frac{(a_1^2 + a_2^2)(b_1^2 + b_2^2) - 3(a_1 b_1 + a_2 b_2)^2}{4} - \sigma_{r_\pm s_\mp}^2, \tag{81}$$

where

$$\mathcal{W}_{\text{MGVT}} = \sigma_{r_\pm}^2 \sigma_{s_\mp}^2 - \frac{\det \bar{V}_\pm'}{4} \tag{82}$$

is the generalized witness functional for the MGVT criteria. Another interesting set of criteria, which contain a sum rather than a product of variances, are the DGCZ criteria [10], expressed through the witness functional

$$\mathcal{W}_{\text{DGCZ}} = \sigma_{r_\pm}^2 + \sigma_{s_\mp}^2 - \det^{1/2}\bar{V}_\pm'. \tag{83}$$

We show the witnessed regions of the DGCZ, MGVT, and our second moment criteria in Figs. 4(a)–4(c), respectively, together with the physically allowed regions by the non-negativity of $\gamma_\pm$ (light gray shaded areas) for $a_1 = b_1 = a_2 = b_2 = 1$, $\vartheta_1 = \vartheta_2 = 0$, and $\xi = 1$.

It is well known that the MGVT criteria imply the DGCZ criteria, which follows from $x^2 y^2 \leqslant \frac{1}{4}(x^2 + y^2)^2$ for all $x, y \in \mathbb{R}$, see Fig. 4(d). Similarly, our second moment criteria (64) imply the DGCZ criteria (83) for all scaling parameters a_1, b_1, a_2, b_2 due to $\sigma_{r_\pm s_\mp}^2 \geqslant 0$ as well, which we exemplify in Fig. 4(e). As the DGCZ criteria are necessary and sufficient for separability in case of Gaussian states when optimized over the scaling parameters a_1, b_1, a_2, b_2, the same holds true for the MGVT and our criteria.

Furthermore, our criteria and the MGVT criteria are equivalent after optimizing both criteria over the set of symplectic transformations under which they are *not* invariant together with an optimization over the scaling parameters a_1, b_1, a_2, b_2. To that end, we consider symplectic transformations $\mathcal{S}_\pm \in Sp(2, \mathbb{R})$ in the nonlocal phase spaces $(r_\pm, s_\mp)$ transforming the Wigner W distribution as $W_\pm(r_\pm, s_\mp) \to \tilde{W}_\pm(\tilde{r}_\pm, \tilde{s}_\mp)$ with $(\tilde{r}_\pm, \tilde{s}_\mp) = \mathcal{S}_\pm(r_\pm, s_\mp)$. Then, the nonlocal covariance matrix transforms as $\gamma_\pm \to \tilde{\gamma}_\pm = \mathcal{S}_\pm \gamma \mathcal{S}_\pm^T$ with $\det \mathcal{S}_\pm = 1$ [13]. This implies for our second moment witness functional (64)

$$\mathcal{W}_{\det V_\pm} \to \tilde{\mathcal{W}}_{\det V_\pm}$$
$$= \det \tilde{V}_\pm - \det \bar{V}_\pm'$$
$$= \det(\mathcal{S}_\pm \gamma_\pm \mathcal{S}_\pm^T + \bar{\gamma}_\pm) - \det \bar{V}_\pm'$$
$$= \det[\gamma_\pm + \bar{\gamma}_\pm(\mathcal{S}_\pm^T \mathcal{S}_\pm)^{-1}] - \det \bar{V}_\pm', \tag{84}$$

showing explicitly that the criteria remain invariant under orthogonal transformations $\mathcal{S}_\pm \mathcal{S}_\pm^T = \mathbb{1}$, i.e., rotations, but not under squeezing for which $\mathcal{S}_\pm = \Xi = \text{diag}(\xi, 1/\xi)$ with $\xi > 0$. In contrast, the MGVT criteria (82) are invariant under squeezing only, in which case

$$\mathcal{W}_{\text{MGVT}} \to \tilde{\mathcal{W}}_{\text{MGVT}} = \xi^2 \sigma_{r_\pm}^2 \frac{1}{\xi^2} \sigma_{s_\mp}^2 - \frac{\det \bar{V}_\pm'}{4}$$
$$= \mathcal{W}_{\text{MGVT}}. \tag{85}$$

Note that the DGCZ criteria (83) are not invariant under either of the two transformations.

Let us now apply a squeezing transformation to our second moment criteria (64) in the sense of (84), which yields

$$
\mathcal{W}_{\det V_\pm} = \left(\xi^2 \sigma_{r_\pm}^2 + \frac{a_1^2 + a_2^2}{2} \right) \left(\frac{1}{\xi^2} \sigma_{s_\mp}^2 + \frac{b_1^2 + b_2^2}{2} \right) - \sigma_{r_\pm s_\mp}^2 - (a_1 b_1 + a_2 b_2)^2. \tag{86}
$$

To optimize over ξ, we minimize the latter expression with respect to ξ, which gives

$$
\xi^2 = \frac{\sigma_{s_\mp}}{\sigma_{r_\pm}} \sqrt{\frac{a_1^2 + a_2^2}{b_1^2 + b_2^2}}, \tag{87}
$$

leading to

$$
\mathcal{W}_{\det V_\pm} = \left(\sigma_{r_\pm} \sigma_{s_\mp} + \frac{\sqrt{(a_1^2 + a_2^2)(b_1^2 + b_2^2)}}{2} \right)^2 - \sigma_{r_\pm s_\mp}^2 - (a_1 b_1 + a_2 b_2)^2. \tag{88}
$$

The MGVT criteria become optimal for an appropriate choice of the coordinate axes such that $\sigma_{r_\pm s_\mp} = 0$, and then the non-negativity of the latter is equivalent to

$$
\sigma_{r_\pm} \sigma_{s_\mp} + \frac{\sqrt{(a_1^2 + a_2^2)(b_1^2 + b_2^2)}}{2} - (a_1 b_1 + a_2 b_2) \geqslant 0. \tag{89}
$$

Finally, noting that $[(a_1^2 + a_2^2)(b_1^2 + b_2^2)]^{1/2} \geqslant a_1 b_1 + a_2 b_2$ with equality for $b_2 = (b_1 a_2)/a_1$ shows that the statements $\mathcal{W}_{\det V_\pm} \geqslant 0$ and $\mathcal{W}_{\mathrm{MGVT}} \geqslant 0$ are indeed equivalent after optimizing over the angle ϕ, the squeezing ξ as well as the scaling parameters a_1, b_1, a_2, b_2.[5]

The different invariances of our criteria (64) and the MGVT criteria (82) cause differences in their performances for fixed a_1, b_1, a_2, b_2, depending on whether the considered Husimi Q distribution is sufficiently rotated or squeezed, see Fig. 4(f). Furthermore, (81) provides an intuition for one set of criteria to outperform the other. Roughly speaking, our criteria outperform (are outperformed by) the MGVT criteria for $\sigma_{r_\pm} \approx \sigma_{s_\mp}$ and $|\sigma_{r_\pm s_\mp}| > 0$ (for $\sigma_{r_\pm} \approx 1/\sigma_{s_\mp}$ and $|\sigma_{r_\pm s_\mp}| \approx 0$).

B. Entropic criteria

Although there is no direct analog to the relation (81), we can derive an inequality between our Wehrl entropic criteria (55) and the marginal-based entropic criteria put forward by WTSTD in Ref. [16], which will provide an intuition for the later comparison.

We start by decomposing the Wehrl entropy (56) in the sense of a joint entropy

$$
S(Q_\pm) = S(F_\pm) + S(G_\mp) - I(F_\pm : G_\mp), \tag{90}
$$

where the marginals of the Husimi Q distribution read

$$
F_\pm \equiv F_\pm(r_\pm) = \int \frac{ds_\mp}{\sqrt{2\pi}} Q_\pm,
$$

$$
G_\mp \equiv G_\mp(s_\mp) = \int \frac{dr_\pm}{\sqrt{2\pi}} Q_\pm. \tag{91}
$$

---

[5]Note that $\mathcal{W}_{\det V_\pm}$ and $\mathcal{W}_{\mathrm{MGVT}}$ themselves are still unequal.

Following (13), these distributions correspond to smeared-out versions of the true marginal distributions associated with measurements of the operators $\boldsymbol{R}_\pm$ or $\boldsymbol{S}_\mp$, i.e., $f_\pm(r_\pm) = \langle r_\pm | \boldsymbol{\rho} | r_\pm \rangle$ or $g_\mp(s_\mp) = \langle s_\mp | \boldsymbol{\rho} | s_\mp \rangle$, respectively, via

$$
F_\pm(r_\pm) = (f_\pm * \bar{f}_\pm)(r_\pm)
$$

$$
= \sqrt{2\pi} \int dr'_\pm f_\pm(r'_\pm) \bar{f}_\pm(r_\pm - r'_\pm), \tag{92}
$$

with $f_\pm(r_\pm)$ being normalized with respect to $dr_\pm/\sqrt{2\pi}$, and similarly for $g_\mp(s_\mp)$.

Analogous to the covariance $\sigma_{r_\pm s_\mp}$, we have a term in the decomposition (90) accounting for the correlations between $r_\pm$ and $s_\mp$ in (90), which is the mutual information

$$
I(F_\pm : G_\mp) = \int \frac{dr_\pm ds_\mp}{2\pi} Q_\pm \ln \frac{Q_\pm}{F_\pm G_\mp}. \tag{93}
$$

Note that the mutual information is a non-negative measure for the total correlations which is zero if and only if $r_\pm$ and $s_\mp$ are uncorrelated. In contrast, a vanishing covariance $\sigma_{r_\pm s_\mp}$ does not allow us to exclude the presence of correlations.

To relate the differential entropies of $f_\pm$ and $F_\pm$ we employ the entropy power inequality

$$
e^{2S(F_\pm/\sqrt{2\pi})} \geqslant e^{2S(f_\pm)} + e^{2S(\bar{f}_\pm)}, \tag{94}
$$

use that the vacuum entropy sum is bounded from below [51,88,89]

$$
S(\bar{f}_\pm) + S(\bar{g}_\mp) = 1 + \ln \pi + \frac{1}{2} \ln \left[(a_1^2 + a_2^2)(b_1^2 + b_2^2) \right]
$$

$$
\geqslant 1 + \ln \pi + \frac{\ln \det \bar{V}_\pm}{2}, \tag{95}
$$

and utilize midpoint concavity of the logarithm, i.e., that $\ln[(x + y)/2] \geqslant (\ln x + \ln y)/2$, to arrive at

$$
\mathcal{W}_1 \geqslant \frac{1}{2} \mathcal{W}_{\mathrm{WTSTD}} - I(F_\pm : G_\mp), \tag{96}
$$

where we introduced the witness functional corresponding to the entropic criteria in Ref. [16] as

$$
\mathcal{W}_{\mathrm{WTSTD}} = S(f_\pm) + S(g_\mp) - 1 - \ln \pi - \frac{\ln \det \bar{V}'_\pm}{2}. \tag{97}
$$

The inequality (96) can be considered the entropic analog of (81). Hence, a necessary (but not sufficient) condition for an entangled state, which is undetected by the WTSTD criteria (97), to be detected by our entropic criteria (55) is that sufficiently strong correlations are present, i.e., that $I(F_\pm : G_\mp) > \mathcal{W}_{\mathrm{WTSTD}} \geqslant 0$ [see Sec. IV C, especially Fig. 7(b)].

Unfortunately, a similar relation cannot be derived for the Rényi-Wehrl criteria (52) and the Rényi entropic criteria by STW, which are given by the witness functional

$$
\mathcal{W}_{\mathrm{STW}} = S_\alpha(f_\pm) + S_\beta(g_\mp) - \frac{\ln \det \bar{V}'_\pm}{2}
$$

$$
+ \frac{1}{2(1-\beta)} \ln \frac{\beta}{\pi} + \frac{1}{2(1-\alpha)} \ln \frac{\alpha}{\pi}, \tag{98}
$$

with the condition

$$
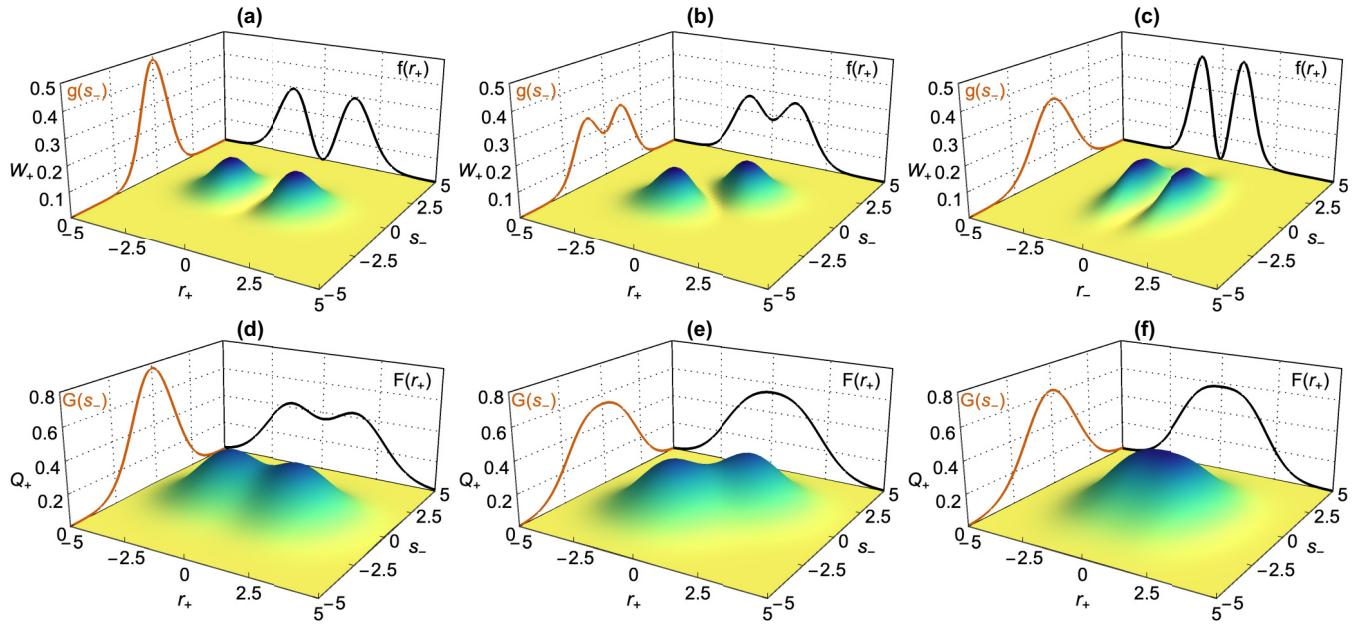\frac{1}{\alpha} + \frac{1}{\beta} = 2, \tag{99}
$$

FIG. 5. The upper and lower rows show the Wigner W and Husimi Q distributions (A1) in the (r_+, s_-) variables, respectively, together with their marginal distributions for $a_1 = b_1 = a_2 = b_2 = 1$, $\sigma_+ = 1$, $\sigma_- = 1.2$ and various choices for the angle ϕ and the squeezing parameter ξ. Panels (a) and (d) represent the case $\phi = 0$, $\xi = 1$. The rotation angle is set to $\phi = \pi/4$ in panels (b) and (e), leading to strong correlations between r_+ and s_-, while in panels (c) and (f) the squeezing parameter is changed to $\xi = 3/2$.

which can be traced back to the Babenko-Beckner inequality [100] appearing in the proof of the corresponding entropic uncertainty relation [88,101]. A decomposition like (90) in terms of Rényi-type entropies would require a possibly negative third term, which does not coincide with the Rényi generalization of the mutual information (93) (an inherently non-negative quantity by definition) in general.

However, there is a crucial difference between the Rényi-type criteria (52) and (98). While our criteria (52) can be optimized over *all* entropic orders β, the marginal based criteria (98) are constrained by the condition (99). More precisely, one marginal distribution is always exponentiated by a real number smaller or equal to one, while the other is always exponentiated by a real number larger or equal to one. As we will see in the following, this is a severe drawback as some entangled states require an optimization where the full phase-space distribution is exponentiated with a small or a large number.

C. Example state

Following Refs. [16,17,25,102–104], we consider the family of pure states corresponding to the global wave function in position space

$$\psi(r_1, r_2) = \frac{r_1 + r_2}{\sqrt{\pi \sigma_- \sigma_+^3}} \exp\left\{-\frac{1}{4}\left[\left(\frac{r_1 + r_2}{\sigma_+}\right)^2 + \left(\frac{r_1 - r_2}{\sigma_-}\right)^2\right]\right\},$$
(100)

which is entangled for all positive σ_+, σ_-. We generalize this setting by explicitly implementing the effects of squeezings Ξ and rotations in the $(r_\pm, s_\mp)$ variables around an angle $\phi \in [0, 2\pi)$ (realized by tuning ϑ_1 and ϑ_2 as discussed in Sec. III C) which leads to the nonlocal Wigner W and

Husimi Q distributions given explicitly for $a_1 = b_1 = a_2 = b_2 = 1$ in the Appendix. We show both distributions for fixed scalings $a_1 = b_1 = a_2 = b_2 = 1$, fixed $\sigma_+ = 1$, $\sigma_- = 1.2$ and the three characteristic choices $(\xi = 1, \phi = 0)$, $(\xi = 1, \phi = \pi/4)$, $(\xi = 3/2, \phi = 0)$ (from left to right) together with their marginal distributions in Fig. 5.

We analyze the performance of all types of criteria considered in Sec. IV in detail. We evaluate the phase-space (solid red curves) and marginal criteria (blue dashed curves) for the triplet of values for ϕ, ξ as shown in Fig. 5 and compare their performance in terms of σ_+, σ_-.

We start with our second moment criteria (64) and the MGVT criteria (82) (first row in Fig. 6). For $\phi = 0$, $\xi = 1$, shown in Fig. 6(a), the MGVT criteria outperform our criteria, while the converse is true for $\phi = \pi/4$, $\xi = 1$ as depicted in Fig. 6(b) as a result of strong correlations being present. When optimizing over ϕ and ξ, the two become equivalent (see Sec. IV A) as indicated in Fig. 6(c) where $\phi = 0$, $\xi = 3/2$. In all cases, the distributions shown in Fig. 5 corresponding to $\sigma_+ = 1$, $\sigma_- = 1.2$ are not witnessed (circle), while the point $\sigma_+ = 4$, $\sigma_- = 1.5$ (triangle) is an example for a state which is witnessed by our criteria but not witnessed by the MGVT criteria when the angle has been chosen unfavorably. The angle dependence of our criteria (light orange dotted curve) and the MGVT criteria (dark orange dashed to solid curve) for this case as well as the covariance (black solid curve) is shown in Fig. 7(a). Following the intuition provided by (81), our criteria can outperform the MGVT criteria whenever $\sigma_{r_\pm s_\mp}^2 > \mathcal{W}_{\text{MGVT}}$, which is fulfilled for $\pi/8 \lesssim \phi \lesssim 3\pi/8$ and $5\pi/8 \lesssim \phi \lesssim \pi$ (orange shaded regions).

The same analysis is carried out for the Wehrl entropic criteria (55) and the WTSTD criteria (97) in the middle row of Fig. 6. We observe similar effects as for the second moment criteria. The main difference is that no straightforward
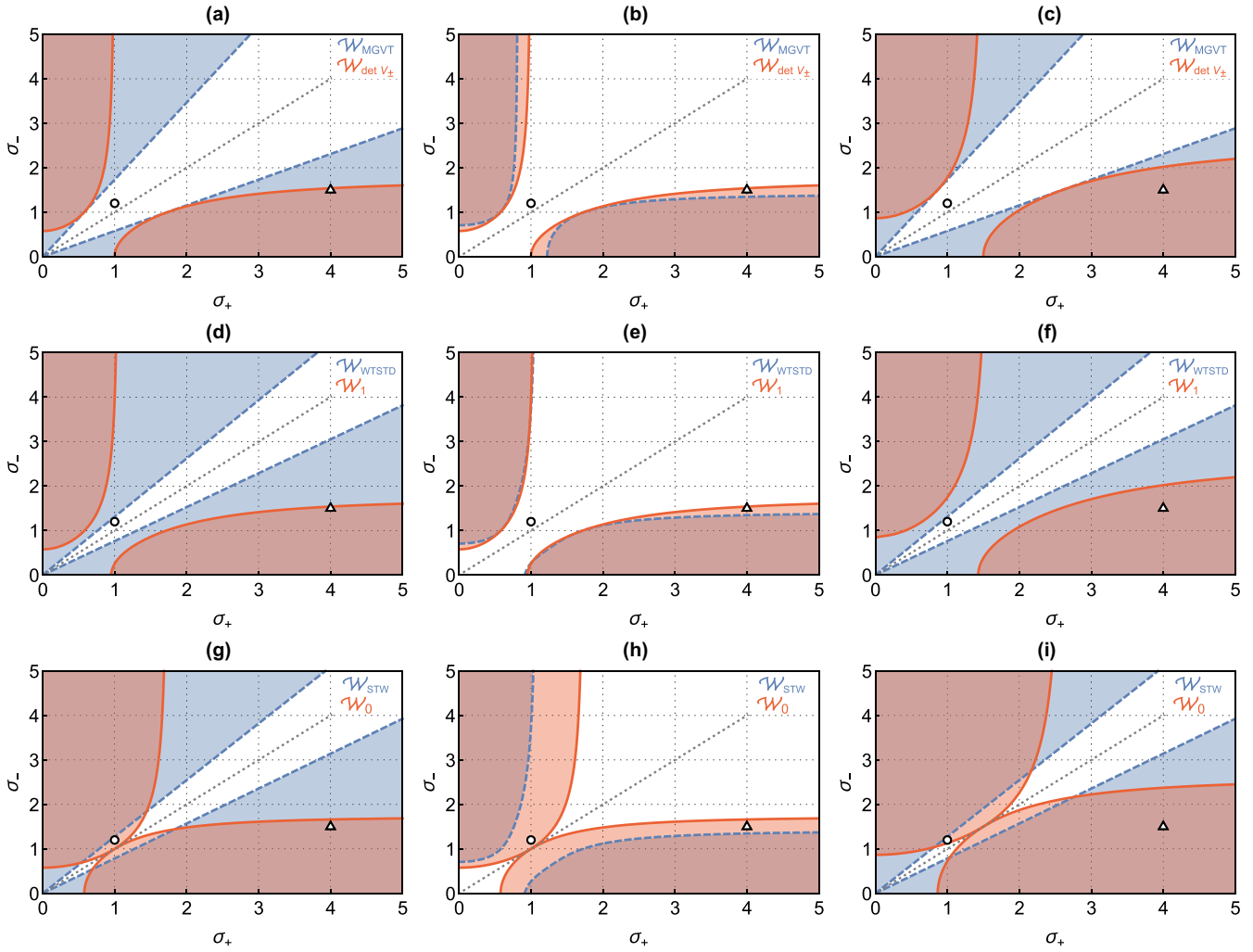
FIG. 6. Witnessed regions of various marginal (dashed blue curves) and phase-space (solid red curves) criteria for the state (100). The upper, middle, lower rows show second moment, entropic, optimized Rényi entropic criteria for the choices $\phi = 0, \xi = 1$ (first column), $\phi = \pi/4$ (second column), and $\phi = 0, \xi = 3/2$ (third column). For $\beta \to 0$ our Rényi-Wehrl criteria (52) certify entanglement for all $\sigma_+ \neq \sigma_-$ after optimizing over ξ as indicated from panels (g) and (i), and hence outperform the STW criteria. In particular, they witness the point $\sigma_+ = 1$, $\sigma_- = 1.2$ (circle) corresponding to the distributions shown in Fig. 5. For highly correlated variables, i.e., for $\phi = \pi/4$ (middle column), the marginal criteria underperform, see, e.g., $\sigma_+ = 4$, $\sigma_- = 1.5$ (triangle) and also Fig. 7. A combined figure composed out of the three lower plots is provided in Ref. [74].
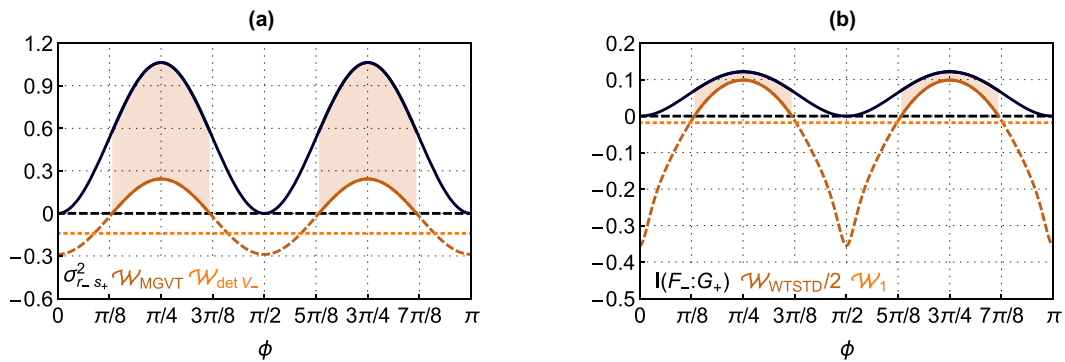

FIG. 7. Illustration of all quantities appearing in the second moment (entropic) relations (81) [(96)] in panels (a) [(b)] for the state (100) with $\sigma_+ = 4$, $\sigma_- = 1.5$ (triangle in Fig. 6) and $\xi = 1$ as functions of ϕ. For sufficiently correlated $r_\pm$ and $s_\mp$, i.e., for $\pi/8 \lesssim \phi \lesssim 3\pi/8$ and $5\pi/8 \lesssim \phi \lesssim \pi$, the marginal criteria (dark orange dashed to solid curves) fail, while the phase-space criteria (light orange dotted curves) still witness entanglement. More precisely, the correlation measures (black solid curves) exceed the values of the marginal witnesses in these regimes (orange shaded regions).
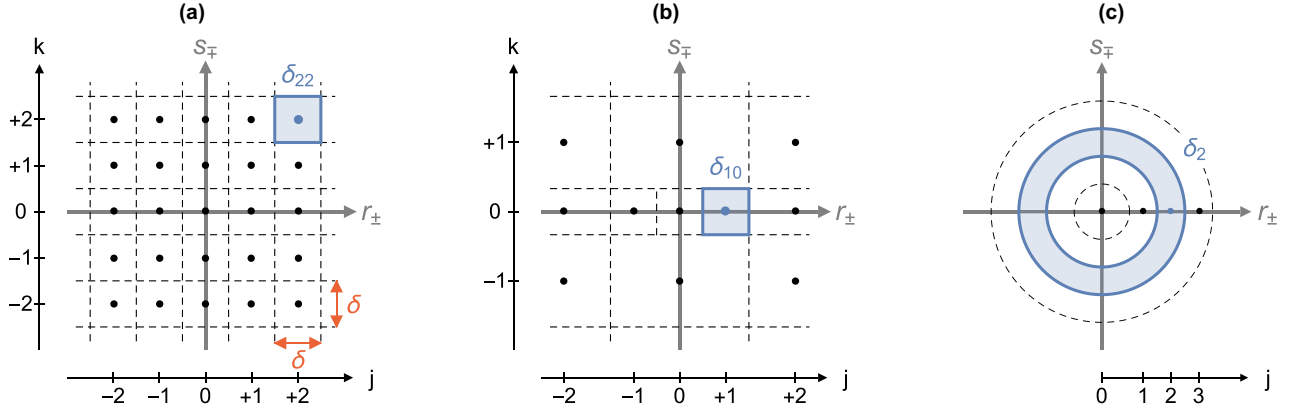
FIG. 8. Three typical discretization schemes: Panel (a) shows regular quadratic tiles with spacing δ along both directions, panel (b) shows an adaptive scheme with smaller tilings closer to the origin, and panel (c) shows a circular tiling around the origin.

optimization leads to an equivalence between them. The point $\sigma_+ = 1$, $\sigma_- = 1.2$ is not witnessed by any of the two and the point $\sigma_+ = 4$, $\sigma_- = 1.5$, which is always witnessed by the Wehrl entropic criteria, is only witnessed by the marginal criteria provided that the correlations between $r_\pm$ and $s_\mp$ remain small. We also plot the quantities appearing in the inequality (96) in Fig. 7(b) and find the same outperformance regions for ϕ as for the second moment criteria.

At last, we compare the Rényi-Wehrl criteria (52) to the STW criteria (98) and optimize over the entropic order for every choice of ϕ, ξ. For $\phi = 0$, the STW criteria become optimal for $\alpha \to 1/2$ and $\beta \to \infty$ in the variables (r_+, s_-) and conversely for (r_-, s_+), i.e.,

$$\mathcal{W}_{\mathrm{STW}} = S_{1/2}(f_+) + S_\infty(g_-) - \ln 4\pi, \tag{101}$$

where $S_\infty(g_-) = -\ln g_-^{\max}$ denotes the min-entropy of g_- with maximum value $g_-^{\max}$ and

$$\mathcal{W}_{\mathrm{STW}} = S_\infty(f_-) + S_{1/2}(g_+) - \ln 4\pi, \tag{102}$$

respectively [17]. For $\phi = \pi/4$, they instead become optimal for $\alpha, \beta \to 1$ and reduce to the WTSTD criteria (97). In contrast, the Rényi-Wehrl criteria (52) become optimal in the limit $\beta \to 0$ independent of ϕ, ξ.

The results are shown in the lower row of Fig. 6. For $\phi = 0, \xi = 1$, our criteria outperform the STW criteria around $\sigma_+ \approx \sigma_-$, see Fig. 6(g). When optimizing over ξ, our criteria witness entanglement for all $\sigma_+, \sigma_- \neq \sigma_+$ (gray dashed line), as indicated in Fig. 6(i), and hence outperform the STW criteria completely. In particular, the point $\sigma_+ = 1$, $\sigma_- = 1.2$ corresponding to the distributions in Fig. 5 can only be witnessed with the phase-space approach after optimization. For $\phi = \pi/4$ the outperformance occurs even without optimizing over ξ, see Fig. 6(h).

V. COARSE-GRAINED MEASUREMENTS

We now discuss the influence of finite resolution on the entanglement criteria (41). In particular, we analyze the possibilities offered by an optimization over the concave function f.

A. Discretization schemes

We aim for a description valid for measurements with constant *and* adaptive resolution. The latter is relevant when an experimental procedure produces samples of the Husimi Q distribution, which are binned in a postmeasurement process ensuring that the underlying distribution is approximated well. A prime example of such schemes is the quadtree method, which has been employed successfully to witness entanglement of a non-Gaussian state using entropic methods in Ref. [95].

To achieve this generality, we discretize phase space into tiles δ_{jk}, where the indices j, k run over integers, but depending on the discretization scheme it may be convenient to draw them from subsets of integers, i.e., $\{j, k\} \in \mathcal{Z} \subseteq \mathbb{Z} \times \mathbb{Z}$. To every tile δ_{jk} we associate its phase-space measure, which we denote by Δ_{jk}, and discrete coordinates $(r_\pm^j, s_\mp^k)$, which are located at the center of δ_{jk} if it corresponds to a simply connected region.

For regular tilings consisting of rectangles with $\Delta \equiv \Delta_{jk}$ there is a simple relation between the continuous and the discrete sets of coordinates which reads

$$r_\pm \to r_\pm^j = j\delta r_\pm, \quad s_\mp \to s_\mp^k = k\delta s_\mp, \tag{103}$$

including the possibility that $\delta r_\pm \neq \delta s_\mp$. Then, the tile δ_{jk} is centered at $(r_\pm^j, s_\mp^k)$ and hence its domain is given by $\delta_{jk} = [\delta r_\pm(j - 1/2), \delta r_\pm(j + 1/2)] \times [\delta s_\mp(k - 1/2), \delta s_\mp(k + 1/2)]$ with constant phase-space measure $\Delta = \delta r_\pm \delta s_\mp/(2\pi)$. In general, such simple relations may only be given recursively or implicitly and hence we work with the discrete indices j, k in the following.

We illustrate three archetypal discretization schemes in Fig. 8. A regular quadratic tiling with $\delta \equiv \delta r_\pm = \delta s_\mp$ and $\Delta = \delta^2/(2\pi)$, most relevant in the context of finite-resolution measurements in quantum optics, is shown in Fig. 8(a). An adaptive scheme with rectangular tiles is sketched in Fig. 8(b), in which case relations like (103) can only be given recursively. However, such schemes allow us to reconstruct the underlying distribution in greater detail in regions where it shows characteristic features, which can be advantageous especially for non-Gaussian distributions. Another adaptive scheme, which is based on radially symmetric tilings, is
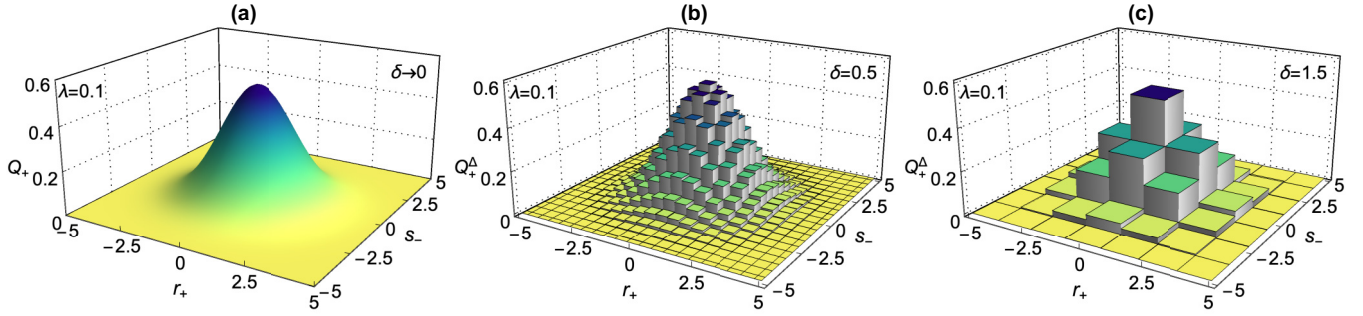
FIG. 9. Coarse-graining of the TMSV state (48) with $\lambda = 0.1$ in the (r_+, s_-) variables following the regular quadratic discretization scheme shown in Fig. 8(a). Starting from the continuum limit $\delta \to 0$ in panel (a), the grid spacing δ increases from panels (a) to (c), resulting in coarser distributions Q_+^Δ.

shown in Fig. 8(c). This method may be employed, for example, when binning sampled data while assuming the underlying distribution to be radially symmetric, such that the discrete coordinates can be labeled with only one non-negative integer-valued index, e.g., $j \in \mathbb{Z}_0^+$.

B. Discretized distributions

The discretization procedure leads to a discrete distribution over the coordinate points $(r_\pm^j, s_\mp^k)$ by integrating the continuous distribution $Q_\pm$ over the corresponding tile δ_{jk}, i.e.,

$$Q_\pm^{jk} \equiv Q_\pm(r_\pm^j, s_\mp^k) = \int_{\delta_{jk}} \frac{dr_\pm ds_\mp}{2\pi} Q_\pm(r_\pm, s_\mp). \quad (104)$$

This distribution is normalized to unity according to

$$\sum_{j,k \in \mathcal{Z}} Q_\pm^{jk} = 1. \quad (105)$$

Uncertainty relations and entanglement criteria for coarse-grained measurements have been studied extensively in the literature, see, e.g., Refs. [89,101,105,106]. The most important conclusion from these studies is that both should *not* be formulated in terms of measures of localization with respect to the discrete distribution $Q_\pm^{jk}$ because they often underestimate their continuous analogs. For example, consider an experimental procedure with sufficiently low resolution such that all measurement outcomes lie within a single tile. Then, variances as well as entropies of the resulting discrete distribution evaluate to zero although their continuous analogs are strictly non-zero.

Therefore, we work with the density of the discrete distribution (104) over every tile instead. It is defined via

$$Q_\pm^\Delta \equiv Q_\pm^\Delta(r_\pm, s_\mp) = \sum_{j,k \in \mathcal{Z}} \begin{cases} \frac{Q_\pm^{jk}}{\Delta_{jk}}, & (r_\pm, s_\mp) \in \delta_{jk} \\ 0 & \text{else} \end{cases} \quad (106)$$

and serves as an approximation to the true continuous Husimi Q distribution $Q_\pm$ albeit being necessarily discontinuous itself. Thus, it is normalized with respect to the standard phase-space measure $dr_\pm ds_\mp/(2\pi)$ and converges to the continuous Husimi Q distribution in the continuum limit, i.e., $Q_\pm^\Delta(r_\pm, s_\mp) \to Q_\pm(r_\pm, s_\mp)$ for $\Delta_{jk} \to 0$.

As an example, we consider the effect of coarse-graining for the Husimi Q distribution $Q_\pm$ of the TMSV state (48),

which we restate here for the reader's convenience:

$$Q_\pm(r_\pm, s_\mp) = \frac{1}{Z} e^{-\frac{1}{2}(r_\pm, s_\mp)^T V_\pm^{-1}(r_\pm, s_\mp)}, \quad (107)$$

with normalization $Z = \det^{1/2} V_\pm$ and covariance matrix $V_\pm = \frac{2}{1\pm\lambda}\mathbb{1}$. We show the case $a_1 = b_1 = a_2 = b_2 = 1$, $\vartheta_1 = \vartheta_2 = 0$, and $\lambda = 0.1$ in Fig. 9 for a regular quadratic tiling with various spacings δ along both axes as shown in Fig. 8(a).

C. Discretized criteria

1. General criteria

To derive entanglement criteria which hold for the discretized approximation $Q_\pm^\Delta$, we apply Jensen's inequality to the integrals appearing in the witness functional (41) for every tile *individually*. Adjusted to this setup, i.e., the continuous Husimi Q distribution $Q_\pm$ restricted to δ_{jk} with measure Δ_{jk} and concave f, Jensen's inequality reads

$$\frac{1}{\Delta_{jk}} \int_{\delta_{jk}} \frac{dr_\pm ds_\mp}{2\pi} f(Q_\pm) \leqslant f\left(\frac{1}{\Delta_{jk}} \int_{\delta_{jk}} \frac{dr_\pm ds_\mp}{2\pi} Q_\pm\right). \quad (108)$$

Now, discretizing the first term in our witness functional $\mathcal{W}_f$ (41) by expanding the phase-space integral over all tiles, inserting a multiplicative one, and using Jensen's inequality (108) together with the definition of the discrete distribution (104) leads to

$$\int \frac{dr_\pm ds_\mp}{2\pi} f(Q_\pm) = \sum_{j,k \in \mathcal{Z}} \Delta_{jk} \frac{1}{\Delta_{jk}} \int_{\delta_{jk}} \frac{dr_\pm ds_\mp}{2\pi} f(Q_\pm)$$

$$\leqslant \sum_{j,k \in \mathcal{Z}} \Delta_{jk} f\left(\frac{1}{\Delta_{jk}} Q_\pm^{jk}\right). \quad (109)$$

Now, the right-hand side of the latter inequality can be rewritten in terms of the continuous approximation

$$\sum_{j,k \in \mathcal{Z}} \Delta_{jk} f\left(\frac{1}{\Delta_{jk}} Q_\pm^{jk}\right) = \int \frac{dr_\pm ds_\mp}{2\pi} f(Q_\pm^\Delta), \quad (110)$$

which is also a useful relation for computing the discretized witness functional in practice as one might prefer to work with $Q_\pm^{jk}$ over $Q_\pm^\Delta$ to simplify calculations.

Defining a discretized witness functional in terms of $Q_\pm^\Delta$ instead of $Q_\pm$,

$$\mathcal{W}_f^\Delta = \int \frac{dr_\pm ds_\mp}{2\pi} [f(Q_\pm^\Delta) - f(\bar{Q}_\pm')], \qquad (111)$$

allows us to conclude $\mathcal{W}_f \leqslant \mathcal{W}_f^\Delta$ and therefore our criteria (41) imply that all separable states fulfill the discretized criteria

$$\boldsymbol{\rho}_{12} \text{ separable} \Rightarrow \mathcal{W}_f^\Delta \geqslant 0, \qquad (112)$$

which are weaker in general. We emphasize that the latter inequality is fulfilled for *arbitrary* discretization schemes. Since the discretized witness functional $\mathcal{W}_f^\Delta$ is of the same form as its continuous counterpart $\mathcal{W}_f$, we can follow the same arguments as above to derive interesting classes of discrete entanglement criteria.

2. Entropic criteria

Similar to (51), we obtain criteria for Rényi-Wehrl entropies of the discretized approximations

$$S_\beta(Q_\pm^\Delta) = \frac{1}{1-\beta} \ln \left[\int \frac{dr_\pm ds_\mp}{2\pi} (Q_\pm^\Delta)^\beta (r_\pm, s_\mp) \right], \quad (113)$$

with entropic orders $\beta \in (0,1) \cup (1, \infty)$, by choosing monomials $f(t) = t^\beta$ and applying a monotonic function in (111), which yields the discretized witness functional

$$\mathcal{W}_\beta^\Delta = S_\beta(Q_\pm^\Delta) - \frac{\ln \beta}{\beta - 1} - \frac{\ln \det \bar{V}_\pm'}{2}, \qquad (114)$$

analogous to (52). In general, we can write the Rényi-Wehrl entropies (113) in terms of the discretized distribution $Q_\pm^{jk}$ as

$$S_\beta(Q_\pm^\Delta) = \frac{1}{1-\beta} \ln \left[\sum_{j,k \in \mathcal{Z}} \Delta_{jk}^{1-\beta} (Q_\pm^{jk})^\beta \right]. \qquad (115)$$

The latter can only be simplified further for regular tilings with $\Delta \equiv \Delta_{jk}$, leading to

$$S_\beta(Q_\pm^\Delta) = S_\beta(Q_\pm^{jk}) + \ln \Delta, \qquad (116)$$

with discrete Rényi-Wehrl entropies

$$S_\beta(Q_\pm^{jk}) = \frac{1}{1-\beta} \ln \left[\sum_{j,k \in \mathcal{Z}} (Q_\pm^{jk})^\beta \right]. \qquad (117)$$

In the limit $\beta \to 1$, the Rényi-Wehrl entropy (113) converges to the Wehrl entropy

$$S_1(Q_\pm^\Delta) = -\int \frac{dr_\pm ds_\mp}{2\pi} Q_\pm^\Delta(r_\pm, s_\mp) \ln Q_\pm^\Delta(r_\pm, s_\mp), \quad (118)$$

for which the discretized witness functional reads

$$\mathcal{W}_1^\Delta = S_1(Q_\pm^\Delta) - 1 - \frac{\ln \det \bar{V}_\pm'}{2}. \qquad (119)$$

Irrespective of the discretization scheme we find the relation

$$S_1(Q_\pm^\Delta) = S_1(Q_\pm^{jk}) + \sum_{j,k \in \mathcal{Z}} \ln(\Delta_{jk}) Q_\pm^{jk}, \qquad (120)$$

with the discrete Wehrl entropy being defined as

$$S_1(Q_\pm^{jk}) = -\sum_{j,k \in \mathcal{Z}} Q_\pm^{jk} \ln Q_\pm^{jk}, \qquad (121)$$

which also follows from (117) in the limit $\beta \to 1$. For regular tilings $\Delta \equiv \Delta_{jk}$, (120) reduces to the simple relation

$$S_1(Q_\pm^\Delta) = S_1(Q_\pm^{jk}) + \ln \Delta. \qquad (122)$$

3. Second moment criteria

Criteria for the second moments of the distribution $Q_\pm^\Delta$, defined via (we assume vanishing expectation values without loss of generality)

$$(\Sigma_{r_\pm}^\Delta)^2 = \int \frac{dr_\pm ds_\mp}{2\pi} r_\pm^2 Q_\pm^\Delta(r_\pm, s_\mp),$$

$$(\Sigma_{s_\mp}^\Delta)^2 = \int \frac{dr_\pm ds_\mp}{2\pi} s_\mp^2 Q_\pm^\Delta(r_\pm, s_\mp),$$

$$\Sigma_{r_\pm s_\mp}^\Delta = \int \frac{dr_\pm ds_\mp}{2\pi} r_\pm s_\mp Q_\pm^\Delta(r_\pm, s_\mp), \qquad (123)$$

can be formulated in terms of the discretized covariance matrix

$$V_\pm^\Delta = \begin{pmatrix} (\Sigma_{r_\pm}^\Delta)^2 & \Sigma_{r_\pm s_\mp}^\Delta \\ \Sigma_{r_\pm s_\mp}^\Delta & (\Sigma_{s_\mp}^\Delta)^2 \end{pmatrix}, \qquad (124)$$

by using that $S_1(Q_\pm^\Delta) \leqslant 1 + \frac{1}{2} \ln \det V_\pm^\Delta$, which leads to the second moment witness functional

$$\mathcal{W}_{\det V_\pm}^\Delta = \det V_\pm^\Delta - \det \bar{V}_\pm'. \qquad (125)$$

One can easily check that the means of the distributions $Q_\pm^\Delta$ and $Q_\pm^{jk}$ agree, i.e.,

$$(\mu_{r_\pm}^\Delta, \mu_{s_\mp}^\Delta) = (\mu_{r_\pm^j}, \mu_{s_\mp^k}), \qquad (126)$$

with the continuous means given by

$$\mu_{r_\pm}^\Delta = \int \frac{dr_\pm ds_\mp}{2\pi} r_\pm Q_\pm^\Delta(r_\pm, s_\mp),$$

$$\mu_{s_\mp}^\Delta = \int \frac{dr_\pm ds_\mp}{2\pi} s_\mp Q_\pm^\Delta(r_\pm, s_\mp), \qquad (127)$$

while the discrete means read

$$\mu_{r_\pm^j} = \sum_{j,k \in \mathcal{Z}} r_\pm^j Q_\pm^{jk},$$

$$\mu_{s_\mp^k} = \sum_{j,k \in \mathcal{Z}} s_\mp^k Q_\pm^{jk}. \qquad (128)$$

In contrast, the second moments of $Q_\pm^\Delta$ and $Q_\pm^{jk}$ differ. For regular tilings with $\Delta = \delta r_\pm \delta s_\mp/(2\pi)$ we find (see also Refs. [105,106])

$$(\Sigma_{r_\pm}^\Delta)^2 = \Sigma_{r_\pm^j}^2 + \frac{(\delta r_\pm)^2}{12},$$

$$(\Sigma_{s_\mp}^\Delta)^2 = \Sigma_{s_\mp^k}^2 + \frac{(\delta s_\mp)^2}{12}, \qquad (129)$$

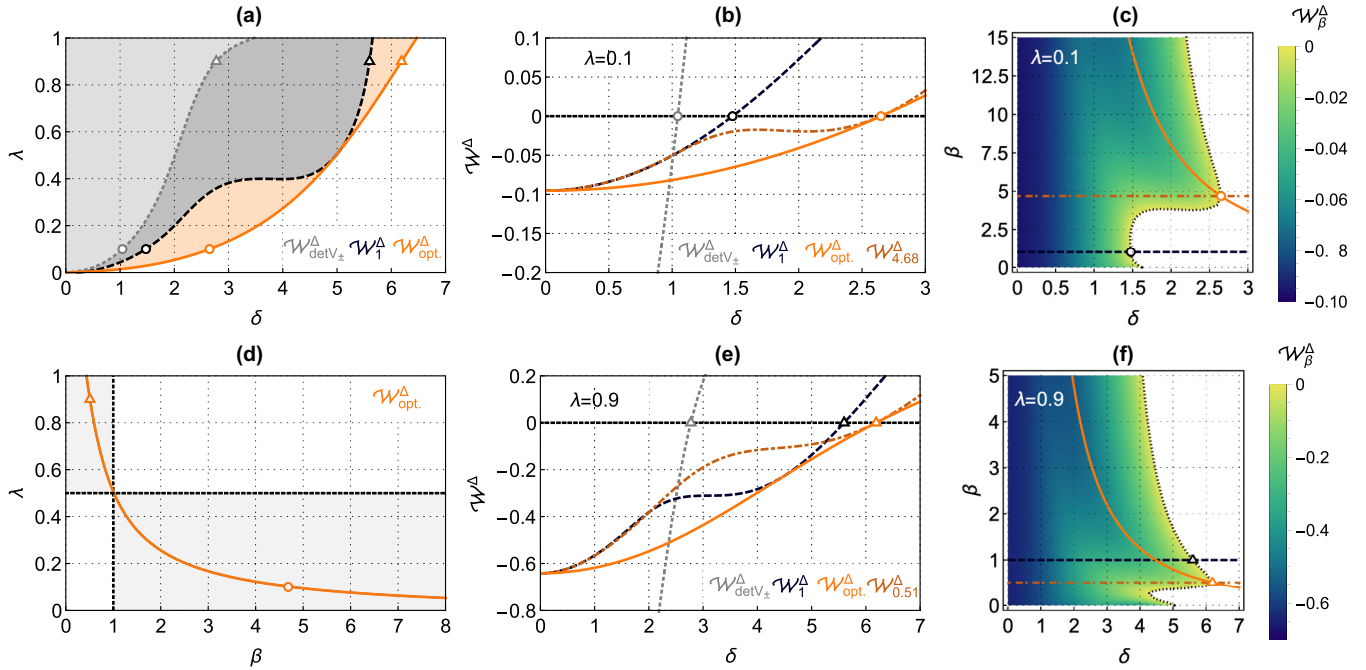$$\Sigma_{r_\pm s_\mp}^\Delta = \Sigma_{r_\pm^j s_\mp^k},$$

FIG. 10. Comparison of discretized entanglement criteria for the TMSV state. As shown in panel (a), there is a strict hierarchy of second moment (gray dotted curves), Wehrl entropic (black dashed curves), and optimized Rényi-Wehrl entropic (light orange solid curves) criteria in the sense that, for a fixed λ, optimization allows us to work with a much coarser grid spacing δ before the witness breaks down (see also Ref. [74]). The entanglement analysis of the Rényi-Wehrl criteria as a function of β is carried out explicitly in panels (c) and (f) for $\lambda = 0.1$ and $\lambda = 0.9$, respectively, where we distinguish between the optimal β overall (dark orange dot-dashed curves) and the optimal β for every δ (light orange solid curves). The values of the four witnesses are shown in panels (b) and (e) as functions of the grid spacing δ for the same choices of λ. The optimal choice for β depends on λ, as plotted in panel (d), showing that for small entanglement $\lambda \leqslant 0.5$ one should take $\beta \geqslant 1$ and vice versa.

with discrete second moments (we assume vanishing expectation values again)

$$\Sigma_{r_\pm^j}^2 = \sum_{j,k \in \mathcal{Z}} (r_\pm^j)^2 Q_\pm^{jk},$$

$$\Sigma_{s_\mp^k}^2 = \sum_{j,k \in \mathcal{Z}} (s_\mp^k)^2 Q_\pm^{jk},$$

$$\Sigma_{r_\pm^j s_\mp^k} = \sum_{j,k \in \mathcal{Z}} r_\pm^j s_\mp^k Q_\pm^{jk}. \tag{130}$$

Equation (129) shows that the discrete second moments $\Sigma_{r_\pm^j}^2$ and $\Sigma_{s_\mp^k}^2$ indeed underestimate the continuous variances $(\Sigma_{r_\pm}^\Delta)^2$ and $(\Sigma_{s_\mp}^\Delta)^2$, an effect which is cured by taking the variances induced by the finite tile sizes into account.

D. Example state

To exemplify the advantages offered by an optimization over f in our general discretized criteria (112) we consider again the Gaussian TMSV state (48) for $a_1 = b_1 = a_2 = b_2 = 1$ and $\vartheta_1 = \vartheta_2 = 0$. The corresponding distribution $Q_\pm$ is discretized following a regular quadratic tiling with grid spacing δ shown in Fig. 8(a), leading to a discretized distribution $Q_\pm^\Delta$, which is illustrated for various δ and $\lambda = 0.1$ in Fig. 9. When statistical errors become negligible, which is a justified assumption for many quantum optics setups, the discretized distribution $Q_\pm^\Delta$ can be measured directly.

Most importantly, the discretization breaks Gaussianity and hence the choice of the function f matters. In Fig. 10(a) we compare the performances of the discretized versions of our second moment criteria (gray dotted curves), the Wehrl entropic criteria (black dashed curves), and the optimized Rényi-Wehrl entropic criteria (light orange solid curves) by plotting the witnessed regions for every λ as a function of the grid spacing δ. In general, we can say that entropic criteria strongly outperform second moment criteria and that an optimization over the order β for every λ provides a second substantial improvement, especially for small $\lambda \approx 0.2$.

We show the regions where the Rényi-Wehrl witness is negative for $\lambda = 0.1$ and $\lambda = 0.9$ as a function of β in Figs. 10(c) and 10(f), respectively, with the black dotted line indicating where the witness evaluates to zero. Note here that the witness functional becomes independent of β when $\delta \to 0$ because the underlying state is Gaussian. Therein, we also show the curves obtained for $\beta = 1$ (black dashed curves), the value for β maximizing entanglement detection (dark orange dot-dashed curves), and the optimal value for β for every δ (light orange solid curves).

The latter three curves are drawn in Figs. 10(b) and 10(e), respectively, together with the curves for the second moment criteria (gray). Considering for example $\lambda = 0.1$ in Fig. 10(b), the Wehrl entropic witness already breaks down around $\delta \approx 1.5$, which corresponds to the level of discreteness shown in Fig. 9(c). In contrast, the optimal Rényi-Wehrl entropic witness certifies entanglement up to $\delta \approx 2.65$.
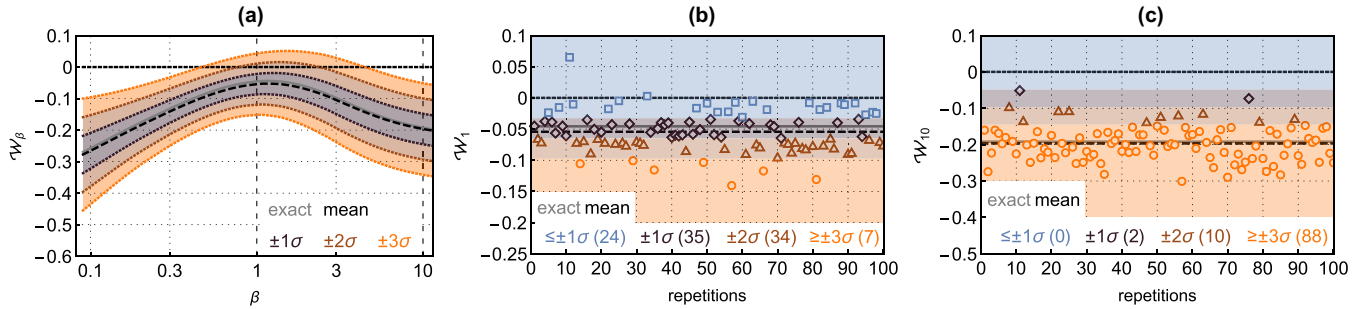
FIG. 11. (a) Mean (black dashed curve) and confidence intervals (shaded regions) of the Rényi-Wehrl criteria (52) for 10^3 samples of the state (131) after 10^2 repetitions of this simulated experiment (see also Ref. [74]). Panels (b) and (c) show the computed values for all repetitions for $\beta \to 1$ and $\beta = 10$, respectively. Optimizing over β, i.e., for β on the order of 10^1, certification of entanglement within large confidence intervals is improved substantially (squares, diamonds, triangles, and circles correspond to $\leqslant \pm 1\sigma$, $\pm 1\sigma$, $\pm 2\sigma$, and $\geqslant \pm 3\sigma$, respectively).

Furthermore, the optimal choice for β is a monotonically decreasing function of λ, which we plot in Fig. 10(d). Roughly speaking, for large entanglement $\lambda \geqslant 0.5$ we need small $\beta \leqslant 1$, while for small entanglement $\lambda \leqslant 0.5$ we should choose large $\beta \geqslant 1$.

VI. SAMPLING MEASUREMENTS

We investigate our entanglement criteria for the second experimentally relevant scenario: Sampling from the Husimi Q distribution with limited statistics.

A. Estimation of functionals of probability densities

Estimating functionals of a probability density function (PDF), such as the witness functional $\mathcal{W}_f$ (41), from samples is a central problem in statistical data analysis. It generally requires density estimation, the construction of an analytical estimate of the underlying PDF based on observed data. A plethora of methods exists for this, including nonparametric approaches like simple data binning or kernel density estimation [107], maximum entropy models [108], and parametric deep-learning-based approaches [109,110].

For the specific case of entropies, studied here, also direct estimation techniques have been devised (see Ref. [111] for a review) including the popular k nearest-neighbor method [112]. Generally, these methods rely on assumptions about the smoothness of the underlying PDF which allow one to bound the approximation error. In this respect, the Husimi Q distribution has favorable properties due to the uncertainty principle which leads to an inherently smooth behavior. In contrast to the Wigner W distribution, which can have arbitrarily sharp features in the nonlocal phase space, the simultaneous detection of two conjugate quadratures leads to a coarse-graining or smoothing of the distribution in the case of Husimi.

We generate synthetic sample data sets drawn from an analytically known Husimi Q distribution and use a Gaussian mixture model to reconstruct $Q_\pm$ and calculate the witness functional $\mathcal{W}_\beta$ (52). Using moderate sample set sizes of $O(10^3)$ we are able to reliably estimate the witness functional. Crucially, by using different Rényi-Wehrl entropies parametrized by β we find that values of small or large β lead to a significantly increased signal-to-noise ratio of the en-

tanglement detection compared with the case of the standard Wehrl entropy ($\beta \to 1$).

B. Example state

We consider a mixture of two displaced TMSV states with equal squeezing parameters λ, opposite displacements $\pm r$ with $r \geqslant 0$ along the $r_\pm$ axis and a mixing probability $p \in [0, 1]$ such that $p = 0$ selects the state displaced towards positive values for $r_\pm$. Its Husimi Q distribution over a pair of nonlocal variables is

$$Q_\pm(r_\pm, s_\mp) = (1-p)\frac{1+\lambda}{2}e^{-\frac{1+\lambda}{4}[(r_\pm - r)^2 + s_\mp^2]}$$

$$+ p\frac{1+\lambda}{2}e^{-\frac{1+\lambda}{4}[(r_\pm + r)^2 + s_\mp^2]}. \qquad (131)$$

Note that this setup is similar to the non-Gaussian state investigated in Refs. [16,17,74].

We test the performance of the Rényi-Wehrl criteria (52) for the state (131) with $\lambda = 0.8$, $r = 2$, $p = 0.3$. This state is not witnessed by the second moment criteria. We simulate measurements by drawing 10^3 samples and evaluate the witness functional using a Gaussian mixture model from the built-in machine-learning density estimation methods in *Mathematica*. This method models the probability density using a mixture of multivariate normal distributions. For details we refer to the documentation of the *Mathematica* function "LearnDensity" and the associated method "GaussianMixture." We gather information on the statistical quantities, in particular mean values and confidence intervals, by repeating this procedure 10^2 times.

The β dependence of the mean value and the three σ intervals is shown in Fig. 11(a). As the mean value (dashed black) aligns with the exact result (solid gray), our estimation method is justified *a posteriori*. We observe that the choice of β has a strong influence on the signal-to-noise ratio, which we exemplify for $\beta \to 1$ and $\beta = 10$ in Figs. 11(b) and 11(c), respectively. While the standard Wehrl witness ($\beta \to 1$) is not able to witness entanglement within $\pm 1\sigma$ for all repetitions, this can be achieved for $\beta = 10$. In the latter case entanglement is even certified within $\pm 3\sigma$ in 88% of all cases [see orange circles in Fig. 11(c)].

VII. CONCLUSIONS AND OUTLOOK

To summarize, using the Husimi Q distribution for constructing entanglement witnesses has several advantages. In contrast to marginal distributions, it contains the full information about the underlying quantum state, which can thus be assessed though measurements in a single experimental setting. Most importantly, the existence of general uncertainty relations permits the derivation of entanglement criteria of much more general form than any known families of marginal based criteria. We showed the resulting strengths in different ways, including derivations of classes of entropic and second moment criteria as well as comparisons with marginal criteria. For a well-known family of states we were able to certify entanglement beyond the capabilities of marginal criteria based on uncertainty relations.

For future work, it would be interesting to investigate whether the requirement $a_1 b_1 = a_2 b_2$ can be relaxed, in which case a tensor product decomposition of the bipartite Hilbert space with respect to the nonlocal variables becomes impossible. However, we believe that our entanglement criteria do still hold. This conjecture is motivated by the fact that the entropic criteria (55) can be derived without the mentioned restriction on the scaling parameters using the entropy power inequality and local uncertainty relations along the lines of Ref. [36].

Whether our criteria are the strongest criteria following from the uncertainty principle in phase space and what are the limitations of the Husimi approach, i.e., whether there exist any states which are witnessed by the STW but not by our criteria, are questions of central importance for future investigations. In general, we expect all entanglement criteria based on the linear nonlocal variables (32) to work best whenever the correlations between the local quadratures are mostly linear. In this case, an analysis of fluctuations in the nonlocal variables resembles the typical statistical analysis with Pearson's coefficient, i.e., the search for linear correlations between the two subsystems. Therefore, we believe that our criteria become weak for states where entanglement manifests itself in higher-order correlations, which are not accurately captured by linear nonlocal variables. Such states are often characterized by partly negative Wigner W distributions, consider, for example, NOON states or the photonic qutrit state. In both cases, we have confirmed that our witness does *not* flag entanglement (in both cases, optimizing over β leads to $\mathcal{W}_\beta \to 0$ from above). Nevertheless, we do not see any arguments that would generally exclude detecting entanglement of states with negativities, and we consider finding such a state as an interesting problem for future work.

Moreover, we wonder whether other positive but not completely positive maps beyond the partial transpose have a faithful representation for the Husimi Q distribution, which would allow us to formulate yet another different class of entanglement criteria. Here, an approach compensating for the lack of detecting bound entanglement is of special interest.

As our derivation relies on the group-theoretic properties of coherent states, another interesting possibility would be to derive analogous criteria for physical systems described by other algebras, for example, spin observables with an SU(2) algebra. Furthermore, it would be desirable to generalize our

findings to incorporate entanglement measures in the spirit of Refs. [22,89,96,97] (see also Ref. [113]) to set measurable lower bounds on the amount of entanglement. Finally, other potential directions where the Husimi approach might reveal its strengths encompass criteria for (genuine) multipartite entanglement as well as criteria based on higher-order moments of the Husimi Q distribution in the sense of Ref. [26], which we expect to be of simpler accessibility than those based on higher-order moments of the Wigner W distribution.

Regarding practical applications, we have shown that the optimization prospects of the generalized entanglement criteria based on the Husimi Q distribution lead to clear performance advantages compared with the Wehrl-entropic criteria in situations with sparse experimental data. In a scenario, where this distribution is only known at a finite number of grid points within phase space, one effectively deals with a step function approximation to the exact continuous distribution. As a result, even for Gaussian states, optimizing the witness functional $\mathcal{W}_f$ over a parametrized family of concave functions f—in our case monomials t^β, relating the witness functional to Rényi-Wehrl entropies—greatly enlarges the range of measurement resolution in which entanglement is detected compared with second-moment-based or Wehrl entropic criteria.

In a second scenario, where experimental measurements correspond to drawing samples from the Husimi Q distribution, we found that both large ($\gg 1$) and small ($\ll 1$) values of β lead to an increased statistical significance of the entanglement detection compared with the standard Wehrl entropy. Intuitively, increasing β allows us to reduce the influence of regions of small probability on $\mathcal{W}_f$, i.e., reduce the influence of the tail behavior of the distribution. This can be beneficial because, for finite statistics, these regions are very sparsely sampled and thus incur a large statistical error. The good performance for small β, where the contribution of the tail regions to $\mathcal{W}_f$ are amplified, might be explained by the employed density estimation routine making justified assumptions about the functional form of the tails.

Overall, not only the generality of our entanglement criteria but also the inherent smoothness of the Husimi Q distribution are found to lead to practical advantages. While one may expect that estimating functionals of a PDF is a much harder task than estimating low moments, it turns out that in the case of the Husimi Q distribution standard density estimation methods work reliably, even for rather small sample sizes. Given that the Husimi Q distribution is a complete representation of the state, this leads to advantages for entanglement detection compared with approaches based on second moments or marginals of the Wigner W distribution.

In the future, one may exploit the prior knowledge about the properties of the Husimi Q distribution—like smoothness, tail behavior, or symmetries—more systematically by using tailored density estimation methods [108]. Moreover, for specific choices of f direct ways of extracting $f(Q_\pm)$, or its phase-space integral [111,112], from samples exist, circumventing full density estimation, which may lead to even more data-efficient estimates. Combined with the flexibility of our entanglement criteria to explore larger families of functions f, this has the potential to significantly reduce the experimental

cost of certifying entanglement in continuous variable systems beyond the results reported here.

ACKNOWLEDGMENTS

APPENDIX: PHASE-SPACE DISTRIBUTIONS

The nonlocal quasiprobability distributions corresponding to the wave function (100) with arbitrary orientation and squeezing read

$$W_\pm(r_\pm, s_\mp) = \frac{1}{2\pi} \exp\left\{-\frac{\sigma_\mp^2}{2}\left[\xi\sin(\phi)r_\pm - \frac{\cos(\phi)}{\xi}s_\mp\right]^2 - \frac{1}{2\sigma_\pm^2}\left[\xi\cos(\phi)r_\pm + \frac{\sin(\phi)}{\xi}s_\mp\right]^2\right\}$$

$$\times \begin{cases} \frac{\sigma_-}{\sigma_+^3}\left[\xi\cos(\phi)r_+ + \frac{\sin(\phi)}{\xi}s_-\right]^2 & \text{for } (r_+, s_-) \\ \frac{\sigma_+^3}{\sigma_-}\left[\xi\sin(\phi)r_- - \frac{\cos(\phi)}{\xi}s_+\right]^2 & \text{for } (r_-, s_+), \end{cases}$$

$$Q_\pm(r_\pm, s_\mp) = \frac{1}{\sqrt{(\xi^2 + \sigma_-^2)(\xi^2 + \sigma_+^2)^5}} \exp\left\{-\frac{\sigma_\mp^2}{2(\xi^2 + \sigma_\mp^2)}[\sin(\phi)r_\pm - \cos(\phi)s_\mp]^2 - \frac{\xi^2}{2(\xi^2 + \sigma_\pm^2)}[\cos(\phi)r_\pm + \sin(\phi)s_\mp]^2\right\}$$

$$\times \begin{cases} \xi^3\sigma_-[\xi^2 + \sigma_+^2(1 + (\cos(\phi)r_+ + \sin(\phi)s_-)^2)] & \text{for } (r_+, s_-) \\ \xi\sigma_+^3[\sigma_+^2 + \xi^2(1 + (\sin(\phi)r_- - \cos(\phi)s_+)^2)] & \text{for } (r_-, s_+). \end{cases} \tag{A1}$$

[1] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777 (1935).

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

[3] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).

[4] O. Gühne and G. Tóth, Entanglement detection, Phys. Rep. **474**, 1 (2009).

[5] 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).

[6] A. Peres, Separability criterion for density matrices, Phys. Rev. Lett. **77**, 1413 (1996).

[7] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: Necessary and sufficient conditions, Phys. Lett. A **223**, 1 (1996).

[8] H. Nha and M. S. Zubairy, Uncertainty inequalities as entanglement criteria for negative partial-transpose states, Phys. Rev. Lett. **101**, 130402 (2008).

[9] M. Horodecki, P. Horodecki, and R. Horodecki, Mixed-state entanglement and distillation: Is there a "bound" entanglement in nature? Phys. Rev. Lett. **80**, 5239 (1998).

[10] L.-M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, Inseparability criterion for continuous variable systems, Phys. Rev. Lett. **84**, 2722 (2000).

[11] S. Mancini, V. Giovannetti, D. Vitali, and P. Tombesi, Entangling macroscopic oscillators exploiting radiation pressure, Phys. Rev. Lett. **88**, 120401 (2002).

[12] V. Giovannetti, S. Mancini, D. Vitali, and P. Tombesi, Characterizing the entanglement of bipartite quantum systems, Phys. Rev. A **67**, 022320 (2003).

[13] R. Simon, Peres-Horodecki separability criterion for continuous variable systems, Phys. Rev. Lett. **84**, 2726 (2000).

[14] A. Serafini, *Quantum Continuous Variables* (CRC Press, Boca Raton, 2017).

[15] L. Lami, A. Serafini, and G. Adesso, Gaussian entanglement revisited, New J. Phys. **20**, 023030 (2018).

[16] S. P. Walborn, B. G. Taketani, A. Salles, F. Toscano, and R. L. de Matos Filho, Entropic entanglement criteria for continuous variables, Phys. Rev. Lett. **103**, 160505 (2009).

[17] A. Saboia, F. Toscano, and S. P. Walborn, Family of continuous-variable entanglement criteria using general entropy functions, Phys. Rev. A **83**, 032307 (2011).

[18] M. D. Reid, Demonstration of the Einstein-Podolsky-Rosen paradox using nondegenerate parametric amplification, Phys. Rev. A **40**, 913 (1989).

[19] Q. Y. He and M. D. Reid, Genuine multipartite Einstein-Podolsky-Rosen steering, Phys. Rev. Lett. **111**, 250403 (2013).

[20] S. P. Walborn, A. Salles, R. M. Gomes, F. Toscano, and P. H. Souto Ribeiro, Revealing hidden Einstein-Podolsky-Rosen nonlocality, Phys. Rev. Lett. **106**, 130402 (2011).

[21] P. Chowdhury, T. Pramanik, A. S. Majumdar, and G. S. Agarwal, Einstein-Podolsky-Rosen steering using quantum correlations in non-Gaussian entangled states, Phys. Rev. A **89**, 012104 (2014).

[22] J. Schneeloch and G. A. Howland, Quantifying high-dimensional entanglement with Einstein-Podolsky-Rosen correlations, Phys. Rev. A **97**, 042338 (2018).

[23] C. Gneiting and K. Hornberger, Detecting entanglement in spatial interference, Phys. Rev. Lett. **106**, 210501 (2011).

[24] M. A. D. Carvalho, J. Ferraz, G. F. Borges, P.-L. de Assis, S. Pádua, and S. P. Walborn, Experimental observation of quantum correlations in modular variables, Phys. Rev. A **86**, 032332 (2012).

[25] G. S. Agarwal and A. Biswas, Inseparability inequalities for higher order moments for bipartite systems, New J. Phys. **7**, 211 (2005).

[26] E. Shchukin and W. Vogel, Inseparability criteria for continuous bipartite quantum states, Phys. Rev. Lett. **95**, 230502 (2005).

[27] C. Griffet, T. Haas, and N. J. Cerf, Accessing continuous-variable entanglement witnesses with multimode spin observables, Phys. Rev. A **108**, 022421 (2023).

[28] M. Gessner, L. Pezzè, and A. Smerzi, Entanglement and squeezing in continuous-variable systems, Quantum **1**, 17 (2017).

[29] Z. Qin, M. Gessner, Z. Ren, X. Deng, D. Han, W. Li, X. Su, A. Smerzi, and K. Peng, Characterizing the multipartite continuous-variable entanglement structure from squeezing coefficients and the Fisher information, npj Quantum Inf. **5**, 3 (2019).

[30] K. Husimi, Some formal properties of the density matrix, Proc. Phys.-Math. Soc. Jap. 3rd Ser. **22**, 264 (1940).

[31] N. D. Cartwright, A non-negative Wigner-type distribution, Physica A (Amsterdam, Neth.) **83**, 210 (1976).

[32] H.-W. Lee, Theory and application of the quantum phase-space distribution functions, Phys. Rep. **259**, 147 (1995).

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

[34] D.-G. Welsch, W. Vogel, and T. Opatrný, *II Homodyne Detection and Quantum-State Reconstruction*, Progress in Optics (Elsevier, 1999), Vol. 39, pp. 63–211.

[35] S. Floerchinger, T. Haas, and H. Müller-Groeling, Wehrl entropy, entropic uncertainty relations, and entanglement, Phys. Rev. A **103**, 062222 (2021).

[36] S. Floerchinger, M. Gärttner, T. Haas, and O. R. Stockdale, Entropic entanglement criteria in phase space, Phys. Rev. A **105**, 012409 (2022).

[37] E. H. Lieb and J. P. Solovej, Proof of an entropy conjecture for Bloch coherent spin states and its generalizations, Acta Math. **212**, 379 (2014).

[38] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States*, 2nd ed. (John Wiley and Sons, New York, 2017).

[39] P. Schupp, Wehrl entropy, coherent states and quantum channels, in *The Physics and Mathematics of Elliott Lieb* (EMS Press, Berlin, 2022), pp. 329–344.

[40] A. Wehrl, General properties of entropy, Rev. Mod. Phys. **50**, 221 (1978).

[41] A. Wehrl, On the relation between classical and quantum-mechanical entropy, Rep. Math. Phys. **16**, 353 (1979).

[42] E. H. Lieb, Proof of an entropy conjecture of Wehrl, Commun. Math. Phys. **62**, 35 (1978).

[43] M. Grabowski, Wehrl-Lieb's inequality for entropy and the uncertainty relation, Rep. Math. Phys. **20**, 153 (1984).

[44] E. A. Carlen, Some integral identities and inequalities for entire functions and their application to the coherent state transform, J. Funct. Anal. **97**, 231 (1991).

[45] S. Luo, A simple proof of Wehrl's conjecture on entropy, J. Phys. A: Math. Theor. **33**, 3093 (2000).

[46] P. Schupp, On Lieb's conjecture for the Wehrl entropy of Bloch coherent states, Commun. Math. Phys. **207**, 481 (1999).

[47] E. H. Lieb and J. P. Solovej, Proof of the Wehrl-type entropy conjecture for symmetric SU(*N*) coherent states, Commun. Math. Phys. **348**, 567 (2016).

[48] E. H. Lieb and J. P. Solovej, Wehrl-type coherent state entropy inequalities for SU(1,1) and its AX + B subgroup, in *Partial Differential Equations, Spectral Theory, and Mathematical Physics*, edited by P. Exner, R. L. Frank, F. Gesztesy, H. Holden, and T. Weidl (EMS Press, Berlin, Germany, 2021), pp. 301–314.

[49] A. Kulikov, Functionals with extrema at reproducing kernels, Geom. Funct. Ana. **32**, 938 (2022).

[50] M. A. Nielsen, Conditions for a class of entanglement transformations, Phys. Rev. Lett. **83**, 436 (1999).

[51] A. Hertz and N. J. Cerf, Continuous-variable entropic uncertainty relations, J. Phys. A: Math. Theor. **52**, 173001 (2019).

[52] Z. Van Herstraeten and N. J. Cerf, Quantum Wigner entropy, Phys. Rev. A **104**, 042211 (2021).

[53] Z. V. Herstraeten, M. G. Jabbour, and N. J. Cerf, Continuous majorization in quantum phase space, Quantum **7**, 1021 (2021).

[54] Z. Van Herstraeten, Ph.D. thesis, Université libre de Bruxelles, 2021.

[55] U. Leonhardt and H. Paul, Measuring the quantum state of light, Prog. Quantum Electron. **19**, 89 (1995).

[56] T. Opatrný and D.-G. Welsch, Density-matrix reconstruction by unbalanced homodyning, Phys. Rev. A **55**, 1462 (1997).

[57] S. Mancini, P. Tombesi, and V. I. Man'ko, Density matrix from photon number tomography, Europhys. Lett. **37**, 79 (1997).

[58] C. Shen, R. W. Heeres, P. Reinhold, L. Jiang, Y.-K. Liu, R. J. Schoelkopf, and L. Jiang, Optimized tomography of continuous variable systems using excitation counting, Phys. Rev. A **94**, 052327 (2016).

[59] G. Kirchmair, B. Vlastakis, Z. Leghtas, S. E. Nigg, H. Paik, E. Ginossar, M. Mirrahimi, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Observation of quantum state collapse and revival due to the single-photon Kerr effect, Nature (London) **495**, 205 (2013).

[60] C. Wang, Y. Y. Gao, P. Reinhold, R. W. Heeres, N. Ofek, K. Chou, C. Axline, M. Reagor, J. Blumoff, K. M. Sliwa, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, A Schrödinger cat living in two boxes, Science **352**, 1087 (2016).

[61] 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).

[62] G. Barontini, L. Hohmann, F. Haas, J. Estève, and J. Reichel, Deterministic generation of multiparticle entanglement by quantum Zeno dynamics, Science **349**, 1317 (2015).

[63] 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).

[64] M. Gärttner, J. G. Bohnet, A. Safavi-Naini, M. L. Wall, J. J. Bollinger, and A. M. Rey, Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet, Nat. Phys. **13**, 781 (2017).

[65] O. Landon-Cardinal, L. C. G. Govia, and A. A. Clerk, Quantitative tomography for continuous variable quantum systems, Phys. Rev. Lett. **120**, 090501 (2018).

[66] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, England, 2013).

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

[68] J. W. Noh, A. Fougères, and L. Mandel, Measurement of the quantum phase by photon counting, Phys. Rev. Lett. **67**, 1426 (1991).

[69] J. W. Noh, A. Fougères, and L. Mandel, Operational approach to the phase of a quantum field, Phys. Rev. A **45**, 424 (1992).

[70] U. Leonhardt and H. Paul, Phase measurement and Q function, Phys. Rev. A **47**, R2460 (1993).

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

[72] P. Kunkel, M. Prüfer, S. Lannig, R. Rosa-Medina, A. Bonnin, M. Gärttner, H. Strobel, and M. K. Oberthaler, Simultaneous readout of noncommuting collective spin observables beyond the standard quantum limit, Phys. Rev. Lett. **123**, 063603 (2019).

[73] P. Kunkel, M. Prüfer, S. Lannig, R. Strohmaier, M. Gärttner, H. Strobel, and M. K. Oberthaler, Detecting entanglement structure in continuous many-body quantum systems, Phys. Rev. Lett. **128**, 020402 (2022).

[74] M. Gärttner, T. Haas, and J. Noll, companion paper, General class of continuous variable entanglement criteria, Phys. Rev. Lett. **131**, 150201 (2023).

[75] C. Gerving, T. Hoang, B. Land, M. Anquez, C. Hamley, and M. Chapman, Non-equilibrium dynamics of an unstable quantum pendulum explored in a spin-1 Bose–Einstein condensate, Nat. Commun. **3**, 1169 (2012).

[76] C. D. Hamley, C. S. Gerving, T. M. Hoang, E. M. Bookjans, and M. S. Chapman, Spin-nematic squeezed vacuum in a quantum gas, Nat. Phys. **8**, 305 (2012).

[77] P. Kunkel, M. Prüfer, H. Strobel, D. Linnemann, A. Frölian, T. Gasenzer, M. Gärttner, and M. K. Oberthaler, Spatially distributed multipartite entanglement enables EPR steering of atomic clouds, Science **360**, 413 (2018).

[78] W.-M. Zhang, D. H. Feng, and R. Gilmore, Coherent states: Theory and some applications, Rev. Mod. Phys. **62**, 867 (1990).

[79] J. M. Radcliffe, Some properties of coherent spin states, J. Phys. A: Gen. Phys. **4**, 313 (1971).

[80] R. Gilmore, On properties of coherent states, Rev. Mex. de Fis. **23**, 143 (1974).

[81] W. P. Schleich, *Quantum Optics in Phase Space* (Wiley-VCH Verlag Berlin, Berlin, 2001).

[82] W. Heisenberg, Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, Eur. Phys. J. A **43**, 172 (1927).

[83] E. H. Kennard, Zur Quantenmechanik einfacher Bewegungstypen, Eur. Phys. J. A **44**, 326 (1927).

[84] E. Schrödinger, Zum Heisenbergschen Unschärfeprinzip, Sitzungsberichte der Preußischen Akademie der Wissenschaften. Physikalisch-mathematische Klasse **14**, 296 (1930).

[85] H. P. Robertson, The uncertainty principle, Phys. Rev. **34**, 163 (1929).

[86] H. P. Robertson, A general formulation of the uncertainty principle and its classical interpretation, Phys. Rev. **35**, 667 (1930).

[87] H. Maassen and J. B. M. Uffink, Generalized entropic uncertainty relations, Phys. Rev. Lett. **60**, 1103 (1988).

[88] I. Białynicki-Birula and J. Mycielski, Uncertainty relations for information entropy in wave mechanics, Commun. Math. Phys. **44**, 129 (1975).

[89] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, Entropic uncertainty relations and their applications, Rev. Mod. Phys. **89**, 015002 (2017).

[90] S. Floerchinger, T. Haas, and B. Hoeber, Relative entropic uncertainty relation, Phys. Rev. A **103**, 062209 (2021).

[91] S. Floerchinger, T. Haas, and M. Schröfl, Relative entropic uncertainty relation for scalar quantum fields, SciPost Phys. **12**, 089 (2022).

[92] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications* (Springer, New York, 2011).

[93] O. Gühne, Characterizing entanglement via uncertainty relations, Phys. Rev. Lett. **92**, 117903 (2004).

[94] O. Gühne and M. Lewenstein, Entropic uncertainty relations and entanglement, Phys. Rev. A **70**, 022316 (2004).

[95] J. Schneeloch, C. C. Tison, M. L. Fanto, P. M. Alsing, and G. A. Howland, Quantifying entanglement in a 68-billion-dimensional quantum state space, Nat. Commun. **10**, 2785 (2019).

[96] B. Bergh and M. Gärttner, Entanglement detection in quantum many-body systems using entropic uncertainty relations, Phys. Rev. A **103**, 052412 (2021).

[97] B. Bergh and M. Gärttner, Experimentally accessible bounds on distillable entanglement from entropic uncertainty relations, Phys. Rev. Lett. **126**, 190503 (2021).

[98] C. Vignat, A. O. Hero, and J. A. Costa, A geometric characterization of maximum rényi entropy distributions, in *2006 IEEE International Symposium on Information Theory* (IEEE, 2006), pp. 1822–1826.

[99] O. Johnson and C. Vignat, Some results concerning maximum Rényi entropy distributions, Ann. I.H.P. Probab. Stat. **43**, 339 (2007).

[100] W. Beckner, Inequalities in Fourier analysis, Ann. Math. **102**, 159 (1975).

[101] I. Białynicki-Birula and Ł. Rudnicki, Entropic uncertainty relations in quantum physics, in *Statistical Complexity*, edited by K. D. Sen (Springer, Dordrecht, 2011).

[102] R. M. Gomes, A. Salles, F. Toscano, P. H. S. Ribeiro, and S. P. Walborn, Quantum entanglement beyond Gaussian criteria, Proc. Natl. Acad. Sci. USA **106**, 21517 (2009).

[103] R. M. Gomes, A. Salles, F. Toscano, P. H. Souto Ribeiro, and S. P. Walborn, Observation of a nonlocal optical vortex, Phys. Rev. Lett. **103**, 033602 (2009).

[104] W. A. T. Nogueira, S. P. Walborn, S. Pádua, and C. H. Monken, Generation of a two-photon singlet beam, Phys. Rev. Lett. **92**, 043602 (2004).

[105] Ł. Rudnicki, S. P. Walborn, and F. Toscano, Heisenberg uncertainty relation for coarse-grained observables, Europhys. Lett. **97**, 38003 (2012).

[106] D. S. Tasca, Ł. Rudnicki, R. M. Gomes, F. Toscano, and S. P. Walborn, Reliable Entanglement Detection under Coarse-Grained Measurements, Phys. Rev. Lett. **110**, 210502 (2013).

[107] M. C. Jones, J. S. Marron, and S. J. Sheather, A brief survey of bandwidth selection for density estimation, J. Am. Stat. Assoc. **91**, 401 (1996).

[108] N. Armstrong, G. J. Sutton, and D. B. Hibbert, Estimating probability density functions using a combined maximum entropy moments and Bayesian method. theory and numerical examples, Metrologia **56**, 015019 (2019).

[109] B. Uria, M.-A. Côté, K. Gregor, I. Murray, and H. Larochelle, Neural Autoregressive Distribution Estimation, J. Mach. Learn. Res. **17**, 1 (2016).

[110] S. E. Reed, A. van den Oord, N. Kalchbrenner, S. G. Colmenarejo, Z. Wang, D. Belov, and N. de Freitas, Parallel Multiscale Autoregressive Density Estimation, in *Proceedings of the 34th International Conference on Machine Learning*, edited by D. Precup and Y. W. Teh (2017), Vol. 70, pp. 2912–2921.

[111] L. Paninski, Estimation of entropy and mutual information, Neural Comput. **15**, 1191 (2003).

[112] A. Kraskov, H. Stögbauer, and P. Grassberger, Estimating mutual information, Phys. Rev. E **69**, 066138 (2004).

[113] G. D. Palma, Uncertainty relations with quantum memory for the Wehrl entropy, Commun. Math. Phys. **108**, 2139 (2017).