Optimizing detection of continuous variable entanglement for limited data

Martin Gärttner,$^{1,2,3,*}$ Tobias Haas,$^{4,†}$ and Johannes Noll$^{3,‡}$

$^1$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
$^4$Centre for Quantum Information and Communication, École polytechnique de Bruxelles, CP 165/39, Université libre de Bruxelles, 1050 Brussels, Belgium

We explore the advantages of a class of entanglement criteria for continuous variable systems based on the Husimi $Q$-distribution in scenarios with sparse experimental data. The generality of these criteria allows optimizing them for a given entangled state and experimental setting. We consider the scenario of coarse grained measurements, or finite detector resolution, where the values of the Husimi $Q$-distribution are only known on a grid of points in phase space, and show how the entanglement criteria can be adapted to this case. Further, we examine the scenario where experimental measurements amount to drawing independent samples from the Husimi distribution. Here, we customize our entanglement criteria to maximize the statistical significance of the detection for a given finite number of samples. 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

Efficient methods for experimentally detecting entanglement are highly relevant for quantum technologies where entanglement is a necessary ingredient for achieving quantum advantage. The task of quantifying entanglement is notoriously difficult and generally requires costly tomography of the prepared quantum state [1, 2]. What is less demanding is detecting the presence of entanglement using experimentally verifiable criteria that can only be met if the prepared state is entangled [3].

Here we consider a family of entanglement criteria$^1$ for continuous variable systems derived in two companion papers [4, 5] and show that it yields practical advantages when dealing with limited experimental data. Concretely, we consider a bipartite system described by the canonical operators $R_j, S_j$ acting on subsystems $j \in \{1, 2\}$ and obeying the commutation relations $[R_j, S_k] = i \delta_{jk} \mathbb{1}$. Our entanglement criteria are based on the Husimi $Q$-distribution, defined as

$$Q(r_1, s_1, r_2, s_2) = \langle \alpha_1 | \otimes \langle \alpha_2 | \rho_{12} (\langle \alpha_1 | \otimes | \alpha_2 \rangle), \quad (1)$$

which is associated with a measurement in the coherent state basis [6–8]. Our criteria build on the intuition that fluctuations in the "non-local" operators [9–11]

$$R_\pm = a_1 R_1 \pm a_2 R_2, \quad S_\pm = b_1 S_1 \pm b_2 S_2, \quad (2)$$

with scalings $a_1, b_1, a_2, b_2 \geq 0$ such that $a_1 b_1 = a_2 b_2$, can only be small if the two subsystems are entangled. Stated quantitatively, for any concave function $f$ of the variables $R_1 + R_2$ and $S_1 - S_2$ is bounded below by the modulus of the expectation value of the commutator for separable states due to uncertainty relations [see also 13, 14] for entropic formulations. In terms of the Husimi $Q$-distribution these uncertainty relations can be formulated in a very general way using the Lieb-Solovay theorem [15, 16]. This results in the non-negativity of the witness functional

$$W_f = \int \frac{dr_+ ds_+}{2\pi} \left[ f(Q_+) - f(\bar{Q}_+) \right] \geq 0, \quad (3)$$

being fulfilled for all separable states, where $f$ can be any concave function with $f(0) = 0$ and we have defined the marginal of the Husimi distribution transformed into the non-local coordinates

$$Q_\pm = Q(r_\pm, s_\pm) = \int \frac{dr_\pm ds_\pm}{2\pi} Q(r_+, s_+, r_-, s_-). \quad (4)$$

$\bar{Q}_\pm$ is the corresponding marginal of the Husimi distribution of the vacuum state, which evaluates to

$$\bar{Q}_\pm(r_\pm, s_\pm) = \frac{1}{a_1 b_1 + a_2 b_2} e^{-\frac{\bar{r}_\pm^2 + \bar{s}_\pm^2}{2(a_1 b_1 + a_2 b_2)}}. \quad (5)$$

For $f = t^\beta$, the integrals in (3) are proportional to the Rényi-Wehrl entropy [17]

$$S_\beta(Q_\pm) = \frac{1}{1 - \beta} \ln \left[ \int \frac{dr_\pm ds_\pm}{2\pi} Q_\beta^\beta(r_\pm, s_\pm) \right] \quad (6)$$

with $\beta > 0, \beta \neq 1$, leading to the criteria

$$W_\beta = S_\beta(Q_\pm) - \frac{\ln \beta}{\beta - 1} - \frac{\ln \det \bar{V}_\beta^\beta}{2} \geq 0, \quad (7)$$

* martin.gaerttner@kip.uni-heidelberg.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.
with \( \hat{V}_\beta^f \equiv (a_1 b_1 + a_2 b_1) \mathbb{1} \), which is the covariance matrix of \( \hat{Q}_\beta^f \). As Gaussian distributions maximize entropies in the limit \( \beta \to 1 \), second moment criteria \( \text{det} V \beta \text{det} V_s = \text{det} V_\Delta - \text{det} \hat{V}_\Delta \geq 0 \) for the covariance matrix \( V_\Delta \) follow from (7). Thus, our criteria flag entanglement when measurements of localization such as the variance or an entropy of \( Q_\pm \) become small, i.e. when the distribution over a pair of non-local variables is more strongly localized, or peaked, than the corresponding marginal of the vacuum Husimi \( Q \)-distribution \( \hat{Q}_\Delta \).

On an operational level, entanglement can be detected by measuring the Husimi \( Q \)-distribution of a prepared state, calculating the marginal \( Q_\pm \) in the non-local variables with arbitrary scalings, and evaluating the witness functional \( W_f \), defined in (3) for arbitrary concave functions \( f \). Finding a choice for which \( W_f < 0 \) proves entanglement of the prepared state. In [5] we show that these criteria outperform commonly used criteria in that they flag entanglement for otherwise undetectable states.

In experiments, of course, it is 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 [18–21]. Thus, the relevant question is not whether our criterion can flag entanglement in the limit of precise knowledge of \( Q_\pm \), but 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 function \( f \).

A common scheme for measuring the Husimi \( Q \)-distribution is the application of a coherent displacement followed by the detection of the vacuum projection [22–24]. Another established way of measuring Husimi distributions is heterodyne detection [8, 32], 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 to simultaneously detect \( R_i \) and \( S_j \) with minimal uncertainty. This corresponds to direct sampling from the Husimi \( Q \)-distribution [33]. This scheme has been realized with optical photons [34–37] and more recently with ultracold atomic gases [38], including the bipartite setting [39]. 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 in section III.

Notation. We use natural units \( \hbar = 1 \), specify quantum operators with bold letters, e.g. \( \rho \), and write vacuum expressions with a bar, e.g. \( \bar{Q} \).

II. COARSE-GRAINED MEASUREMENTS

We start with discussing the influence of finite resolution on the entanglement criteria (3). 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 post-measurement 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 [40].

To achieve this generality, we discretize phase space into compact \( \delta_{jk} \), 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. \( \{j, k\} \in \mathbb{Z} \subseteq \mathbb{Z} \times \mathbb{Z} \). To every tile \( \delta_{jk} \) we associate its phase space measure, which we denote by \( \Delta_{jk} \), and discrete coordinates \( (r_{\pm}^j, s_{\pm}^k) \), which are located at the center of \( \delta_{jk} \) if it corresponds to a simply connected region.

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

\[
r_{\pm} \to r_{\pm}^j = j \delta r_{\pm}, \quad s_{\pm} \to s_{\mp}^k = k \delta s_{\mp},
\]

including the possibility that \( \delta r_{\pm} \neq \delta s_{\mp} \). Then, the tile \( \delta_{jk} \) is centered at \( (r_{\pm}^j, s_{\mp}^k) \) and hence its domain is given by \( \delta_{jk} = \left[ \delta r_{\pm} (j - 1/2), \delta r_{\pm} (j + 1/2) \right] \times \left[ \delta s_{\mp} (j - 1/2), \delta s_{\mp} (j + 1/2) \right] \) 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 Figure 1. A regular quadratic tiling with \( \delta = \delta_{\pm} = \delta_{\mp} \) and \( \Delta = \delta^2 / (2\pi) \), most relevant in the context of finite resolution measurements in quantum optics, is shown in
which can be advantageous especially for non-Gaussian distributions. Another adaptive scheme, which is based on radially symmetric tilings, is shown in Figure 1\textbf{c}). Here, the discrete coordinates may be labeled with only one non-negative integer-valued index, e.g. $j \in \mathbb{Z}^+_0$.

\section*{B. Discretized distributions}

The discretization procedure leads to a discrete distribution over the coordinate points $(r^j_{\pm}, s^k_{\mp})$ by integrating the continuous distribution $Q_{\pm}$ over the corresponding tile $\delta_{jk}$, i.e.

$$Q^{jk}_{\pm} \equiv Q_{\pm}(r^j_{\pm}, s^k_{\mp}) = \int_{\delta_{jk}} \frac{dr_{\pm} ds_{\mp}}{2\pi} Q_{\pm}(r_{\pm}, s_{\mp}).$$

(9)

This distribution is normalized to unity according to

$$\sum_{j, k \in \mathbb{Z}} Q^{jk}_{\pm} = 1.$$  

(10)

Uncertainty relations and entanglement criteria for coarse-grained measurements have been studied extensively in the literature, see e.g. [41–44]. 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 $Q^{jk}_{\pm}$ as they often underestimate their continuous analogs. For example, consider an experimental procedure with sufficiently large 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 positive.

Therefore, we work with the density of the discrete distribution (9) over every tile instead. It is defined via

$$Q^\Delta_{\pm} \equiv Q^\Delta_{\pm}(r_{\pm}, s_{\mp}) = \sum_{j, k \in \mathbb{Z}} \left\{ \begin{array}{ll} Q^{jk}_{\pm}, & (r_{\pm}, s_{\mp}) \in \delta_{jk}, \\ 0, & \text{else,} \end{array} \right.$$  

(11)

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. $Q^\Delta_{\pm}(r_{\pm}, s_{\mp}) \rightarrow Q_{\pm}(r_{\pm}, s_{\mp})$ for $\Delta_{jk} \rightarrow 0$.

As an example, we consider the effect of coarse-graining for the Husimi $Q$-distribution $Q_{\pm}$ of the two-mode squeezed vacuum state (TMSV) state

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

(12)

with $\lambda \in [0, 1]$ being the squeezing parameter, which reads

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

(13)

with normalization $Z = \det V_{\pm}^{1/2}$ and covariance matrix $V_{\pm} = \frac{1}{4} 1$. We show the case $a_1 = b_1 = a_2 = b_2 = 1, \vartheta_1 = \vartheta_2 = 0$ and $\lambda = 0.1$ in Figure 2 for a regular quadratic tiling with various spacing $\delta$ along both axes as shown in Figure 1\textbf{a}).

\section*{C. Discretized criteria}

\subsection*{1. General criteria}

In order to derive entanglement criteria which hold for the discretized approximation $Q^\Delta_{\pm}$, we apply Jensen’s inequality to the integrals appearing in the witness functional (3) for every tile individually. Adjusted to this
setup, i.e. the continuous Husimi $Q$-distribution $Q_{\pm}$ restricted to $\delta_{jk}$ with measure $\Delta_{jk}$ and concave $f$, Jensen’s inequality reads
\[
\frac{1}{\Delta_{jk}} \int_{\delta_{jk}} \frac{df}{2\pi} \frac{ds_{\pm}}{2\pi} f(Q_{\pm}) \leq f \left( \frac{1}{\Delta_{jk}} \int_{\delta_{jk}} \frac{df}{2\pi} \frac{ds_{\pm}}{2\pi} Q_{\pm} \right). \tag{14}
\]
Now, discretizing the first term in our witness functional $W_f$ (3) by expanding the phase space integral over all tiles, inserting a multiplicative one and using Jensen’s inequality (14) together with the definition of the discrete distribution (9) leads to
\[
\int \frac{dr_{\pm} ds_{\pm}}{2\pi} f(Q_{\pm}) = \sum_{j,k \in \mathbb{Z}} \Delta_{jk} \frac{1}{\Delta_{jk}} \int_{\delta_{jk}} \frac{dr_{\pm} ds_{\pm}}{2\pi} f(Q_{\pm}) \leq \sum_{j,k \in \mathbb{Z}} \Delta_{jk} f \left( \frac{1}{\Delta_{jk}} Q_{\pm}^{jk} \right). \tag{15}
\]
Now, the right hand side of the latter inequality can be rewritten in terms of the continuous approximation
\[
\sum_{j,k \in \mathbb{Z}} \Delta_{jk} f \left( \frac{1}{\Delta_{jk}} Q_{\pm}^{jk} \right) = \int \frac{dr_{\pm} ds_{\pm}}{2\pi} f(Q_{\pm}), \tag{16}
\]
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 $Q_{\pm}$. Defining a discretized witness functional in terms of $Q_{\pm}^{jk}$ instead of $Q_{\pm}$
\[
W_f^{\Delta} = \int \frac{dr_{\pm} ds_{\pm}}{2\pi} \left[ f(Q_{\pm}) - f(Q_{\pm}') \right], \tag{17}
\]
allows to conclude $W_f \leq W_f^{\Delta}$ and therefore our criteria (3) imply that all separable states fulfill the discretized criteria
\[
\rho_{12} \text{ separable } \Rightarrow W_f^{\Delta} \geq 0, \tag{18}
\]
which are weaker in general. We emphasize that the latter inequality is fulfilled for arbitrary discretization schemes. Since the discretized witness functional $W_f^{\Delta}$ is of the same form as its continuous counterpart $W_f$, we can follow the same arguments as in [4, 5] to derive interesting classes of discrete entanglement criteria. As it is often more convenient to work with the discrete distribution $Q_{\pm}^{jk}$ instead of $Q_{\pm}$ we will also give relations in terms of the former whenever possible.

2. Entropic criteria

Similar to (6), we obtain criteria for Rényi-Wehrl entropies of the discretized approximations
\[
S_\beta(Q_{\pm}^{jk}) = \frac{1}{1 - \beta} \ln \left[ \int \frac{dr_{\pm} ds_{\pm}}{2\pi} \left( Q_{\pm}^{jk} \right)^\beta (r_{\pm}, s_{\pm}) \right] \tag{19}
\]
with entropic orders $\beta \in (0, 1) \cup (1, \infty)$, by choosing monomials $f(t) = t^\beta$ and applying a monotonic function in (17), which yields the witness functional
\[
W_\beta^{\Delta} = S_\beta(Q_{\pm}^{jk}) = \frac{1}{1 - \beta} \ln \frac{\beta}{\beta - 1} - \ln \det V_{\pm}^{jk}, \tag{20}
\]
apalogous to (7). In general, we can write the Rényi-Wehrl entropies (19) in terms of the discretized distribution $Q_{\pm}^{jk}$ as
\[
S_\beta(Q_{\pm}^{jk}) = \frac{1}{1 - \beta} \ln \left[ \sum_{j,k \in \mathbb{Z}} \Delta_{jk}^{1 - \beta} \left( Q_{\pm}^{jk} \right)^\beta \right]. \tag{21}
\]
The latter can only be simplified further for regular tilings with $\Delta \equiv \Delta_{jk}$, leading to
\[
S_\beta(Q_{\pm}^{jk}) = S_\beta(Q_{\pm}^{jk}) + \ln \Delta, \tag{22}
\]
with discrete Rényi-Wehrl entropies
\[
S_\beta(Q_{\pm}^{jk}) = \frac{1}{1 - \beta} \ln \left[ \sum_{j,k \in \mathbb{Z}} \left( Q_{\pm}^{jk} \right)^\beta \right]. \tag{23}
\]
In the limit $\beta \to 1$, the Rényi-Wehrl entropy (19) converges to the Wehrl entropy [45, 46]

$$S_1(Q^\Delta_\pm) = -\int \frac{dr_\pm ds_\pm}{2\pi} Q^\Delta_\pm(r_\pm, s_\pm) \ln Q^\Delta_\pm(r_\pm, s_\pm),$$

for which the discretized witness functional reads

$$W^\Delta_\pm = S_1(Q^\Delta_\pm) - 1 - \ln \det V^\Delta_\pm.$$  (24)

Note that the continuous analog of (25) has been derived in [47] for $a_1 = b_1 = a_2 = b_2 = 1$. Irrespective of the discretization scheme we find the relation

$$S_1(Q^\Delta_\pm) = S_1(Q^{jk}_\pm) + \sum_{j,k \in \mathbb{Z}} \ln(\Delta_{jk}) Q^{jk}_\pm$$  (26)

with the discrete Wehrl entropy being defined as

$$S_1(Q^{jk}_\pm) = -\sum_{j,k \in \mathbb{Z}} Q^{jk}_\pm \ln Q^{jk}_\pm,$$  (27)

which also follows from (23) in the limit $\beta \to 1$. For regular tilings $\Delta = \Delta_{jk}$ (26) reduces to the simple relation

$$S_1(Q^\Delta_\pm) = S_1(Q^{jk}_\pm) + \ln \Delta.$$  (28)

### 3. Second moment criteria

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

$$\left(\Sigma^\Delta_{r_\pm}\right)^2 = \int \frac{dr_\pm ds_\pm}{2\pi} r_\pm^2 Q^\Delta_\pm(r_\pm, s_\pm),$$
$$\left(\Sigma^\Delta_{s_\pm}\right)^2 = \int \frac{dr_\pm ds_\pm}{2\pi} s_\pm^2 Q^\Delta_\pm(r_\pm, s_\pm),$$
$$\Sigma^\Delta_{r_\pm s_\pm} = \int \frac{dr_\pm ds_\pm}{2\pi} r_\pm s_\pm Q^\Delta_\pm(r_\pm, s_\pm),$$

can be formulated in terms of the discretized covariance matrix

$$V^\Delta_\pm = \begin{pmatrix} \left(\Sigma^\Delta_{r\pm}\right)^2 & \Sigma^\Delta_{r\pm s\pm} \\ \Sigma^\Delta_{s\pm s\pm} & \left(\Sigma^\Delta_{s\pm}\right)^2 \end{pmatrix},$$  (30)

by using that $S_1(Q^\Delta_\pm) \leq 1 + \frac{1}{2} \ln \det V^\Delta_\pm$, which leads to the second moment witness functional

$$W^\Delta_\pm = \det V^\Delta_\pm - \det \bar{V}^\Delta_\pm.$$  (31)

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

$$\left(\mu^\Delta_{r\pm}, \mu^\Delta_{s\pm}\right) = \left(\mu^{jk}_{r\pm}, \mu^{jk}_{s\pm}\right),$$  (32)

with the continuous means given by

$$\mu^\Delta_{r\pm} = \int \frac{dr_\pm ds_\pm}{2\pi} r_\pm Q^\Delta_\pm(r_\pm, s_\pm),$$
$$\mu^\Delta_{s\pm} = \int \frac{dr_\pm ds_\pm}{2\pi} s_\pm Q^\Delta_\pm(r_\pm, s_\pm),$$  (33)

while the discrete means read

$$\mu^{jk}_{r\pm} = \sum_{j,k \in \mathbb{Z}} r_{jk}^j Q^{jk}_\pm,$$
$$\mu^{jk}_{s\pm} = \sum_{j,k \in \mathbb{Z}} s_{jk}^k Q^{jk}_\pm.$$  (34)

In contrast, the second moments of $Q^\Delta_\pm$ and $Q^{jk}_\pm$ differ. For regular tilings with $\Delta = \Delta_{jk}$ (26) reduces to the simple relation

$$\left(\Sigma^\Delta_{r_{\pm}}\right)^2 = \Sigma^2_{r_{\pm}} + \frac{(\delta r_{\pm})^2}{12},$$
$$\left(\Sigma^\Delta_{s_{\pm}}\right)^2 = \Sigma^2_{s_{\pm}} + \frac{(\delta s_{\pm})^2}{12},$$
$$\Sigma^\Delta_{r_{\pm} s_{\pm}} = \Sigma_{r_{\pm} s_{\pm}}^{jk},$$  (35)

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

$$\Sigma^2_{r_{\pm}} = \sum_{j,k \in \mathbb{Z}} \left(r_{jk}^j\right)^2 Q^{jk}_\pm,$$
$$\Sigma^2_{s_{\pm}} = \sum_{j,k \in \mathbb{Z}} \left(s_{jk}^k\right)^2 Q^{jk}_\pm,$$
$$\Sigma_{r_{\pm} s_{\pm}}^{jk} = \sum_{j,k \in \mathbb{Z}} r_{jk}^j s_{jk}^k Q^{jk}_\pm.$$  (36)

Eq. (35) shows that the discrete second moments $\Sigma^2_{r_{\pm}}$ and $\Sigma^2_{s_{\pm}}$ indeed underestimate the continuous variances $(\Sigma^\Delta_{r_{\pm}})^2$ and $(\Sigma^\Delta_{s_{\pm}})^2$, 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 (18) we consider again the Gaussian TMSV state (12) for $a_1 = b_1 = a_2 = b_2 = 1$ and $\delta_1 = \delta_2 = 0$. The corresponding distribution $Q_\pm$ is discretized following a regular quadratic tiling with grid spacing $\delta$ shown in Figure 1 a), leading to a discretized distribution $Q^\Delta_\pm$, which is illustrated for various $\delta$ and $\lambda = 0.1$ in Figure 2. When statistical errors become negligible, which is a justified assumption for many quantum optics setups, the discretized distribution $Q^\Delta_\pm$ can be measured directly.

Most importantly, the discretization breaks Gaussianity and hence the choice of the function $f$ matters. In Figure 3 a) we compare the performances of the discretized
versions of our second moment criteria (gray), the Wehrl entropic criteria (black) and the optimized Rényi-Wehrl entropic criteria (light orange) as functions of the grid spacing $\delta$. In general, we can say that entropic criteria strongly outperform second moment criteria and that an optimization over the order $\beta$ for every $\lambda$ 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 $\beta$ in c) and f), respectively, with the black dotted line indicating where the witness evaluates to zero. Note here that the witness functional becomes independent of $\beta$ overall (dark orange) and the optimal $\beta$ for every $\delta$ (light orange). The values of the four witnesses are shown in b) and e) as functions of the grid spacing $\delta$ for the same choices of $\lambda$. The optimal choice for $\beta$ depends on $\lambda$, as plotted in d), showing that for small entanglement $\lambda \leq 0.5$ one should take $\beta \geq 1$ and vice versa.

Further, the optimal choice for $\beta$ is a monotonically decreasing function of $\lambda$, which we plot in d). Roughly speaking, for large entanglement $\lambda \geq 0.5$ we need small $\beta \leq 1$, while for small entanglement $\lambda \leq 0.5$ we should choose large $\beta \geq 1$.

### III. 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 $W_f$ (3), 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 non-parametric approaches like simple data binning or kernel density estimation [48], maximum entropy models [49] and parametric deep-learning based approaches [50, 51].

For the specific case of entropies, studied here, also direct estimation techniques have been devised (see [52] for a review) including the popular $k$-nearest neighbor
method [53]. Generally, these methods rely on assumptions about the smoothness of the underlying PDF which allow to bound the approximation error. In this respect, the Husimi Q-distribution has favorable properties due to the quantum uncertainty principle which leads to an inherently smooth behavior. In contrast to the Wigner W-distribution, which can have arbitrarily sharp features in the non-local phase space, the simultaneous detection of two conjugate quadratures leads to a coarse graining in the non-local phase space, the simultaneous detection of two conjugate quadratures leads to a coarse graining or smoothing of the phase space 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 $W_\beta$ (7). Using moderate sample set sizes $O(10^3)$ we are able to reliably estimate the witness functional. Crucially, by using different Rényi-Wehrl entropies parameterized by $\beta$ we find that values of small or large $\beta$ lead to a significantly increased signal-to-noise ratio of the entanglement detection compared to 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 $\lambda$, opposite displacements $\pm r$ with $r \geq 0$ along the $r_{\pm}$ axis and a mixing probability $p \in [0,1]$ s.t. $p = 0$ selects the state displaced towards positive values for $r_{\pm}$. Its Husimi Q-distribution over a pair of non-local variables is

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

(37)

Note that this setup is similar to the non-Gaussian state investigated in [4, 5, 13, 14].

We test the performance of the Rényi-Wehrl criteria (7) for the state 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 $\beta$-dependence of the mean value and the three $\sigma$-intervals is shown in Figure 4 a). As the mean value (black dashed) aligns with the exact result (gray), our estimation method is justified a posteriori. We observe that the choice of $\beta$ has a strong influence on the signal-to-noise ratio, which we exemplify for $\beta \to 1$ and $\beta = 10$ in b) and 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 $\beta = 10$. In the latter case entanglement is even certified within $\pm 3\sigma$ in 88% of all cases.

**IV. CONCLUSIONS AND OUTLOOK**

We have shown that generalized entanglement criteria based on the Husimi Q-distribution lead to clear performance advantages 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 $W_f$ over a parameterized family of concave functions $f$ — in our case monomials $t^3$, relating the witness functional to Rényi-Wehrl entropies — greatly enlarges the range of measurement resolution in which entanglement is detected compared to second moment based or Wehrl entropic criteria. In a second scenario, where experimental measurements correspond to drawing samples from the Husimi

Figure 4. a) Mean and confidence intervals of the Rényi-Wehrl criteria (7) for $10^3$ samples of the state (37) after $10^2$ repetitions of this simulated experiment. b) and c) show the computed values for all repetitions for $\beta \to 1$ and $\beta = 10$, respectively. Optimizing over $\beta$, i.e. for $\beta$ on the order of $10^3$, certification of entanglement within large confidence intervals is improved substantially.
While one may expect that estimating functionals of a $L.\text{-M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, Inseparable Wigner to approaches based on second moments or marginals of leads to advantages for entanglement detection compared to rather small sample sizes. Given that the Husimi standard density estimation methods work reliably even turns out that in the case of the Husimi PDF is a much harder task than estimating low moments, Q-criteria, but also the inherent smoothness of the Husimi density estimation routine making justified assumptions for small thus incur a large statistical error. The good performance for small $\beta$, where the contribution of the tail regions to $W_f$ are amplified, might be explained by the employed density estimation routine based on second moments or marginals of the Wigner $W$-distribution (see also [5]).

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 [49]. Moreover, for specific choices of $f$ direct ways of extracting $f(Q_\beta)$, or its phase space integral [52, 53], 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.

ACKNOWLEDGEMENTS

We thank Oliver Stockdale, Stefan Floerchinger and Markus Oberthaler for discussions on the subject. M.G. is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy EXC 2181/1 - 390900948 (the Heidelberg STRUCTURES Excellence Cluster) and under SFB 1225 ISOQUANT - 273811115. T.H. acknowledges support by the European Union under project ShoQC within ERA-NET Cofund in Quantum Technologies (QuantERA) program.

[1] M. B. Plenio and S. Virmani, An Introduction to Entanglement Measures, Quantum Inf. Comput. 7, 1–51 (2007).
[2] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, Efficient quantum state tomography, Nat. Comm. 1, 149 (2019).
[3] O. Gühne and G. Tóth, Entanglement detection, Phys. Rep. 474, 1 (2009).
[4] M. Gärttner, T. Haas, and J. Noll, General class of continuous variable entanglement criteria, Companion paper appearing in the same arXiv posting (2022).
[5] M. Gärttner, T. Haas, and J. Noll, Detecting continuous variable entanglement in phase space with the $Q$-distribution, Companion paper appearing in the same arXiv posting (2022).
[6] K. Husimi, Some formal properties of the density matrix, Proc. Phys.-Math. Soc. Jap. 3rd Ser. 22, 264 (1940).
[7] W. P. Schleich, Quantum Optics in Phase Space (Wiley-VCH Verlag Berlin, 2001).
[8] L. Mandel and E. Wolf, Optical Coherence and Quantum Optics (Cambridge University Press, 2013).
[9] A. Einstein, B. Podolsky, and N. Rosen, Can Quantum-Mechan-ical Description of Physical Reality Be Considered Complete?, Phys. Rev. 47, 777 (1935).
[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] V. Giovannetti, S. Mancini, D. Vitali, and P. Tombesi, Characterizing the entanglement of bipartite quantum systems, Phys. Rev. A 67, 022320 (2003).
[12] S. Mancini, V. Giovannetti, D. Vitali, and P. Tombesi, Entangling Macroscopic Oscillators Exploiting Radiation Pressure, Phys. Rev. Lett. 88, 120401 (2002).
[13] 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).
[14] A. Saboia, F. Toscano, and S. P. Walborn, Family of continuous-variable entanglement criteria using general entropy functions, Phys. Rev. A 83, 032307 (2011).
[15] 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).
[16] P. Schupp, Wehrl entropy, coherent states and quantum channels, in The Physics and Mathematics of Elliott Lieb (EMS Press, 2022) pp. 329–344.
[17] I. Bengtsson and K. Zyczkowski, Geometry of Quantum States, 2nd ed. (John Wiley and Sons, 2017).
[18] U. Leonhardt and H. Paul, Measuring the quantum state of light, Prog. Quantum. Electron. 19, 89 (1995).
[19] S. L. Braunstein and P. van Loock, Quantum information with continuous variables, Rev. Mod. Phys. 77, 513 (2005).
[20] 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).
[21] A. Serafini, Quantum Continuous Variables (CRC Press, 2017).
[22] T. Opatrný and D.-G. Welsch, Density-matrix reconstruction by unbalanced homodyning, Phys. Rev. A 55, 1462 (1997).
